การระบุและลักษณะสมบัติของยีนที่มีหน้าที่เกี่ยวกับการสืบพันธุ์ของกุ้งกุลาดำ

Penaeus monodon

นางสาว รุ่งนภา ลีละธนาวิทย์

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาเทคโนโลยีชีวภาพ

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2551

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

IDENTIFICATION AND CHARACTERIZATION OF GENES FUNCTIONALLY RELATED TO REPRODUCTION OF THE GIANT TIGER SHRIMP *Penaeus monodon*

Miss Rungnapa Leelatanawit

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy Program in Biotechnology

Faculty of Science

Chulalongkorn University

Academic Year 2008

Copyright of Chulalongkorn University

Thesis Title	Identification and characterization of genes functionally related to reproduction of the giant tiger shrimp *Penaeus monodon*
By	Miss Rungnapa Leelatanawit
Field of study	Biotechnology
Thesis Principal Advisor	Professor Piamsak Menasveta, Ph.D.
Thesis Co-advisor	Sirawut Klinbunga, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

Deputy Dean for Administrative Affairs,
.... Acting Dean, The Faculty of Science

(Associate Professor Vimolvan Pimpan, Ph.D.)

THESIS COMMITTEE

.... Chairman

(Associate Professor Charoen Nitithamyong, Ph.D.)

.... Thesis Principal Advisor

(Professor Piamsak Menasveta, Ph.D.)

.... Thesis Co-advisor

(Sirawut Klinbunga, Ph.D.)

.... External Member

(Sittiruk Roytrakul, Ph.D.)

.... Member

(Associate Professor Thaithaworn Lirdwitayaprasit, Ph.D.)

.... Member

(Associate Professor Warawut Chulalaksananukul, Ph.D.)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

รุ่งนภา ลีละธนาวิทย์ : การระบุและลักษณะสมบัติของยีนที่มีหน้าที่เกี่ยวกับการสืบพันธุ์ของกุ้งกุลาดำ *Penaeus monodon* (IDENTIFICATION AND CHARACTERIZATION OF GENES FUNCTIONALLY RELATED TO REPRODUCTION OF THE GIANT TIGER SHRIMP *Penaeus monodon*) อ. ที่ปรึกษาวิทยานิพนธ์หลัก : ศ.ดร. เปี่ยมศักดิ์ เมนะเศวต, อ. ที่ปรึกษาวิทยานิพนธ์ร่วม : ดร. ศิราวุธ กลิ่นบุหงา, 307 หน้า.

กุ้งกุลาดำจะมีความสมบูรณ์พันธุ์ต่ำเมื่ออยู่ในภาวะการเลี้ยง ซึ่งเป็นอุปสรรคสำคัญต่อการปรับปรุงพันธุ์ที่มีประสิทธิภาพของสัตว์เศรษฐกิจชนิดนี้ ดังนั้นองค์ความรู้เกี่ยวกับกลไกการพัฒนารังไข่และอัณฑะของกุ้งกุลาดำในระดับโมเลกุล จึงมีความสำคัญและสามารถนำไปประยุกต์ใช้ในอุตสาหกรรมการเลี้ยงกุ้ง ในการศึกษานี้จึงสืบค้นยีนที่เกี่ยวข้องกับการพัฒนาอัณฑะของกุ้งกุลาดำ โดยทำการหาลำดับนิวคลีโอไทด์ของอีเอสทีจากห้องสมุดยีนอัณฑะ จำนวนทั้งหมด 896 โดยพบว่าอีเอสทีจำนวน 606 โคลน (67.6%) มีลำดับนิวคลีโอไทด์ที่เหมือนกับยีนในฐานข้อมูล GenBank (E-value < 1e-04) นอกจากนี้ได้ทำการสร้างห้องสมุดยีนโดยวิธี suppression subtractive hybridization (SSH) ระหว่าง cDNA ของอัณฑะของกุ้งกุลาดำขนาดขนาดพ่อแม่พันธุ์และกุ้งวัยรุ่นอายุ 4 เดือน และทำการหาลำดับนิวคลีโอไทด์ของ 178 และ 187 โคลนที่ได้จากห้องสมุด forward และ reverse SSH พบอีเอสทีที่มีลำดับนิวคลีโอไทด์เหมือนกับยีนใน GenBank จำนวน 67 โคลน (37.1%) และ 104 โคลน (54.0%) ตามลำดับ โดยพบยีนที่มีหน้าที่สำคัญเกี่ยวกับการพัฒนาการของอัณฑะ เช่น *small ubiquitin-like modifier 1 (SUMO-1)*, *cyclophilin A* และ *dynactin subunit 5* นอกจากนี้ยังพบยีน *transformer-2* (*Tra-2*) ซึ่งเป็นยีนที่เกี่ยวข้องกับการกำหนดเพศอีกด้วย

จากการหาลำดับนิวคลีโอไทด์ที่สมบูรณ์ของยีนต่าง ๆ จำนวน 16 ยีน ด้วยวิธี RACE-PCR พบว่าสามารถหาลำดับนิวคลีโอไทด์ที่สมบูรณ์ของยีนได้จำนวน 16 ยีน ประกอบด้วย *PMTST1*, *multiple inositol polyphosphate phosphatase 2 (MIPP2)*, *prohibitin-2*, *cell division kinase 7 (cdk7)*, *flotillin 2*, *growth factor receptor-bound protein*, *innexin 1*, *innexin 2*, *Rac-GTPase activating protein 1*, *transformer 2 (Tra-2)*, *Dmc1*, *progestin membrane receptor component 1 (PGMRC1)*, *saposin*, *troponin T isoform 3*, *Ero1L CG1333-PB isoform B*, และ *dihydrolipoamide dehydrogenase*

ทำการศึกษารูปแบบการแสดงออกของยีนในอัณฑะและรังไข่ของกุ้งกุลาดำขนาดอายุ 4 เดือนและขนาดพ่อแม่พันธุ์ (N = 4 ของแต่ละกลุ่ม) จำนวน 59 ยีนด้วยเทคนิค RT-PCR พบว่า *PMTST1* มีการแสดงออกเฉพาะในอัณฑะแต่ไม่แสดงออกในรังไข่ของกุ้งกุลาดำ และพบว่า *MIPP*, *MIPP2*, *Dmc1* และ *HSP70-2* มีแนวโน้มการแสดงออกในอัณฑะมากกว่าในรังไข่ของกุ้งกุลาดำ นอกจากนี้พบยีนที่มีแนวโน้มการแสดงออกในรังไข่มากกว่าในอัณฑะจำนวนทั้งหมด 36 ยีน

ตรวจสอบการแสดงออกของยีนจำนวน 12 ยีน ด้วยวิธี semiquatitative RT-PCR หรือ quantitative real-time PCR ในกลุ่มตัวอย่างต่างๆ พบว่า *PMTST1* แสดงออกเฉพาะในอัณฑะเท่านั้น โดยพบระดับการแสดงออกของ *Dmc1*, *saposin*, *spermatogonial stem-cell renewal factor*, *MIPP2* และ *HSP70-2* ในอัณฑะสูงกว่าในรังไข่ ($P < 0.05$) ในขณะที่ระดับการแสดงออกของ *CYA* และ *Trap240* ในรังไข่สูงกว่าในอัณฑะ ($P < 0.05$) โดย *SUMO-1*, *Tra-2* และ *prohibitin2* มีการแสดงออกที่ไม่แตกต่างกันในอัณฑะและรังไข่ของกุ้งกุลาดำ ทั้งนี้ระดับการแสดงออกที่ลดลงของ *SUMO-1*, *Dmc1*, และ *spermatogonial stem-cell renewal factor* และ การแสดงออกที่เพิ่มขึ้นของ *prohibitin2* ในกุ้งขนาดพ่อพันธุ์ที่ทำการคัดพันธุ์ อาจใช้เป็นเครื่องหมายโมเลกุลเพื่อบ่งบอกระดับการลดลงของความสมบูรณ์เพศในกุ้งกุลาดำขนาดพ่อแม่พันธุ์ที่ทำการเพาะเลี้ยงในโรงเพาะเลี้ยง นอกจากนี้ยังพบว่า Dopamine ที่ปริมาณ 10^{-6} โมล/ตัว ทำให้ระดับการแสดงออกของ *PGMRC1* เพิ่มขึ้นที่ 3 ชั่วโมงหลังจากการฉีด ($P < 0.05$) แต่ไม่มีผลต่อการแสดงออกของ *Dmc1* ($P > 0.05$)

นอกจากนี้ยังทำการสร้างโปรตีนลูกผสมของยีน *Dmc1*, *spermatogonial stem-cell renewal factor*, *SUMO-1* และ *CYA* ใน *E.coli* นำโปรตีนลูกผสมของ Dmc1 spermatogonial stem-cell renewal factor และ CYA ที่ทำบริสุทธิ์ไปผลิต polyclonal antibody ในกระต่าย เพื่อใช้ตรวจสอบระดับการแสดงออกในระดับการแปลรหัสและหน้าที่ของโปรตีนดังกล่าวต่อไป

สาขาวิชา.......เทคโนโลยีชีวภาพ......... ลายมือชื่อนิสิต.......รุ่งนภา ลีละธนาวิทย์..........

ปีการศึกษา.........2551.................... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก..........................

ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม..............................

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

4773831023 : MAJOR BIOTECHNOLOGY

KEY WORD : *Penaeus monodon* / GIANT TIGER SHRIMP / TESTICULAR DEVELOPMENT / cDNA LIBRARY / RT-PCR

RUNGNAPA LEELATANAWIT: IDENTIFICATION AND CHARACTERIZATION OF GENES FUNCTIONALLY RELATED TO REPRODUCTION OF THE GIANT TIGER SHRIMP *Penaeus monodon*. THESIS PRINCIPAL ADVISOR: PROF. PIAMSAK MENASVETA, Ph.D. THESIS COADVISOR: SIRAWUT KLINBUNGA, Ph.D., 307 pp.

Low degrees of reproductive maturation of the giant tiger shrimp (*Penaeus monodon*) in captivity have limited the ability to genetically improve this important species effectively. Therefore, mechanisms governing gonadal maturation of *P. monodon* at the molecular level are important and can be directly applied to the shrimp industry. Genes expressed in testes of *P. monodon* were identified and characterized by EST analysis. A total of 896 clones from the conventional testis cDNA library were sequenced and 606 ESTs (67.6%) significantly matched sequences in the GenBank (E-value < 1e-04). In addition, 178 clones from the forward and 187 clones from the reverse SSH libraries between cDNA in testes of broodstock and juvenile *P. monodon* were also constructed and sequenced. Of which, 67 ESTs (37.1%) and 104 ESTs (54.0%) significantly matched known genes. Several genes functionally involved in testicular development were found such as *small ubiquitin-like modifier* (*SUMO-1*), *cyclophilin A* and *dynactin subunit 5*. In addition, *transformer-2* (*Tra-2*), a gene involving sex determination cascades, was also found.

Apart from the full length cDNA that found in the established libraries, additional 16 functionally important gene homologues including *low molecular weight neurofilament protein X NF-L* (termed *P. monodon testis-specific transcript 1, PMTST1*), *multiple inositol polyphosphate phosphatase 2* (*MIPP2*), *prohibitin-2, cell division kinase 7 (cdk7), flotillin 2, growth factor receptor-bound protein, innexin 1, innexin 2, Rac-GTPase activating protein 1, transformer 2 (Tra-2), meiotic recombination protein DMC1/LIM15 homolog isoform 1* (*Dmc1*), *progestin membrane receptor component 1 (PGMRC1), saposin, troponin T isoform 3, Ero1L CG1333-PB isoform B,* and *dihydrolipoamide dehydrogenase* were successfully characterized by RACE-PCR.

Expression patterns of 59 gene homologues in testes and ovaries of juvenile and broodstock *P. monodon* ($N = 4$ for each group) were non-quantitatively examined by RT-PCR. *P MTST1* was only expressed in testes ($N = 8$) but not ovaries ($N = 8$) whereas *MIPP, MIPP2, Dmc1*, and *HSP70-2* exhibited a trend of preferential expression in testes of *P. monodon*. Thirty-six genes showed a trend of greater expression levels in ovaries than testes.

Semiquatitative RT-PCR and quantitative real-time PCR were carried out to examine expression levels of 12 gene homologues in different groups of shrimp. Testis-specific expression of *PMTST1* was confirmed. *CYA* and *Trap240* were more abundantly expressed in ovaries than testes ($P < 0.05$). *Dmc1, saposin, spermatogonial stem-cell renewal factor, MIPP* and *HSP70-2* were preferentially expressed in testes to ovaries ($P < 0.05$). Expression levels of *SUMO-1, Tra-2* and *prohibitin2* in ovaries and testes of *P. monodon* were not significantly different ($P > 0.05$). *PMTST1* was up-regulated but that of the remaining genes in testes of *P. monodon* broodstock was down-regulated after shrimp were molted ($P < 0.05$). Significant reduction of *SUMO-1, Dmc1*, and *spermatogonial stem-cell renewal factor* and increment of *prohibitin2* transcripts in domesticated broodstock ($P < 0.05$) suggested that these reproductively related genes may be used as biomarkers to evaluate reduced degrees of the reproductive maturation in domesticated *P. monodon*. In addition, effects of dopamine on expression of *PGMRC1* and *Dmc1* in testes of juvenile *P. monodon* (3, 6, 12, and 24 hr post injection) were examined. Dopamine administration (10^{-6} mol/shrimp) resulted in up-regulation of *PGMRC1* in testes of juvenile *P. monodon* at 3 h post treatment ($P < 0.05$) but had no effect on *Dmc1* ($P > 0.05$).

Recombinant proteins of *Dmc1, spermatogonial stem-cell renewal factor, SUMO-1*, and *CYA* were successfully expressed *in vitro*. The polyclonal antibody was produced from recombinant Dmc1, spermatogonial stem-cell renewal factor, and CYA proteins for further functional analysis of these genes at the translational level.

Field of study........Biotechnology......... Student's signature........R. Leelatanawit..........

Academic year.........2008.................. Principal advisor's signature..................

Co-advisor's signature..................

ACKNOWLEDGMENTS

I would like to express my deepest sense of gratitude to my advisor Professor Dr. Piamsak Menasveta and my co-advisor, Dr. Sirawut Klinbunga for their great help, guidance, encouragement, valuable suggestion and supports throughout my study.

My sincere gratitude is also extended to Associate Professor Dr. Charoen Nitithamyong, Assistant Professor Dr. Thaithaworn Lirdwitayaprasit, Associate Professor Dr. Warawut Chulalaksananukul, and Dr. Sittiruk Roytrakul who serve as the committee of my thesis.

I would particularly like to thank the Center of excellence for Marine Biotechnology, Faculty of Science, Chulalongkorn University and Aquatic Animal Genetics and Biotechnology Laboratory, National Center for Genetic Engineering and Biotechnology (BIOTEC), National Science and Technology Development Agency (NSTDA) for providing facilities and the Golden Jubilee PhD program, the Thailand Research Funds (TRF) for a student grant.

I would like to extend my special thank to Professor Dr. Takashi Aoki, Professor Dr. Ikuo Hirono, Laboratory of Genome Science, Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, Japan and many thanks are also expressed to Dr. Ubolsree Leartsakulpanich for kindly suggestions.

ศูนย์วิทยทรัพยากร

In addition, many thanks are also expressed to my beloced friend; Ms. Kanchana Sittikhankeaw, Ms. Natechanok Tamneamdee, and Ms. Sirikarn Prasertlux), Mister Arun Buaklin, and everyone in our laboratory for their helps and friendly assistance.

จุฬาลงกรณ์มหาวิทยาลัย

Finally, I would like to express my deepest gratitude to my parents and members of my family for their love, understanding and encouragement extended throughout my study.

CONTENTS

	Page
THAI ABSTRACT………………………………………………………..……...	iv
ENGLISH ABSTRACT………………………………………………………....	v
ACKNOWLEDGMENTS..……………………………………………..….....	vi
CONTENTS…………………………………………………………………..…	vii
LIST OF TABLES………………………………………………………….....	xiii
LIST OF FIGURES……………………………………………………..….	xvi
LIST OF ABBREVIATIONS………………………………………………..….	xxix
CHAPTER I INTRODUCTION...	1
1.1 General introduction..	1
1.2 Taxonomy of *P. monodon*...	4
1.3 The reproductive organs of marine shrimp.....................................	4
1.4 The importance for domestication of *P. monodon*............................	4
1.5 Molecular genetic approaches used in this thesis............................	7
1.5.1 PCR...	7
1.5.2 Expression Sequence Tag (EST) analysis...........................	7
1.5.3 Suppression subtractive hybridization (SSH)......................	9
1.5.4 Reverse transcription-polymerase chain reaction (RT-PCR)...	12
1.5.5 Rapid amplification of cDNA ends-polymerase chain reaction RACE-PCR...	14
1.5.6 Real-time PCR...	14
1.5.7 Microarray analysis...	17
1.6 Sex determining systems..	17
1.7 Spermatogenesis..	22
1.8 Histology of testes of penaeid shrimp...	23
1.9 Candidate genes involved with testicular development and spermatogenesis of *P. monodon*...	28
1.10 Objectives..	35
CHAPTER II MATERIALS AND METHODS...	36
2.1 Experimental animals..	36
2.2 RNA extraction..	36

Page

2.3 Measuring concentrations of extracted RNA using spectrophotometry and electrophoresis.. 37
2.4 Construction of conventional cDNA libraries from heart and testes and suppression subtractive hybridization (SSH) cDNA libraries from different stages of testes of *P. monodon*.................... 37
2.4.1 Purification of mRNA.. 37
2.4.2 Construction of the conventional testis and heart cDNA libraries of *P. monodon*.. 38
2.4.2.1 First and second strand cDNA synthesis.................... 38
2.4.2.2 Bunting the cDNA Termini....................................... 39
2.4.2.3 Ligation of *Eco*R I Adapters..................................... 39
2.4.2.4 Phosphorylation of *Eco*R I Ends............................... 40
2.4.2.5 *Xho* I digestion.. 40
2.4.2.6 Size Fractionation... 40
2.4.2.6.1 Assembling and loading the drip column......... 40
2.4.2.6.2 Collecting the sample fractions........................ 41
2.4.2.7 Processing the cDNA fraction................................... 42
2.4.2.8 Ligation of the cDNA insert...................................... 42
2.4.2.9 Packaging and tittering.. 42
2.4.2.9.1 Packaging.. 42
2.4.2.9.2 Titering of the primary library......................... 43
2.4.2.10 Amplification of the library..................................... 43
2.4.2.11 *In vivo* excision to convert the lambda library to the phagemid library... 44
2.4.2.12 Colony PCR.. 45
2.4.2.13 Plasmid DNA extraction... 45
2.4.2.14 Sequencing, EST clustering and assembly............... 46
2.4.3 Construction of suppression subtractive hybridization (*SSH*) cDNA libraries.. 46
2.4.3.1 First strand cDNA synthesis..................................... 46
2.4.3.2 Second strand cDNA synthesis and purification......... 47
2.4.3.3 *Rsa* I Digestion... 47
2.4.3.4 Adaptor Ligation.. 48
2.4.3.5 First Hybridization.. 49

Page

2.4.3.6 Second Hybridization.. 51
2.4.3.7 PCR Amplification... 51
2.4.3.8 Ligation of PCR products to the pGEM®–T easy Vector.. 52
2.4.3.9 Transformation of ligation products to *E. coli* host cells.. 53
2.4.3.10 Colony PCR of the recombinant clones................... 53
2.4.3.11 Extraction of recombinant plasmid DNA................. 53
2.4.3.12 Sequencing, EST clustering and assembly............... 54
2.5 Isolation and characterization of the full length cDNA of functionally important gene homologues of *P. monodon* using Rapid Amplification of cDNA Ends-Polymerase Chain Reaction (RACE-PCR).. 54
2.5.1 Preparation of the 5′ and 3′ RACE template........................ 54
2.5.2 Primer design.. 55
2.5.3 RACE-PCR.. 55
2.5.4 Agarose gel electrophoresis.. 59
2.5.5 Elution of DNA fragments from agarose gels..................... 59
2.5.6 Ligation of PCR products to the pGEM®–T easy vector...... 60
2.5.7 Transformation of ligation products to *E. coli* host cells...... 60
2.5.7.1 Preparation of competent cells.................................. 60
2.5.7.2 Transformation... 60
2.5.8 DNA sequencing and assembly sequence............................ 61
2.5.9 Phylogenetic analysis of *CyclophilinA, SUMO, Tra-2, PGRMC1,* and *Dmc1*... 61
2.6 Examination of expression patterns of genes related with testicular development by RT-PCR and Tissue distribution analysis.............. 61
2.6.1 Primer design.. 61
2.6.2 First strand cDNA synthesis... 62
2.6.3 RT-PCR analysis.. 62
2.6.4 Tissue distribution analysis.. 69
2.7 Semiquantitative RT-PCR of functionally important gene homologues in testes and ovaries of broodstock and juvenile *P. monodon*.. 70

Page

2.7.1 Experiment animals.. 70

2.7.2 Primers.. 70

2.7.3 Total RNA extraction and the first strand cDNA synthesis.. 70

2.7.4 Optimization of semiquantitative RT-PCR conditions......... 71

2.7.4.1 Primer concentration.. 71

2.7.4.2 $MgCl_2$ concentration.. 71

2.7.4.3 Cycle numbers.. 71

2.7.4.4 Gel electrophoresis, quantitative, and data analysis.. 71

2.8 Examination of expression levels of interesting genes in testes of *P. monodon* by real-time PCR... 72

2.8.1 Effects of dopamine on expression of genes in testes of juvenile *P. monodon*... 72

2.8.2 Expression of genes in testes in different groups of *P. monodon*.. 72

2.8.3 Primers and construction of the standard curve................... 73

2.8.4 Quantitative real-time RT-PCR.. 73

2.9 Large scale examination of gene expression of *P. monodon* by microarray analysis.. 74

2.9.1 Total RNA extraction... 74

2.9.2 Synthesis of target cDNA... 74

2.9.3 Hybridization and detection... 76

2.10 *In vitro* expression of the full length cDNA using the bacterial expression system... 76

2.10.1 Designation of primers... 76

2.10.2 Construction of recombinant plasmid in cloning and expression vectors... 77

2.10.3 Expression of recombinant proteins.................................. 78

2.10.4 Detection of recombinant proteins.................................... 78

2.10.5 Mass spectrometric analysis of recombinant proteins........ 79

2.10.6 Purification of recombinant proteins................................. 79

2.10.7 Polyclonal antibody production.. 79

Page

CHAPTER III RESULTS.. 80

3.1 Construction of conventional testis cDNA library of *P. monodon*… 80

3.2 Construction of suppression subtractive hybridization (*SSH*) cDNA libraries from testes juvenile and broodstock of *P. monodon*........... 88

3.3 Construction of conventional heart cDNA library of *P. monodon*.... 97

3.4 Isolation and characterization of the full length cDNA of functionally important gene homologues of *P. monodon*................. 102

3.5 Examination of expression patterns and tissue distribution analysis of genes related to testicular development by RT-PCR.................... 167

3.6 Semiquantitative RT-PCR of functionally important gene homologues in testes and ovaries of broodstock and juvenile *P. monodon*.. 178

3.6.1 Optimization of the primer concentration, $MgCl_2$ concentration, and cycle numbers............................ 178

3.6.2 Semiquantitative RT-PCR analysis...................................... 180

3.7 Quantitative analysis of interesting genes in testes by real-time PCR.. 188

3.7.1 Expression level of *Dmc1* and *progestin receptor membrane component 1 (PGRMC1)* in testes of juvenile *P. monodon* in response to dopamine administration.......... 188

3.7.2 Expression levels of *SUMO-1, cyclophilinA (CYA), Dmc1*, *saposin* and *spermatogonial stemcell renewal factor* at different stage of testes and ovaries of *P. monodon*............. 191

3.8 Large scale examining expression patterns of genes in testes of *P. monodon* by microarray analysis.. 198

3.8.1 Gene differentially expressed in testes of juvenile and broodstock *P. monodon*... 198

3.8.2 Gene expression of dopamine-treated testes at 3 and 6 hours post injection of juvenile *P. monodon*..................... 201

3.9 *In vitro* expression of interesting genes using the bacterial expression system.. 206

3.9.1 Construction of recombinant plasmid in cloning and expression vector.. 206

3.9.2 *In vitro* expression of recombinant protein......................... 208

Page

3.9.3 Peptide sequencing of recombinant cyclophilin A and SUMO-1 217
3.9.4 Purification of recombinant protein 217
3.9.4.1 The first trial 217
3.9.4.2 The second trial 220

CHAPTER IV DISCUSSION 226
CHAPTER V CONCLUSION 238
REFERENCES 240
APPENDEICES 250
Appendix A 251
Appendix B 284
Appendix C 293
Biography 306

LIST OF TABLES

		Page
Table 1.1	Thai Frozen Shrimp Export between 2001-2006..........................	2
Table 1.2	Comparisons between Thai frozen shrimp export between January-June 2007 and 2008..	3
Table 2.1	Sequences of the PCR-Select cDNA synthesis primer, adaptors and PCR primers..	48
Table 2.2	Ligation reactions of cDNA of ovaries (tester 1-1 and 1-2) and testes (tester 2-1 and 2-2)..	49
Table 2.3	Composition of the first hybridization reaction of each subtraction...	49
Table 2.4	Preparation of the primary PCR master mix................................	51
Table 2.5	Preparation of the secondary PCR master mix.............................	52
Table 2.6	Primer sequences for the first strand cDNA synthesis and RACE – PCR..	55
Table 2.7	Gene-Specific Primers (GSPs) used for further characterization of the full length cDNA of functionally important gene homologues in *P. monodon*..	56
Table 2.8	Internal primers used for primer walking sequencing of the full length cDNA of functionally important gene homologues in *P. monodon*...	57
Table 2.9	Composition of 5′ RACE–PCR..	58
Table 2.10	Composition of 3′ RACE–PCR..	58
Table 2.11	Sequence, melting temperature and expected product sizes from primers designed from testis cDNA library and SSH cDNA libraries of *P. monodon*..	63
Table 2.12	Sequence, melting temperature and expected product sizes from primers designed from heart cDNA library of *P. monodon*..........	69
Table 2.13	Nucleotide sequences of primers used for real-time PCR analysis of *Dmc1, PGRMLC1, SUMO-1*, *cyclophilinA*, *spermatogonial stem-cell renewal factor* and *saposin* in *P. monodon*....................	75
Table 2.14	Nucleotide sequences of primers used for *in vitro* expression of *cyclophilin A, SUMO-1, spermatogonial stem-cell renewal factor*, *PGRMC1* and *Dmc1* in *P. monodon*................................	77

Page

Table 3.1	Functional categories of ESTs significantly matched with data in GenBank of testis cDNA library...	83
Table 3.2	Functionally important gene homologues found in a conventional testis cDNA library of *P. monodon*.........................	84
Table 3.3	Functional categories of ESTs in testes of *P. monodon* identified by SSH analysis..	91
Table 3.4	Known transcripts found in the forward SSH library of testes of *P. monodon*...	93
Table 3.5	Known transcripts found in the reverse SSH library of testes of *P. monodon*...	95
Table 3.6	Functional categories of ESTs of the conventional heart cDNA library of *P. monodon* significantly matched previously deposited sequences in GenBank..	99
Table 3.7	Functionally important gene homologues from a conventional heart cDNA library of *P. monodon*...	100
Table 3.8	Summary of the full length cDNA of gene homologues found in the testes and heart cDNA libraries..	112
Table 3.9	Summary of the full lengths of gene homologues from testis of *P. monodon* using RACE-PCR...	160
Table 3.10	Summary of expression patterns of various gene homologues analyzed by RT-PCR...	167
Table 3.11	Gene homologues that specifically expressed in testes but not ovaries and those showing a treand of preferentially expressed in testes to ovaries of *P. monodon*..	169
Table 3.12	Gene homologues showing a trend of more abundant expression patterns in ovaries than testes of *P. monodon*..............................	172
Table 3.13	Gene homologues that did not exhibit differential expression in testes and ovaries of *P. monodon*..	175
Table 3.14	Gene homologues that exhibit non-specific amplification from RT-PCR analysis..	175
Table 3.15	Optimal primer and $MgCl_2$ concentrations and the number of PCR cycles for semiquantitative RT-PCR analysis of gene homologues in *P. monodon*...	179

Page

Table 3.16	A summary on the number of up- or down-regulated genes in testes of broodstock and juvenile *P. monodon* analyzed by microarrays..........	199
Table 3.17	Examples of up- or down-regulated genes in testes of broodstock in comparison with those in testes of juvenile *P. monodon*..........	200
Table 3.18	A summary on the number of up- or down-regulated genes of testes of dopamine-treated *P. monodon* at 3 hours and 6 hours post injection..........	202
Table 3.19	Examples of up- or down-regulated genes in testes of normal and dopamine-treated juveniles of *P. monodon*..........	204
Table 3.20	Summary of conditions used for purification of recombinant proteins in this study..........	223
Table 3.21	ELISA results illustrating the titers of polyclonal antibodies after rabbits were boosted by recombinant spermatogonial stem cell renewal factor, Dmc1 and cyclophilin A after rabbits were boosted 4 times..........	224
Table 3.22	ELISA results illustrating the titers of polyclonal antibodies after rabbits were boosted by recombinant spermatogonial stem cell renewal factor, Dmc1 and cyclophilin A after rabbits were boosted 5 times..........	225

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

LIST OF FIGURES

		Page
Figure 1.1	World shrimp production since 1996-2003 (Globefish)....................	1
Figure 1.2	Diagram of male, dorsal view (A) and lateral view (B) to show reproductive organ. T = testis, VD = vas deferens, and TA = terminal ampoule..	5
Figure 1.3	Sexes of juveniles and broodstok of penaeid shrimp can be externally differentiated by petasma of male (A) and thelycum of female (B)...	5
Figure 1.4	General illustration of the polymerase chain reaction (PCR) for amplification of the target DNA..	8
Figure 1.5	Overview for construction of cDNA inserts (A) and automated DNA sequencing (single-pass) of randomly selected cDNA clones (the entire process simply called EST analysis)...............................	10
Figure 1.6	Overview of the Clontech PCR-Select^{TM} procedure. The cDNA in which specific transcripts are to be found is called “tester” and the reference cDNA is called “driver”...	11
Figure 1.7	Overall concept of the RT-PCR..	13
Figure 1.8	Overview of the SMART^{TM} RACE cDNA Amplification Kit..........	15
Figure 1.9	An overall concept of the Real-time PCR procedure.......................	16
Figure 1.10	An overall concept for large scale screening of gene expression based on microarray analysis...	18
Figure 1.11	General diagram of spermatogenesis..	23
Figure 1.12	(1) Longitudinal section of a less than 1.0 gram juveniles showing the testes (Tes)/vas deferens (Vsd) junction. (2) Cross sectional orientation view of a laterally projecting testicular lobe (Tes) from a 2.0 gram juvenile...	25
Figure 1.13	(1) Longitudinal orientation section from a less than 1.0 gram juvenile showing the proximal vas deferens (Vsd). (2-4) Longitudinal orientation section of the medial vas deferens (Vsd), from a 2.0 gram juvenile...	26
Figure 1.14	(1) Longitudinal orientation view of the enlarged distal vas deferens or terminal ampoule (Tam) from a less than 1.0 gram juvenile. (2) Longitudinal section of the terminal ampoule (Tam), (itself sectioned transversely or obliquely), from a less than 2.0	

Page

gram juvenile. (3) Longitudinal view of the terminal ampoule from a less than 2.0 gram juvenile. and (4) Orientation view of the terminal ampoule from a mature 35.0+ gram shrimp........................ 27

Figure 1.15 (1) Enlarged view of early testicular development stages from a 2.0 gram juvenile. (2) Enlarged view of mature testes from 15.0+ gram shrimp. (3) Orientation view of mature spermatozoa arranged in bundles, each of which is presumably surrounded by an acellular "primary spermatophore layer". The bundles appear very similar to seminiferous tubules with spermatozoa (Sp4) in the central core and the periphery partially made up of presumed nurse cells (hollow arrows). (4) Enlarged view of mature sperm of spermatozoa from terminal ampoule... 29

Figure 2.1 Assembly of the drip column for size-fctionation of syhthesizeed cDNA.. 41

Figure 2.2 Preparation of adaptor-ligated tester cDNA for hybridization and PCR... 50

Figure 3.1 Total RNA and mRNA (A) of *P. monodon* testes were used to construct testis cDNA library. First and second strand cDNA (B) were synthesized from purified mRNA... 80

Figure 3.2 Size-fractionation of testes cDNA of *P. monodon*........................... 81

Figure 3.3 Colony PCR for determining sizes of inserts of positive clones from a conventional testis cDNA library of *P. monodon*................. 82

Figure 3.4 The relative discovery rate of new genes from a conventional testis cDNA library of *P. monodon*... 82

Figure 3.5 Agarose gel electrophoresis illustrating digested cDNA of *P. monodon* with *Rsa* I... 89

Figure 3.6 Primary and secondary suppression PCR of subtracted and nonsubtracted of forward (A) and reverse (B) SSH of testis cDNA of *P. monodon*... 89

Figure 3.7 Colony PCR for determining sizes of inserts of positive clones from the forward SSH testis library of *P. monodon*......................... 90

Figure 3.8 Colony PCR for determining sizes of inserts of positive clones from the reverse SSH testis library of *P. monodon*.......................... 90

Page

Figure 3.9 Total RNA (A) and mRNA (B) of heart of *P. monodon* used to construct a conventional cDNA library. The first and second strand cDNA (C) were synthesized from purified mRNA of heart of juvenile *P. monodon*.. 98

Figure 3.10 Size-fractionation of heart cDNA of *P. monodon*............................ 98

Figure 3.11 Colony PCR for determining sizes of inserts of positive clones from a conventional heart cDNA library of *P. monodon*................. 99

Figure 3.12 The full length cDNA sequences of *SUMO* (1305 bp in length with an ORF of 282 bp corresponding to a polypeptide of 93 aa) of *P. monodon*... 103

Figure 3.13 The full length cDNA sequences of *cyclophilin A* (929 bp in length with an ORF of 495 bp corresponding to a polypeptide of 164 aa) of *P. monodon*... 104

Figure 3.14 The full length cDNA sequences of *translationally controlled tumor protein* (730 bp in length with an ORF of 507 bp corresponding to a polypeptide of 168 aa) of *P. monodon*.............. 104

Figure 3.15 The full length cDNA sequences of *small nuclear ribonucleoprotein polypeptide G* (694 bp in length with an ORF of 234 bp corresponding to a polypeptide of 77 aa) of *P. monodon*... 105

Figure 3.16 The full length cDNA sequences of *dynactin subunit 5* (1436 bp in length with an ORF of 594 bp corresponding to a polypeptide of 197 aa) of *P. Monodon*... 106

Figure 3.17 The full length cDNA sequences of *ubiquitin conjugating enzyme 2* (1416 bp in length with an ORF of 675 bp corresponding to a polypeptide of 224 aa) of *P. monodon*.............. 106

Figure 3.18 The full length cDNA sequences of *BUB3 budding uninhibited by benzimidazoles 3 homolog* (3554 bp in length with an ORF of 981 bp corresponding to a polypeptide of 326 aa) of *P. monodon*…....... 108

Figure 3.19 The full length cDNA sequences of *cell division control protein 2* (*cdc2*, 1696 bp in length with an ORF of 900 bp corresponding to a polypeptide of 299 aa) of *P. Monodon*... 109

Page

Figure 3.20 The full length cDNA sequences of *thioredoxin 1* (731 bp in length with an ORF of 318 bp corresponding to a polypeptide of 105 aa) of *P. Monodon*.. 110

Figure 3.21 The full length cDNA sequences of *multiprotein bridging factor 1* (715 bp in length with an ORF of 450 bp corresponding to a polypeptide of 149 aa) of *P. monodon*... 110

Figure 3.22 The full length cDNA sequences of *myosin 1 light chain* (1414 bp in length with an ORF of 465 bp corresponding to a polypeptide of 154 aa) of *P. monodon*.. 111

Figure 3.23 5′RACE-PCR (A) and 3′RACE-PCR (B) products of *low molecular weight neurofilament protein* (called *PMTST1*) and 3′RACE-PCR products of *prohibitin-2* (B), and *growth factor receptor-bound protein* (C) of *P. monodon*.................................... 113

Figure 3.24 The full length cDNA and deduced protein sequences (2194 bp in length with an ORF of 339 bp corresponding to a polypeptide of 112 aa) of a short form of *PMTST1* of *P. monodon* (*PMTST1-s*)..... 115

Figure 3.25 The full length cDNA and deduced protein sequences (2406 bp in length with an ORF of 339 bp corresponding to a polypeptide of 112 aa) of a long form of *PMTST1* of *P. monodon* (*PMTST1-l*)...... 116

Figure 3.26 Pairwise alignments of different isoforms of *PMTST1* cDNAs of *P. monodon*.. 117

Figure 3.27 The full length cDNA and deduced protein sequences of *prohibitin-2* (1380 bp; ORF 891 bp, 296 aa) of *P. monodon*........... 118

Figure 3.28 The full length cDNA and deduced protein sequences of the short form of *growth factor receptor-bound protein* (*GFRBP-s*, 1188 bp; ORF 636 bp, 211 aa) of *P. monodon*....................................... 119

Figure 3.29 The full length cDNA and deduced protein sequences of the long form of *growth factor receptor-bound protein* (*GFRBP-l*, 1885 bp; ORF 636 bp, 211 aa) of *P. monodon*.. 120

Figure 3.30 Pairwise alignments of different isoforms of *growth factor receptor-bound protein* (*GFRBP*) of *P. monodon*........................... 121

Figure 3.31 5′RACE-PCR products of *MIPP* (A) and 3′ and 5′RACE-PCR of *MIPP2* (B) of *P. monodon*... 122

Page

Figure 3.32 The full length cDNA and deduced amino acid sequences of *multiple inositol polyphosphate phosphatase* (*MIPP*, 1644 bp; ORF 1374 bp, 457 aa) of *P. monodon*.. 123

Figure 3.33 The full length cDNA and deduced amino acid sequences of *multiple inositol polyphosphate phosphatase 2* (*MIPP2*, 1746 bp; ORF 1194 bp, 397 aa) of *P. monodon*.. 124

Figure 3.34 Amino acid alignments of two isoforms of *multiple inositol polyphosphate phosphatase* (*MIPP*) of *P. monodon*........................ 125

Figure 3.35 Pairwise alignment of different forms of *P. monodon multiple inositol polyphosphate phosphatase* (*MIPP*).................................. 126

Figure 3.36 5′RACE-PCR products of *transformer-2* (A) and *Rac GTPase-activating protein 1* (B), and 3′RACE-PCR products of *flotillin-2* (C) of *P. monodon*... 127

Figure 3.37 The full length cDNA and deduced protein sequences of the long form of *Tra-2* (*Tra-2l*, 2673 bp; ORF 804 bp, 267 aa) of *P. monodon*... 128

Figure 3.38 The full length cDNA and deduced protein sequences of the short form of *Tra-2* (*Tra-2s*, 2658 bp; ORF 789 bp, 262 aa) of *P. monodon*... 129

Figure 3.39 Amino acid alignments of two isoforms of *Tra-2* of *P. monodon*.... 130

Figure 3.40 Pairwise alignments of different forms of *Tra-2* of *P. monodon*...... 132

Figure 3.41 The full length cDNA and deduced protein sequences of *Rac GTPase-activating protein 1* (2838 bp; ORF 1812 bp, 603 aa) of *P. monodon*... 133

Figure 3.42 The full length cDNA and deduced protein sequences of *flotillin 2* (1937 bp; ORF 1320 bp, 439 aa) of *P. monodon*........................... 134

Figure 3.43 The RACE-PCR products of 5′- and 3′- *DMC1/LIM15 homolog isoform 1* (A), 5′- *PGMRC1* (B), and 3′- *PGMRC1* (C) of *P. monodon*... 135

Figure 3.44 The full length cDNA and deduced protein sequences of *DMC1/LIM15 homolog isoform 1* (1661 bp; ORF 1026 bp, 341 aa) of *P. monodon*... 136

Page

Figure 3.45 The full length cDNA and deduced protein sequences of a short form of *progestin receptor membrane component 1* (*PGMRC1-s*) (2028 bp; ORF 573 bp, 190 aa) of *P. monodon*.............................. 137

Figure 3.46 The full length cDNA and deduced protein sequences of the medium form of *PGMRC1* (*PGMRC1-m*, 2896 bp; ORF 573 bp, 190 aa) of *P. monodon*.. 138

Figure 3.47 The full length cDNA and deduced protein sequences of the long form of *PGMRC1* (*PGMRC1-l*, 3019 bp; ORF 573 bp, 190 aa) of *P. monodon*... 139

Figure 3.48 Multiple alignment of three different forms of *PGMRC1* (*PGMRC1-s*, *PGMRC1-m* and *PGMRC1-l*) of *P. monodon*............ 142

Figure 3.49 5′ RACE-PCR product of *innexin1* (A) and 5′ and 3′ RACE-PCR products *innexin2* (B) of *P. monodon*.. 143

Figure 3.50 The full length cDNA and deduced protein sequences of *innexin1* (2505 bp; ORF 1143 bp, 380 aa) of *P. monodon*............................ 144

Figure 3.51 The full length cDNA and deduced protein sequences of *innexin2* (1615 bp; ORF 1077 bp, 358 aa) of *P. monodon*............................ 145

Figure 3.52 The RACE-PCR products of 5′-*Saposin* (A and B), 3′-*CDK7* (C) and 3′ *Dihydrolipoamide dehydrogenase* (C) of *P. monodon*.......... 146

Figure 3.53 The full length cDNA and deduced protein sequences of *saposin* (3034 bp; ORF 2589 bp, 862 aa) of *P. monodon*............................ 148

Figure 3.54 The full length cDNA and deduced protein sequences of *cdk7* (1431 bp; ORF 1062 bp, 353 aa) of *P. monodon*............................ 149

Figure 3.55 3′RACE-PCR of *troponinT* (A), 5′RACE-PCR of *dihydrolipoamide dehydrogenase* (B) and *ERO1* (C) and 3′RACE-PCR of *ERO1* (D) of *P. monodon*.................................... 151

Figure 3.56 The full length cDNA and deduced protein sequences of *troponin T isoform 3* (1430 bp; ORF 1107 bp, 368 aa) of *P. monodon*.......... 152

Figure 3.57 The full length cDNA and deduced protein sequences of the short form of *Ero1L, isoform B* (*ERO1-s*, 1756 bp; ORF 1413 bp, 470 aa) of *P. monodon*.. 153

Figure 3.58 The full length cDNA and deduced protein sequences of the long form of *Ero1L, isoform B* (*ERO1-l*; 1774 bp; 1434 bp, 477 aa) of *P. monodon*... 154

Page

Figure 3.59 Pairwise alignments of deduced amino acid sequences of the short and long forms of *P. monodon Ero1 isoform B*.............................. 154

Figure 3.60 Pairwise alignments of different isoforms of *Ero1, isoform B* of *P. monodon*... 155

Figure 3.61 The full length cDNA sequences of the short form of *dihydrolipoamide dehydrogenase* (*DLADH-s*, 1770 bp; ORF 1521 bp, 506 aa) of *P. monodon*.. 156

Figure 3.62 The full length cDNA sequences of the long form of *dihydrolipoamide dehydrogenase* (*DLADH-l*, 2050 bp; ORF 1521 bp, 506 aa) of *P. monodon*... 158

Figure 3.63 Pairwise alignments of different isoforms of *Dihydrolipoamide dehydrogenase* of *P. monodon*.. 159

Figure 3.64 A bootstrapped neighbor-joining tree illustrating relationships between *Tra-2* of various taxa.. 162

Figure 3.65 A bootstrapped neighbor-joining tree illustrating relationships between *Dmc1* of various taxa.. 163

Figure 3.66 A bootstrapped neighbor-joining tree illustrating relationships between *PGRMC* of various taxa.. 164

Figure 3.67 A bootstrapped neighbor-joining tree illustrating relationships between a homologue of *Cyclophilin A* from *P. monodon* (*PMCYA*), and that of various taxa... 165

Figure 3.68 A bootstrapped neighbor-joining tree illustrating relationships between a homologue of small ubiquitin modifier-1 from *P. monodon* (PMSUMO-1), and that of various taxa....................... 166

Figure 3.69 RT-PCR of *PMTST1* (374 bp) using the first strand cDNA of testes of broodstock and juveniles and ovaries of broodstock and juveniles of *P. monodon* for 25 (A), 30 (B) and 35 (C) cycles......... 168

Figure 3.70 RT-PCR of *MIPP*2 (A) and *MIPP* (B), using the first strand cDNA of testes of broodstock and juveniles and ovaries of broodstock and juveniles of *P. monodon* for 25 cycles..................................... 170

Figure 3.71 RT-PCR of *Dmc1* (A) and *HSP70-2* (B), using the first strand cDNA of testes of broodstock and juveniles and ovaries of broodstock and juveniles of *P. monodon* for 25 cycles.................... 171

Page

Figure 3.72 RT-PCR of *SUMO-1* (A), *CYA* (B), *Tra-2* (C), *prohibitin-2* (D), *Trap240* (E), *growth factor receptor-bound protein* (F), and *D2LIC* (G) using the first strand cDNA of testes of broodstock and juveniles and ovaries of broodstock and juveniles of *P. monodon* for 25 cycles.. 173

Figure 3.73 RT-PCR of *carbonyl reductase 1* (A), *Cdk2* (B), *ubiquitin carboxyl-terminal hydrolase 14* (C), *ubiquitin carboxyl-terminal hydrolase* 5 (D), *cdc2* (E), *E1B-55kDa-associated protein 5 isoform 5* (F), and *oncoprotein nm23* (G) using the first strand cDNA of testes of broodstock and juveniles and ovaries of broodstock and juveniles of *P. monodon* for 25 cycles.................... 174

Figure 3.74 RT-PCR of *innexin 2* (A), *PGRMC1* (B) using the first strand cDNA of testes of broodstock and juveniles and ovaries of broodstock and juveniles of *P. monodon* for 25 cycles.................... 176

Figure 3.75 Tissue distribution of *prohibitin-2* (A), *Trap240* (B), *Tra-2* (C), *SUMO-1* (D), *CYA* (E), *growth factor receptor-bound protein* (F), *D2LIC* (G) and *saposin* (H) using the first strand cDNA of testes (T), ovaries (O), eyestalks (E), gills (G), heart (H), hepatopancreas (HP), hemocytes (HC), lymphoid organs (LO), intestine (IT), stomach (ST), pleopods (PL), and thoracic ganglion (TG) of *P. monodon* broodstock for 25 cycles................................. 177

Figure 3.76 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *PMTST1* and *EF-1α* of testis from wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juveniles (CJ-TT), and ovaries from wild broodstock (WB-OV), and cultured juveniles (CJ-OV) and testes from broodstock after molting for Day 1-5, Day 6-9, and Day 10-16................................ 182

Figure 3.77 Histograms showing the expression levels of *PMTST1* in gonads of *P. monodon*.. 182

Figure 3.78 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *MIPP* and *EF-1α* of testis from wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juveniles (CJ-TT), and ovaries from wild broodstock (WB-OV),

Page

and cultured juveniles (CJ-OV) and testes from broodstock after molting for Day 1-5, Day 6-9, and Day 10-16................................ 183

Figure 3.79 Histograms showing the expression levels of *MIPP* in gonads of *P. monodon*... 183

Figure 3.80 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *Tra-2* and *EF-1α* of testis from wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juveniles (CJ-TT), and ovaries from wild broodstock (WB-OV), and cultured juveniles (CJ-OV) and testes from broodstock after molting for Day 1-5, Day 6-9, and Day 10-16................................ 184

Figure 3.81 Histograms showing the expression levels of *Tra-2* in gonads of *P. monodon*... 184

Figure 3.82 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *prohibitin 2* and *EF-1α* of testis from wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juveniles (CJ-TT), and ovaries from wild broodstock (WB-OV), and cultured juveniles (CJ-OV) and testes from broodstock after molting for Day 1-5, Day 6-9, and Day 10-16................................ 185

Figure 3.83 Histograms showing the expression levels of *prohibitin 2* in gonads of *P. monodon*... 185

Figure 3.84 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *Trap240* and *EF-1α* of testis from wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juveniles (CJ-TT), and ovaries from wild broodstock (WB-OV), and cultured juveniles (CJ-OV) and testes from broodstock after molting for Day 1-5, Day 6-9, and Day 10-16................................ 186

Figure 3.85 Histograms showing the expression levels of *Trap240* in gonads of *P. monodon*... 186

Figure 3.86 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *HSP70-2* and *EF-1α* of testis from wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juveniles (CJ-TT), and ovaries from wild broodstock (WB-OV), and cultured juveniles (CJ-OV) and testes from broodstock after molting for Day 1-5, Day 6-9, and Day 10-16................................ 187

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Page

Figure 3.87 Histograms showing the expression levels of *HSP70-2* in gonads of *P. monodon*.......... 187

Figure 3.88 Standard curve of *Dmc1* (A; r^2 for the standard curve = 0.996820) and EF-1α (B; r^2 for the standard curve = 0.99327) in testes of juvenile *P. monodon*.......... 188

Figure 3.89 Standard curve of *PGRMC1* (A; r^2 for the standard curve = 0.99820) and *EF-1α* (B; r^2 for the standard curve = 0.99520) in testes of juvenile *P. monodon*.......... 189

Figure 3.90 Real-time PCR analysis illustrating relative expression levels of *Dmc1* in testes of juvenile *P. monodon* after injected with normal saline or dopamine (10^{-6} mol/shrimp) for 3, 6, 12 and 24 hr.......... 190

Figure 3.91 Standard curve of *SUMO-1* (Efficiency for the standard curve = 1.993) and *CYA* (Efficiency for the standard curve = 1.980).......... 192

Figure 3.92 Standard curve of *Dmc1* (Efficiency for the standard curve = 1.971) and *saposin* (Efficiency for the standard curve = 1.984).......... 192

Figure 3.93 Standard curve of *spermatogonial stem-cell renewal factor* (Efficiency for the standard curve = 2.027) and *EF-1α* (Efficiency for the standard curve = 1.945).......... 192

Figure 3.94 Real-time PCR analysis illustrating the relative expression levels of *CYA* (A) and *SUMO-1* (B) in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting.......... 193

Figure 3.95 Real-time PCR analysis illustrating the relative expression levels of *saposin* in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting.......... 194

Figure 3.96 Real-time PCR analysis illustrating the relative (A) and absolute (B) expression levels of *spermatogonial stem cell renewal factor* in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting.......... 196

Figure 3.97 Real-time PCR analysis illustrating the relative (A) and absolute (B) expression levels of *Dmc1* in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting.......... 197

Page

Figure 3.98 Microarray analysis using cDNA of testes of juveniles and broodstock of *P. monodon* labeled with Cy3 and Cy5 as the probes, respectively........ 199

Figure 3.99 Microarray analysis using cDNA of broodstock and juvenile testes of *P. monodon* labeled with Cy3 and Cy5 as the probes, respectively........ 200

Figure 3.100 Microarray analysis showing up- and/or down-regulation of examined genes in testes of broodstock (A) and juvenile (B) *P. monodon*........ 201

Figure 3.101 Microarray analysis using cDNA of normal testes and dopamine-treated testes of juvenile shrimps (10^{-6} mol/shrimp) at 3 hours post injection labeled with Cy3 and Cy5 as the probes, respectively........ 203

Figure 3.102 Microarray analysis using cDNA of normal testes and dopamine-treated testes of juvenile shrimps (10^{-6} mol/shrimp) at 6 hours post injection labeled with Cy3 and Cy5 as the probes, respectively........ 203

Figure 3.103 Microarray analysis showing up- and/or down-regulation of examined genes in testes of dopamine-treated juveniles of *P. monodon* at 3 hr (A) and 6 hr post injection........ 205

Figure 3.104 Alignments between the full length cDNA of *spermatogonial stem-cell renewal factor* initially characterized from testes (ORF-TT) and ovaries (ORF-OV) of *P. monodon*........ 207

Figure 3.105 Alignments of deduced amino acid sequences of *spermatogonial stem-cell renewal factor* from testes (ORF-TT) and ovaries (ORF-OV) of *P. monodon*........ 208

Figure 3.106 SDS-PAGE showing *in vitro* expression of four recombinant clones of *spermatogonial stem-cell renewal factor* (A) and *Dmc1* (B) of *P. monodon* at 0 hr, 3 hr, and 6 hr after induced by 1.0 mM IPTG........ 209

Figure 3.107 SDS-PAGE showing *in vitro* expression of four recombinant clones of *SUMO-1* (A) and *cyclophilin A* (B) of *P. monodon* at 0 hr, 3 hr, and 6 hr after induced by 1.0 mM IPTG........ 209

Page

Figure 3.108 *In vitro* expression of eight recombinant clones of *PGRMC1* of *P. monodon* after induced by 1 mM IPTG overnight using SDS-PAGE (A) and western blot analysis (B).............................. 210

Figure 3.109 *In vitro* expression of recombinant spermatogonial stem cell renewal factor of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG using SDS-PAGE (A) and western blot analysis (B)... 211

Figure 3.110 *In vitro* expression of recombinant Dmc1 of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG using SDS-PAGE (A) and western blot analysis (B).............................. 211

Figure 3.111 *In vitro* expression of recombinant SUMO-1 of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG using SDS-PAGE (A) and western blot analysis (B).............................. 212

Figure 3.112 *In vitro* expression of recombinant cyclophilin A of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG using SDS-PAGE (A) and western blot analysis (B).............................. 213

Figure 3.113 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant spermatogonial stem cell renewal factor after IPTG induction for 6 hr at 37°C as the insoluble protein.. 213

Figure 3.114 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant spermatogonial stem cell renewal factor after IPTG induction overnight at 15°C as both soluble and insoluble proteins... 214

Figure 3.115 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant Dmc1 after IPTG induction for 6 hr at 37°C as the insoluble protein.. 214

Figure 3.116 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant Dmc1 after IPTG overnight at 15°C as the insoluble protein.. 215

Figure 3.117 SDS-PAGE (A) and western blot analysis (B) showing more expression of recombinant cyclophilin A after IPTG induction for 3 hr at 37°C as the insoluble protein than the soluble protein.. 215

Page

Figure 3.118 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant cyclophilin A after IPTG induction overnight at 15°C as both soluble and insoluble proteins............. 216

Figure 3.119 SDS-PAGE (A) and western blot analysis (B) showing equal expression of recombinant SUMO-1 after IPTG induction for 6 hr at 37°C as the soluble and insoluble proteins........................ 216

Figure 3.120 Purification of recombinant spermatogonial stem cell renewal factor of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).. 218

Figure 3.121 Purification of recombinant Dmc1 of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D)..................... 218

Figure 3.122 Purification of recombinant SUMO-1 of *P. monodon* (cultured at 37°C, for 6 hr). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D)..................... 219

Figure 3.123 Purification of recombinant cyclophilin of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D)............ 219

Figure 3.124 Purification of recombinant spermatogonial stem cell renewal factor of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).. 221

Figure 3.125 Purification of recombinant Dmc1 of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D)..................... 221

Figure 3.126 Purification of recombinant cyclophilin A of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).. 222

LIST OF ABBREVIATIONS

bp	base pair
°C	degree Celcius
DEPC	Diethylpyrocarbonate
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
dGTP	deoxyguanosine triphosphate
dTTP	deoxythymidine triphosphate
DNA	deoxyribonucleic acid
HCl	hydrochloric acid
IPTG	isopropyl-thiogalactoside
Kb	kilobase
M	Molar
$MgCl_2$	magnesium chloride
mg	Milligram
ml	Millilitre
mM	Millimolar
ng	Nanogram
OD	optical density
PCR	polymerase chain reaction
RNA	Ribonucleic acid

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

RNase A	ribonuclease A
rpm	revolution per minute
RACE	Rapid Amplification of cDNA ends
RT	Reverse transcription
SDS	sodium dodecyl sulfate
Tris	tris (hydroxyl methyl) aminomethane
μg	Microgram
μl	Microlitre
μM	Micromolar
UV	ultraviolet

CHAPTER I

INTRODUCTION

1.1 General introduction

A total of 343 economically important shrimp species are reported by the FAO (Bailey-Brock and Moss, 1992). Among cultured marine shrimp, one of the most economically important species is the giant tiger shrimp, *Penaeus monodon*. The world shrimp aquaculture production (Fig. 1.1), which had stabilized in the 1990s, has shown strong increases in subsequent years. In 2003, shrimp aquaculture exceeded 1.6 million metric tons (MT).

In Thailand, *P. monodon* had been intensively cultured for more than two decades and had contributed approximately 60% of the total cultivated shrimp production. The reasons for this are supported by several factors including the appropriate farming areas without serious disturbing from typhoons or cyclone, small variable of seawater during seasons, and ideal soils for pond construction. Culture of *P. monodon* had increased the national revenue, therefore *P. monodon* was, until recently, the most economically important cultured species in Thailand.

The production of *P. monodon* is largely constrained by the current dependency on wild-caught broodstock which varies in both quality and quantity. Recently, the farming of *P. monodon* in the region has significantly declined. Previously, annual production of farmed *P. monodon* in Thailand alone has reached or exceeded 200,000 metric tons since 1993 (Asian Shrimp Culture Council, 1996) but the production was recently dropped owing to shrimp diseases and the lack of high quality wild and/or domesticated broodstock of *P. monodon* (Limsuwan, C., 2004). Therefore, *Litopenaeus vannamei* was introduced as a new cultured species and significantly contributed to the production since 2004.

The amount of exported frozen shrimp is relatively stable (Tables 1.1 and 1.2). However, the annual production of *P. monodon* has consistently decreased even though values of frozen shrimp export increase since 2003 (Table 1.1).

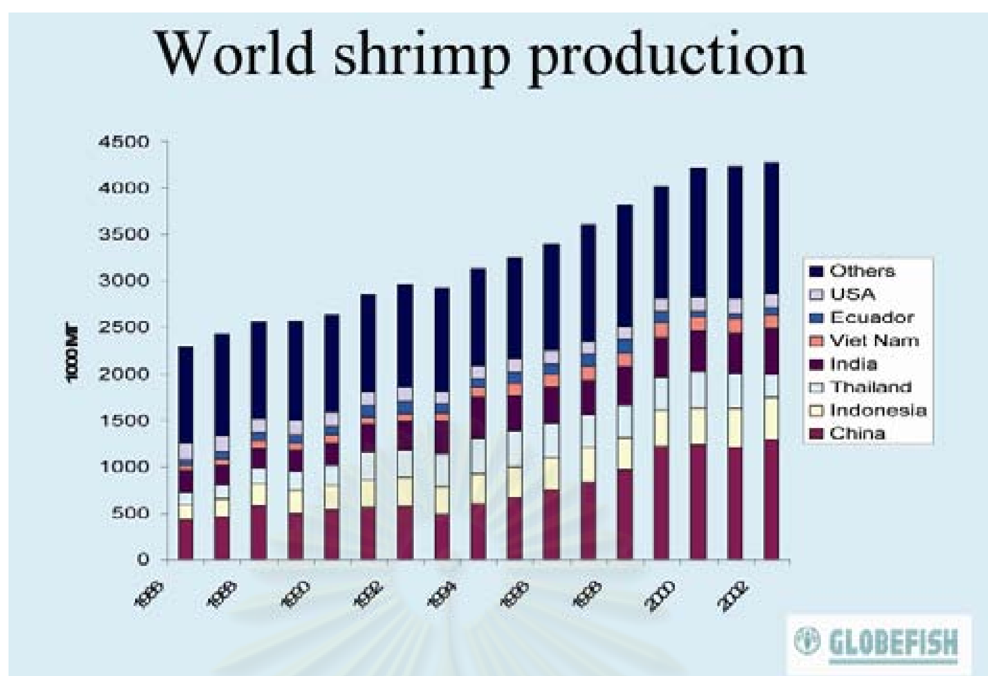


Figure 1.1 World shrimp production since 1996-2003 (Globefish).

Table 1.1 Thai frozen shrimp export between 2001-2006.

Item	2001	2002	2003	2004	2005	2006
Value (MB)						
Total export	**2,923,941.4**	**3,325,630.1**	**3,874,823.8**	**4,439,310.6**	**4,439,310.6**	**4,938,508.2**
Chilled and frozen shrimp	**34,406.2**	**35,921.2**	**32,536.1**	**37,730.3**	**37,730.3**	**37,802.5**
Black tiger shrimp	28,283	24,179	15,029	8,571	8,571	6,991
Giant freshwater prawn	494	603	796	1,356	1,356	1,714
Other shrimp	5,629	11,139	16,711	7,804	27,804	29,098

Source: Ministry of Commerce Thailand

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 1.2 Comparisons between Thai frozen shrimp export between January-June 2007 and 2008

Country	Jan -Jun 2007		Jan -Jun 2008		% difference	
	Quantity (MT)	Value (MB)	Quantity (MT)	Value (MB)	Quantity (MT)	Value (MB)
Asia	**41,612**	**9,805**	**49,249**	**10,381**	**18.35**	**5.87**
China	1,779	357	3,588	452	101.69	26.61
Japan	26,591	6,642	30,789	7,392	15.79	11.29
- others	13,242	2,806	14,872	2,537	12.31	-9.59
USA	**74,186**	**17,161**	**70,837**	**15,722**	**-4.51**	**-8.39**
EU	**12,696**	**3,244**	**14,476**	**3,532**	**14.02**	**8.88**
Australia	**3,420**	**777**	**2,609**	**595**	**-23.71**	**-23.42**
Others	**14,451**	**3,077**	**11,990**	**2,458**	**-17.03**	**-20.12**
Total	**146,365**	**34,064**	**149,161**	**32,688**	**1.91**	**-4.04**

Source: Ministry of Commerce Thailand (Amounts–tons and values–million bahts)

Farming of *P. monodon*, in Thailand relies almost entirely on wild-caught broodstock for supply of juveniles because pond-reared *P. monodon* rarely produced sufficient quality of larvae required by the industry (Withyachumnarnkul et al., 1998). Low degrees of reproductive maturation of captive *P. monodon* females and low quality of spermatozoa of captive males have limited the potential of genetic improvement resulted in remarkably slow domestication and selective breeding programs of *P. monodon* in Thailand (Withyachumnarnkul *et al.*, 1998; Preechaphol *et al.*, 2007). Selective breeding for the improvement of particular performance traits, such as growth and survival, is important for the future growth of penaeid aquaculture production.

Practically, breeding of *P. monodon* using spermatozoa of captive males yields low quality offspring. The use of spermatozoa from wild males with either wild or pond-reared females has resolved the problem successfully (B. Withyachumnarnkul, personal communication).

The basic information related to testicular development and sperm quality is rather limited in penaeid shrimp (Benzie, 1998). An initial step toward understanding

molecular mechanisms of testicular and spermatozoa development in *P. monodon* is to identify and characterize differentially expressed genes in different stages of testicular development of this species.

In addition, the fundamental controls of growth in penaeid shrimp are largely unstudied. Genes encoding vertebrate-like growth factors and cell cycle regulating proteins should be characterized.

1.2 Taxonomy of *P. monodon*

The giant tiger shrimp (*P. monodon*) is taxonomically classified as a member of Phylum Arthropoda; Subphylum Crustacea; Class Malacostraca; Subclass Eumalacostraca; Order Decapoda; Suborder Natantia; Infraorder Penaeidea; Superfamily Penaeoidea; Family Penaeidae Rafinesque, 1985; Genus *Penaeus* Fabricius, 1798 and Subgenus *Penaeus*. (Bailey-Brock and Moss, 1992).

The scientific name of giant tiger shrimp is *Penaeus monodon* where the English common name is the giant (black) tiger shrimp.

1.3 The reproductive organs of marine shrimp

The male reproductive system includes paired testes, paired vas deferens, and a petasma (Fig. 1.3). Mating of *P. monodon* occurs at night after the female molts. Sperm is deposited into a special structure called the thelycum on the underside of the female's thorax (Fig. 1.3). A single female usually produces 250,000-800,000 eggs, which are released into the water and hatch within 18 hours into nauplii larvae. Larvae pass through six nonfeeding nauplii, 3 zoeal and 3 mysis stages and metamorphoresed to the postlarva. Life span of shrimp is probably less than two years.

1.4 The importance for domestication of *P. monodon*

Farming of *P. monodon* presently relies almost entirely on wild-caught broodstock for the seed supply because breeding of *P. monodon* in captivity is extremely difficult. This open reproductive cycle and reliance on wild stocks of *P. monodon* results in heavy exploitation of female broodstock from wild populations.

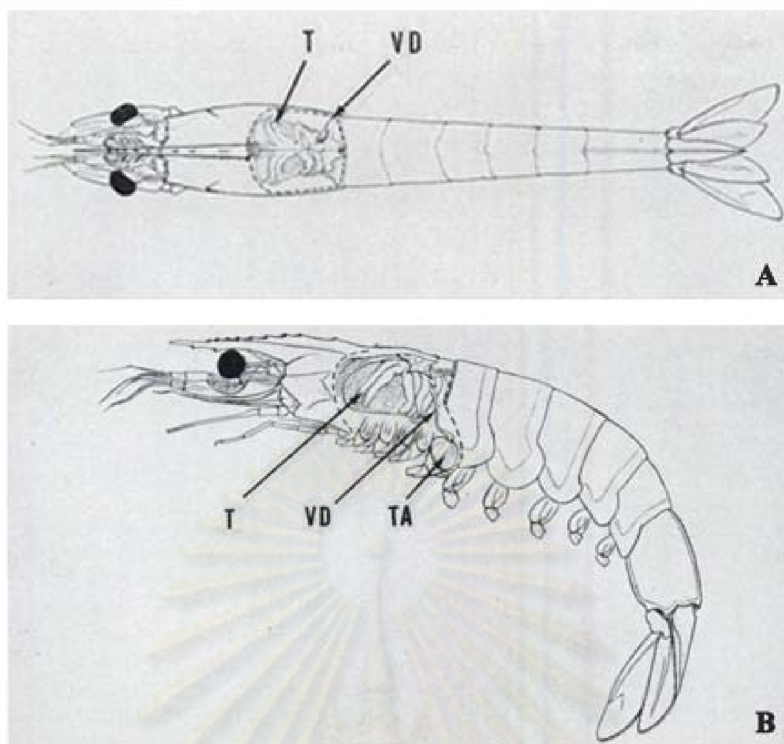


Figure 1.2 Diagram of male, dorsal view (A) and lateral view (B) to show reproductive organ. T = testis, VD = vas deferens, and TA = terminal ampoule (after King, 1998).

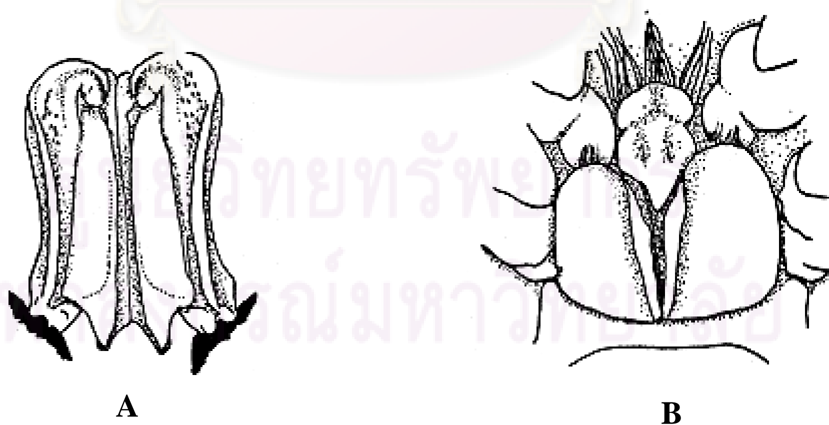


Figure 1.3 Sexes of juveniles and broodstock of penaeid shrimp can be externally differentiated by petasma of males (A) and thelycum of females (B) (after King, 1998).

The lack of high quality wild and/or domesticated broodstock of *P. monodon* has possibly caused an occurrence of a large portion of stunted shrimps at the harvest time (3-5 g rather than approximately 25 g body weight at 4 month cultivation period). As a result, the farmed production of *P. monodon* has significantly decreased since the last few years.

Progress in genetic and biotechnology researches in penaeid shrimps have been slow because a lack of knowledge on fundamental aspects of their biology (Benzie, 1998). A research concerning domestication of *P. monodon* is being carried out in Thailand by production of high quality pond-reared *P. monodon* broodstock. Subsequently, it is expected that selective breeding programs of *P. monodon* will be the key to provide shrimps having commercially desired phenotypes (e.g. high growth rate and/or disease resistance) and to produce *P. monodon* stocks with the ability to induce high quality egg development in domesticated females without the irreversible side-effects caused by a typical eyestalk ablation technique (Lyons and Li, 2002).

Despite the success of the farmed production, the shrimp industry has encountered several problems including environmental degradation, outbreaks of diseases, and shortages of high quality broodstock. The white shrimp (*Litopenaeus vannamei*) was then introduced into the country and initially contributed on the cultured production of Thailand significantly. However, the price of *L. vannamei* is quite low and broodstock used relies almost entirely on genetic improved stocks brought from different sources. In addition, the labor costs in Thailand are higher than other countries (e.g. Vietnam and China) preventing the advantage of competition for the world market. In contrast, the market of premium-sized *P. monodon* is still open. Applications of the knowledge for genetic selection and biotechnology of *P. monodon* should be studied and practically implemented to fulfill that purpose.

Determining the relative effect that male and female broodstock quality is having on reproductive performance, particularly for key parameters such as hatching rate, will enhance the rate at which improvements reproductive performance of domesticated stocks can be achieved. Previous studies assessing the reproductive performance of reciprocally crossed wild and pond-reared broodstock found that the

wild females outperformed domesticated females in terms of maturation, spawning and total egg production in *P. monodon* (Menasveta et al., 1993)

Although phenotypic improvement can be accomplished through conventional breeding programs, knowledge from genome studies and molecular markers linked to important traits (marker-assisted selection, MAS) can also be directly applied to improve artificial selection processes more efficiently.

1.5 Molecular genetic approaches used in this thesis

1.5.1 PCR

The introduction of the polymerase chain reaction (PCR; Mullis et al.,1987) has opened a new approach for molecular genetic studies. This method is technique for enzymatically replicating DNA without using a living organism, such as *E. coli* or yeast and is a method using specific DNA sequences by the two oligonucleotide primers, usually 18 - 27 nucleotides in length. Million copies of the target DNA sequence can be synthesized from the low amount of starting DNA template within a few hours.

The PCR reaction components are composed of DNA template, a pair of primers for the target sequence, dNTPs (dATP, dCTP, dGTP and dTTP), PCR buffer and heat-stable DNA polymerase (usually *Taq* polymerase). The amplification reaction typically consists of three steps; denaturation of double stranded DNA at high temperature, annealing to allow primers to form hybrid molecules at the optimal temperature, and extension of the annealed primers by heat-stable DNA polymerase. The cycle is repeated for 30 - 40 times (Figure 1.4). The amplification product is determined by agarose or polyacrylamide gel electrophoresis.

1.5.2. Expression Sequence Tag (EST) analysis

ESTs are large-scale single-pass sequences of randomly picked clones from a cDNA library usually constructed from mRNA at a particular developmental stage and/or tissue. This method has been widely employed for discovering novel and uniquely expressed genes, and for characterizing the gene expression profiles of several tissues.

The general principles for construction of a cDNA library (Figure 1.5) begin with purification of the target mRNA that is reverse-transcribed to the first-strand cDNA. This step is catalyzed by reverse transcriptase using the oligo (dT) primer as the synthesizing primer. The second-strand DNA is then copied from the first-strand cDNA using *E. coli* DNA polymerase I. The double-strand cDNA is ligated to adapter and subsequently to an appropriate vector using T4 DNA ligase. The recombinant vector-cDNA molecules are packaged (λ vector) *in vitro* and transfected to the appropriate host. If a plasmid is used recombinant plasmid is transformed into *E. coli* host cells to generate a cDNA library.

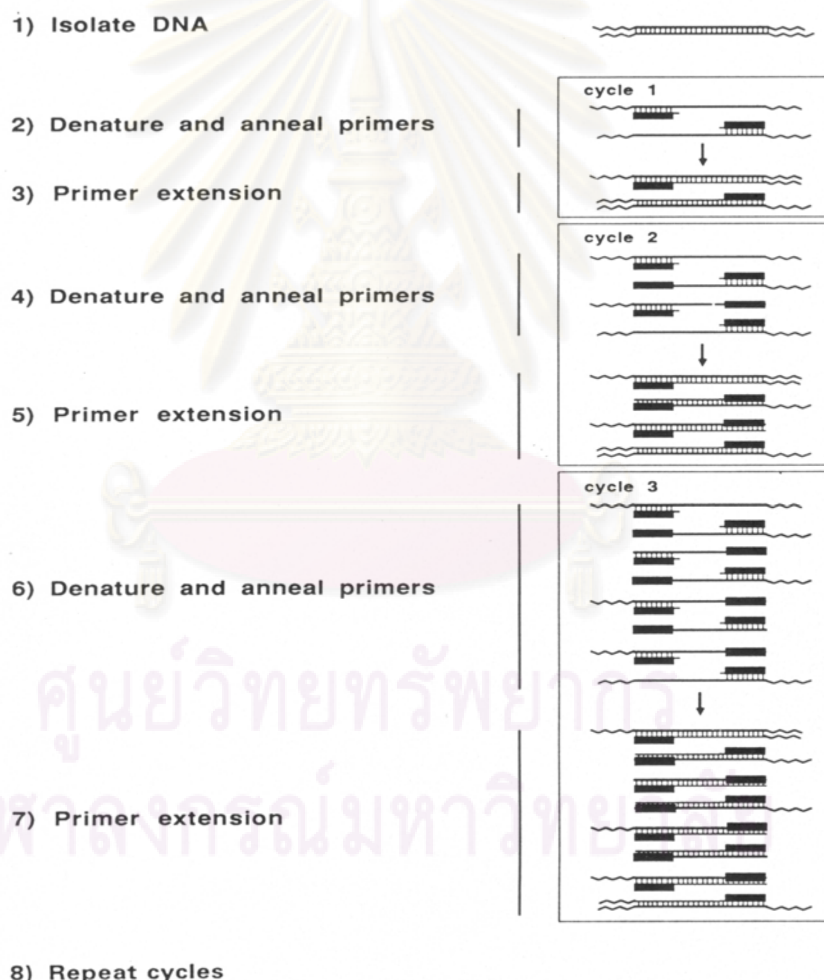


Figure 1.4 General illustration of the polymerase chain reaction (PCR) for amplification of the target DNA (Avise, 1994).

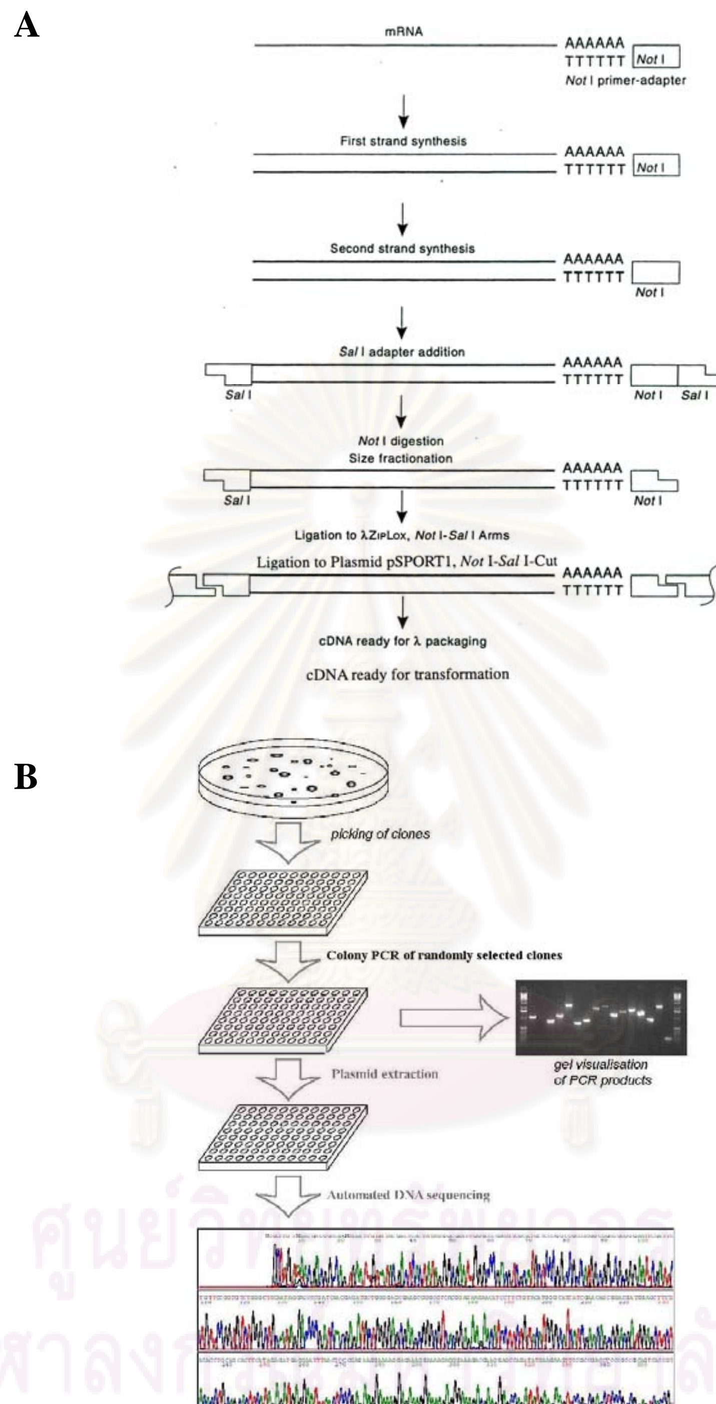
ESTs can be sequenced from either 5′ or 3′ ends of cloned cDNA. The 3′ end of the cloned insert is usually marked by the poly A stretch which is often problematic for thermostable polymerase sequencing, and sequencing through poly T can reduce the length and quality of the subsequent sequence. Nevertheless, 3′ UTR usually exhibit high polymorphism and is a promising location for SNP identification. The 5′ ESTs have the advantage of being more likely to include some of the open reading frame (ORF) of the cDNA and thus facilitate identification of the encoded product.

EST sequences are used as the tag to homology search through the sequence data in the GenBank (Altschl et al., 1990). The Blast*N* program uses nucleotide sequence to compare against the NCBI nucleotide database whereas the BLAST*X* uses the translated protein products to compare against the NCBI protein database in all possible 6 reading frames. Generally, sequences are considered to be significantly matched when the possibility value (E-value) is less than 10^{-4} and the match length is > 100 nucleotides for Blast*N* and a match length is > 10 amino acid residues for Blast*X*, respectively, (Anderson and Brass, 1998).

EST analysis is an important tool for several applications. This approach has mainly applied for rapid gene discovery of genes, comparative genomics and functional genomics in various organisms. After characterization and annotation, cDNA or designed oligonucleotides of transcripts can be further used for microarray analysis. Construction of genetic linkage maps and/or physical maps of interesting species can be carried out by development and sequencing of EST-derived markers using genomic DNA of species under investigation (Liu and Cordes, 2004).

1.5.3 Suppression subtractive hybridization (SSH)

Suppression subtractive hybridization is a powerful technique to compare two populations of mRNA to obtain clones of genes that are expressed (or differential expressed) in one population but not in the other. Although there are several different methods, a Clontech PCR select cDNA subtraction method is convenient for generation of differentially expressed sequences (Fig. 1.6).



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Figure 1.5 Overview for construction of cDNA inserts (A, www.bdbiosciences.com) and automated DNA sequencing (single-pass) of randomly selected cDNA clones. The entire process simply called EST analysis.

First, cDNA is synthesized from two types of mRNA populations and cDNA that contains specific (differentially expressed) transcripts is regarded as where the tester, and the reference cDNA is regarded as the driver. The tester and driver are digested with *Rsa* I. The tester cDNA is divided to two portions, and each is ligated with different cDNA adaptor. The ends of adaptor do not have a phosphate group,

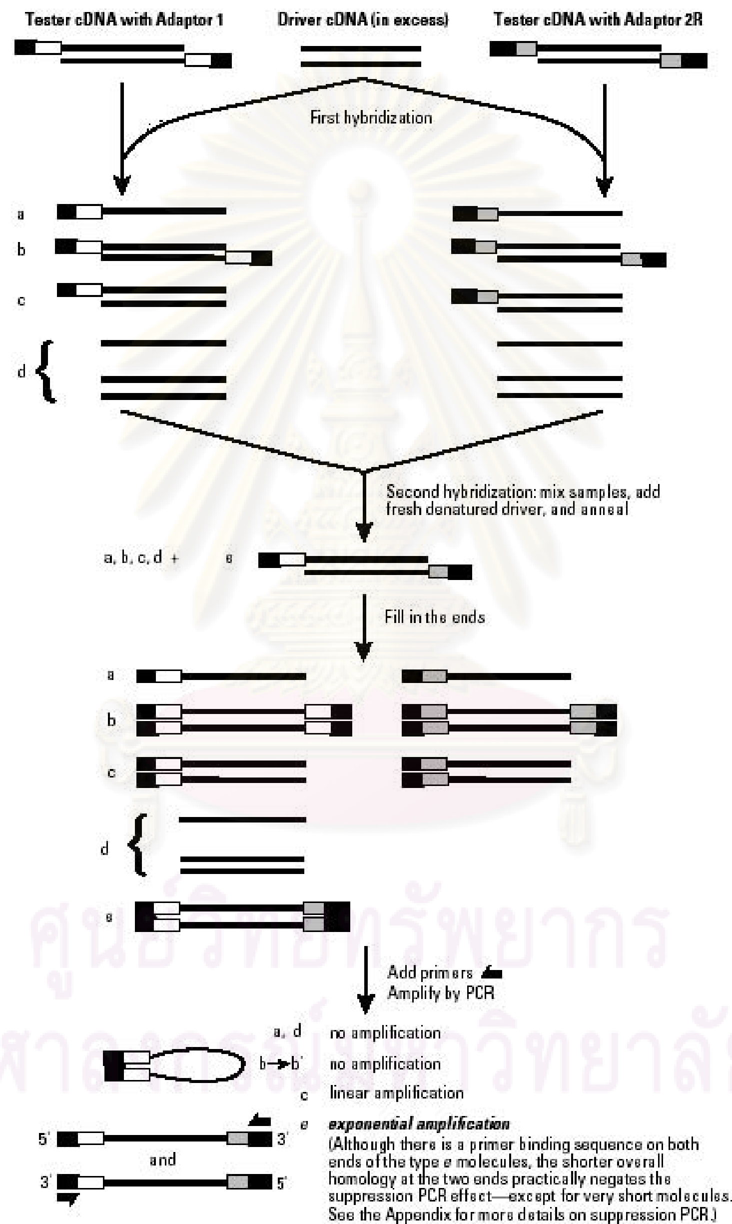


Figure 1.6 Overview of the Clontech PCR-Select^{TM} procedure (www.bdbiosciences.com). The cDNA in which specific transcripts are to be found is called “tester” and the reference cDNA is called “driver”.

therefore only one strand of each adaptor attaches to the 5′ ends of cDNA. The two adaptors have stretches of identical sequences to allow annealing of the PCR primer once recessed ends have been filled in.

Tester and driver cDNAs are hybridized. In the first hybridization, an excess of the driver is added to each tester. The samples are heat-denatured and allow to anneal Differential expressed transcripts are then enriched. During the second hybridization, template for PCR amplification is generated from differentially expressed transcripts. Two rounds of suppression PCR was carried out and only differentially expressed transcripts are amplified exponentially. Subtractive cDNA products are then cloned into the T-vector. Positive clones are characterized by hybridization or sequencing.

1.5.4 Reverse transcription-polymerase chain reaction (RT-PCR)

RT-PCR is a comparable method of conventional PCR but the first strand cDNA template rather than genomic DNA was used as the template in the amplification reaction (Fig. 1.7). It is a direct method for examination of gene expression of known sequence transcripts in the target species. The template for RT-PCR can be the first stranded cDNA synthesized from total RNA or poly A^+ RNA. Reverse transcription of total RNA can be performed with oligo(dT) or random primers using a reverse transcriptase The product is then subjected to the second strand synthesis using a gene-specific forward primer.

RT-PCR can also be used to identify homologues of interesting genes by using degenerate primers and/or conserved gene-specific primers from the original species and the first strand cDNA of the interesting species is used as the template. The amplified product is further characterized by cloning and sequencing.

Semi-quantitative RT-PCR is an quantitative approach where the target genes and the internal control (e.g. a housekeeping gene) were separately or simultaneously amplified using the same template. The internal control (such as *β-actin*; *elongation factor*, *EF-1α* or *G3PDH*) is used under the assumption that those coding genes are transcribed constantly and independently from the extracellular environment stimuli

and that their transcripts are reverse transcribed with the same efficiency as the product of interesting transcript.

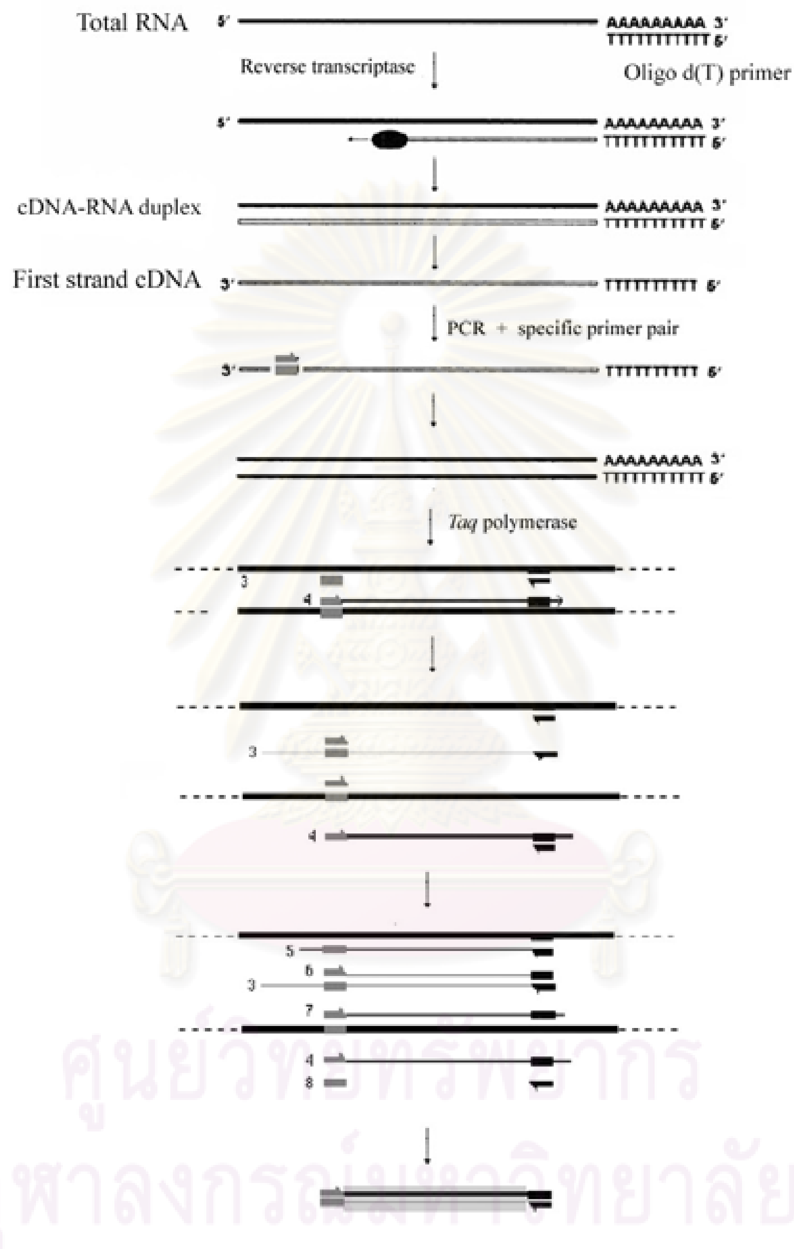


Figure 1.7 Overall concepts of RT-PCR. During the first strand cDNA synthesis, an oligo d(T) (or random primers) primer anneals and extends from sites present within mRNA. The second strand cDNA synthesis primed by the 18 – 25 base specific primer proceeds during a single round of DNA synthesis catalyzed by thermostable DNA polymerase (e.g. *Taq* polymerase) (www.copewithcytokines.de).

1.5.5 Rapid amplification of cDNA ends-polymerase chain reaction RACE-PCR

RACE-PCR is the common approach used for isolation of the full length of characterized cDNA. Using SMART (**S**witching **M**echanism **A**t 5′ end of RNA **T**ranscript) technology, terminal transferase activity of Powerscript Reverse Transcriptase (RT) adds 3 - 5 nucleotides (predominantly dC) to the 3′ end of the first-strand cDNA. This activity is harnessed by the SMART oligonucleotides whose terminal stretch of dG can anneal to the dC-rich cDNA tail and serve as an extended template for reverse transcriptase. A complete cDNA copy of original mRNA is synthesized with the additional SMART sequence at the end (Fig. 1.8).

The first strand cDNA of 5′ and 3′ RACE is synthesized using a modified oligo (dT) primers and serve as the template for RACE PCR reactions. Gene specific primers (GSPs) are designed from interested gene for 5′- RACE PCR (antisense primer) and 3′-RACE PCR (sense primer) and used with the universal primer (UPM) that recognize the SMART sequence. RACE products are characterized. Finally, the full length cDNA is constructed.

1.5.6 Real-time PCR

Real Time PCR is a kinetic approach based on the polymerase chain reaction, which is used to amplify and simultaneously quantify a target DNA molecule. It enables both detection and quantification (as absolute number of copies or relative amount when normalized to DNA input or additional normalizing genes) of a specific sequence in the sample.

The procedure follows the general principle of PCR. Its key feature is that the amplified DNA is quantified as it accumulates in the reaction in *real time* after each amplification cycle. Two common methods of quantification are the use of fluorescent dyes that intercalate with double-stranded DNA such as SYBR green, and modified DNA oligonucleotide probes that are fluorescent when hybridized with a complementary DNA.

The general principle of SYBR green polymerase chain reaction composed of first step, denaturation: at the beginning of amplification, the unbound dye molecules weakly fluorescence, the second step, annealing: after annealing of the primer, a few

dye molecules bind to the double strand. The last step, extension: during elongation, more dye molecules bind to the newly synthesized DNA. Fluorescence measurement at the end of the elongation step of every PCR cycle is performed to monitor the increasing amount of amplified DNA (Fig. 1.9).

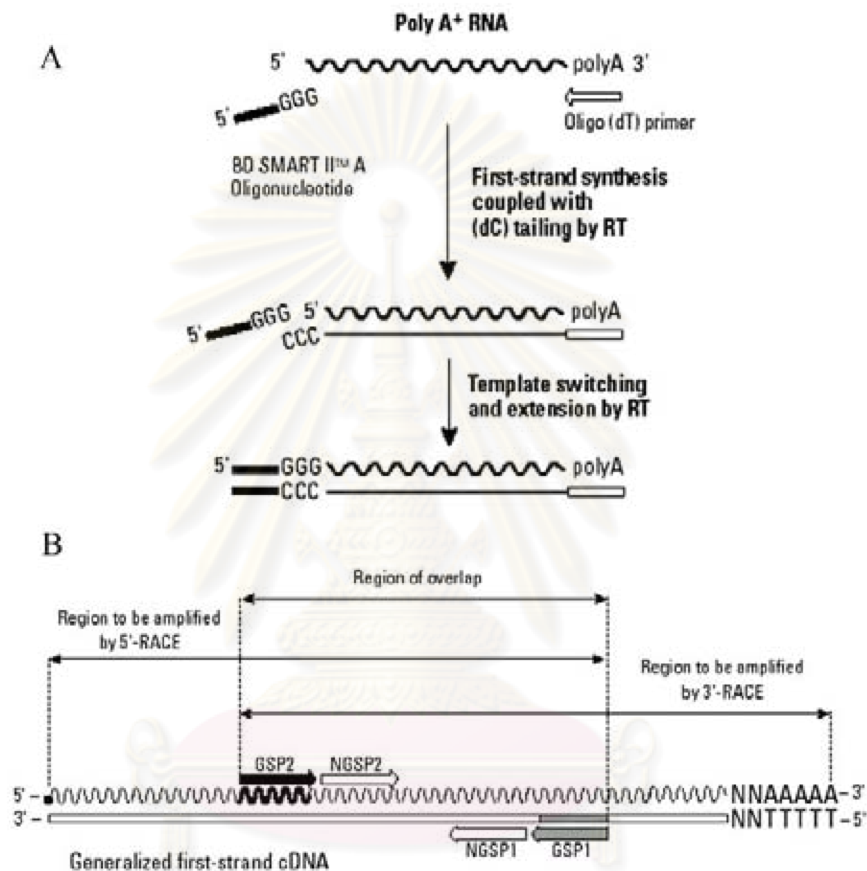


Figure 1.8 Overview of the SMART^{TM} RACE cDNA Amplification Kit (www.bdbiosciences.com).

A. Mechanism of SMART cDNA synthesis. First strand synthesis is primed using a modified oligo (dT) primer. After reverse transcriptase reaches the end of the mRNA template, it added several dC residues. The SMART II A Oligonucleotide annels to the tail of the cDNA and serves as an extended template for PowerScript-RT.

B. Relationships of gene-specific primers to the cDNA template. This diagram shows a generalized first strand cDNA template.

Real-time polymerase chain reaction in the laboratory can be applying numerous applications. It is commonly used for both diagnostic and research applications. For diagnostic purposes, real-time PCR is applied to rapidly detect the presence of genes involved in infectious diseases, cancer and genetic abnormalities. In the research applications, real-time PCR is mainly used to provide highly sensitive quantitative measurements of gene transcription.

The technology may be used in determining how the genetic expression of a particular gene changes over time, such as in the response of tissue and cell cultures to an administration of a pharmacological agent, progression of cell differentiation, or in response to changes in environmental conditions.

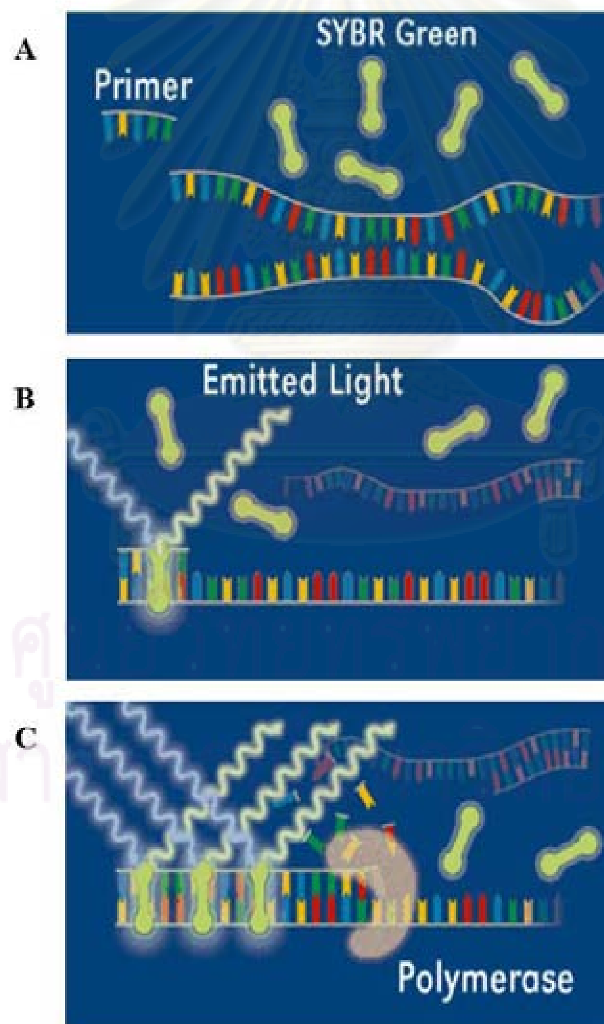


Figure 1.9 An overall concept of the Real-time PCR procedure (www.thaiscience.com/lab_vol/p23/Real-Time_PCR.asp).

1.5.7 Microarray analysis

A microarray is a tool for analyzing expression of a large number of genes simultaneously. A microarray consists of a small membrane or glass slide containing samples of many genes arranged in a regular pattern. Advances in biotechnology allow the identification and classification of this DNA sequence information and the assignment of functions to these new genes: the emergence of DNA microarray technology.

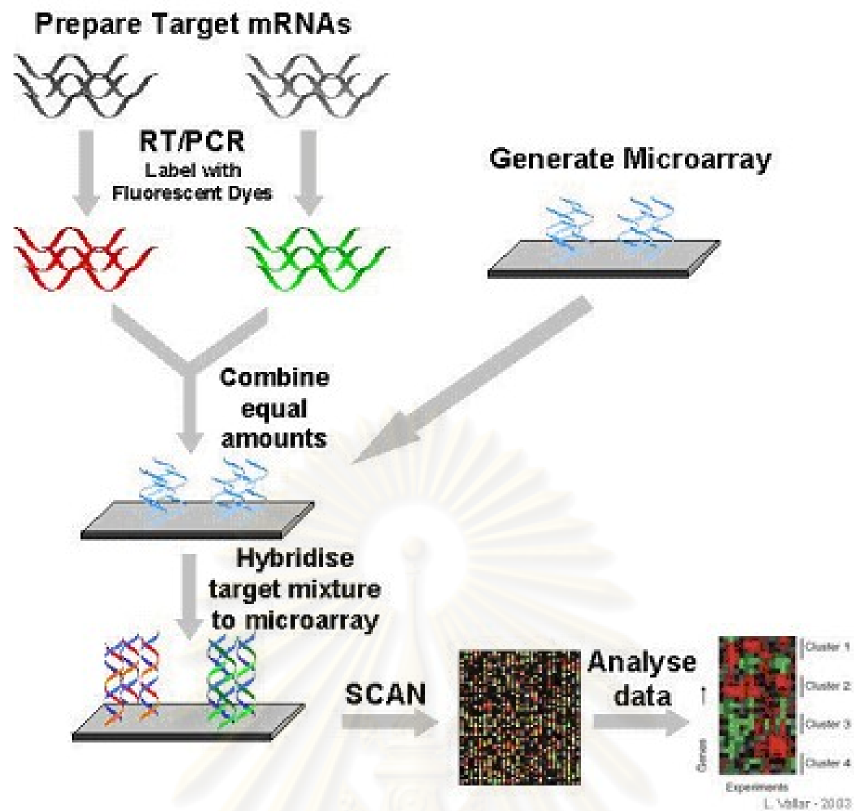
Microarrays are therefore useful when one wants to survey a large number of genes quickly. Microarrays may be used to assay gene expression within a single sample or to compare gene expression in two different cell types or tissue samples, such as in healthy and diseased tissue. Because a microarray can be used to examine the expression of hundreds or thousands of genes at once, it promises to revolutionize the way scientists examine gene expression. This technology is still considered to be in its infancy; therefore, many initial studies using microarrays have represented simple surveys of gene expression profiles in a variety of cell types.

A microarray works by exploiting the ability of a given mRNA molecule to bind specifically to, or hybridize to, the DNA template from which it originated. By using an array containing many DNA samples, scientists can determine, in a single experiment, the expression levels of hundreds or thousands of genes within a cell by measuring the amount of mRNA bound to each site on the array. With the aid of a computer, the amount of mRNA bound to the spots on the microarray is precisely measured, generating a profile of gene expression in the cell.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

1.6 Sex determining systems

Apart from the control of reproductive maturation, manipulation of sex ratio is an important tool for crustacean aquaculture. An understanding of sex determination and differentiation is, therefore, necessary for designing appropriate breeding programs in penaeid species.



)

Figure 1.10 An overall concept for large scale screening of gene expression based on microarray analysis (www.microarray.lu/en/MICROARRAY_Overview.shtml).

Sex determination is problematic in researches of many species. This can usually be solved by the application of DNA based technology but this is only possible if a sexspecific (located on unique sex chromosomes) marker is available. The lack of sex chromosomes reported in P. monodon and other penaeid shrimps implied that development of genomic DNA-based sex determination markers in P. monodon may not be possible.

Manipulation of sex ratio and sexual maturation are important tools for crustacean aquaculture. Although sex determination in crustaceans has been reviewed and interested during the last two decades, the genetic basis for sex determination in decapods has not been studied.

Vertebrates have sex determination systems but do not conserve across distant related species. Sex determination systems consist of species illustrating sex chromosomes (e.g. XX/XY, XX/XO and ZZ/ZW systems) and those lacking sex chromosomes but sex differentiation of those species is controlled by autosomal genes (Baker et al., 1976). Sex determination systems in invertebrates completely differ from those of vertebrates because most invertebrates do not possess sex chromosomes.

The XX/XY system is found in mammals. Males possess heterogametic sex chromosomes (XY) and females exhibit homogametic sex chromosomes (XX). In contrast, the XX/XO system which is found in grasshoppers, crickets, roaches and some insects, males exhibit a single set of X chromosome (XO) but females possess homogametic sex chromosomes (XX). In addition, ZZ/ZW system is found in birds, some fishes and some insects including butterflies and moths. Sex heterogametic chromosomes (ZW) are found in females but sex homogametic chromosomes (ZZ) are found in males.

In insects, sex determination system has been well studied in *Drosophila melanogaster* because several spontaneous mutants affecting sex determination were found. The ratio of X chromosomes (X) to autosomes (A) is the primary signal for sex determination in this species. The gene *Sex-lethal* (*Sxl*) controls sex determination, dosage compensation, and oogenesis in *D. melanogaster* and can be activated only when the X/A was 1 or more and a female develops. Alternatively, a ratio of 0.5 (X: AA) leaves *sxl* inactivate and male development occurs.

Sxl participates in the female-specific splicing of its own pre-mRNA. The downstream target of Sxl is the *transformer* (*Tra*) which encodes a non-functional truncated Tra protein in males (Inoue et al., 1990). The female Tra protein induces female-specific splicing of the *doublesex* (*dsx*) pre-mRNA in cooperation with the Tra-2 gene product promoting female sexual development (Burtis et al., 1991; Jursnich and Burtis, 1993; An and Wensink, 1995).

The *dsx* gene is known as the final gene of the sex-determining cascade in *D. melanogaster*. Female- or male-specific proteins of *dsx* regulate the expression of sex-specific differentiation gene such as yolk protein genes (Burtis et al., 1991; Jursnich et

al., 1993; An et al., 1995a, b). The *dsx* proteins have a zinc finger-like domain called DM domain (Erdman et al., 1993; Erdman et al., 1996; Raymond et al., 1998).

The *dsx* homologues have been identified in many other species including mab-3 from *Caenorhabditis elegans* (Shen et al., 1988), DMRT1 from human (Raymond et al., 1998; Moniot et al., 2000) and *Dmrt*1 from mouse and chicken (Raymond et al., 1999; De Grandi et al., 2000). These *dsx* homologues were all contain DM domain and considered to regulate sexual differentiation and have been evolutionarily conserved as sex-determining genes.

Sex determination mechanisms have long been of major interest from both developmental and evolutionary points of view (Delvin and Nagahama, 2002). An understanding of sexual biology of any sexual-reproducing species is important for designing breeding programs in that species (Preechaphol et al., 2007). In *Drosophila*, sex determination is under the control of the sex lethal (*Sxl*) gene (Burtis et al., 1991).

In *M. rosenbergii*, implantation of androgenic glands in sexually undifferentiated females produced reproductively competent neomales. Mating of neomales with normal females suggested a complex heterogametic (ZW) system of females in this species.

In mammalian species, X-linked zinc figer proten gene (Zfx) and Y-linked zinc finger protein gene (Zfy) usually showed fixed single nucleotide polymorphism (SNP) between males and females. Previously, a homologue of zinc finger protein gene was isolated in *P. monodon*. The amplification product of this gene in females and males of *P. monodon* was characterized by sequencing. Unlike mammalian species, this gene homologue did not exhibit fixed polymorphism between genders (Leelatanawit, 2003).

In *P. monodon*, females exhibit approximately 10%–20% greater growth rate than do males (Browdy, 1998). The diploid chromosome numbers of penaeid shrimps have been reported in *P. esculentus*, *P. monodon*, *Farfantepenaeus aztecus*, *Fenneropenaeus chinensis*, *Fenneropenaeus merguiensis*, *Fenneropenaeus penicillatus* and *Marsupenaeus japonicus*, and (2N = 88), *P. semisulcatus* and *Litopenaeus setiferus* (2N = 90) and *Farfantepenaeus californiensis* and *Litopenaeus*

occidentalis (2N = 92) (Benzie, 1998). Neither sex chromosomes nor environmental sex determination have been reported in penaeid shrimps.

Sex chromosomes have not been cytologically identified in penaeid species. Recently, Li *et al*. (2003) constructed genetic linkage maps of the kuruma shrimp, *Masupenaeus japonicus*, based on AFLP analysis and revealed that sex of female progeny (*N* = 54) was tightly mapped to the linkage group 28 of the female map (LOD = 5.0) which led to the argument of female heterogamy (ZW) in this species. Moreover, triploidy affects the sex ratio in *Feneropenaeus chinensis* (Li *et al*., 2003) and *M. japonicus* (Preston *et al*., 2004) where the female-to-male ratio was almost 4:1 in the former but all triploids were female in the latter. These further support complex heterogametic sex in penaeid shrimp (Preston *et al*., 2004).

Wilson et al. (2002) used an identical approach to construct the male and female genetic linkage maps in *P. monodon*. A total of 673 polymorphic AFLP loci that confirmed to Mendelian segregation ratios were scored in three families and used to constructed separate male and female linkage maps for each family. Common markers found in two or more reference families were used to construct a common linkage map across three families. Nevertheless, sex-linked AFLP markers were not found in *P. monodon*. This indirectly implied that female-linked markers found in *M. japonicus* placed in the female map may be resulted artifact AFLP bands.

Staelens *et al*. (2008) constructed sex-specific high-density linkage maps and identified sex-linked markers for *P. monodon*. In total, 44 male and 43 female linkage groups from the analysis of 2306 AFLP markers segregating in three full-sib families, covering 2378 and 2362 cM, respectively. Twenty-one putatively homologous linkage groups, including the sex-linkage groups, were identified between the female and male linkage maps. Six sex-linked AFLP marker alleles were inherited from female parents in the three families, suggesting that the *P. monodon* adopts a WZ–ZZ sex-determining system. Two sex-linked AFLP markers, one of which was converted into an allele-specific assay, confirmed their association with sex in a panel of 52 genetically unrelated animals.

1.7 Spermatogenesis

Spermatogenesis is a complex cell differentiation process required a coordinated series of both mitosis and meiosis cycle events (Abe, 1987) and consists of a series of complex cellular events, in which different genes express to ensure the proper development of spermatozoa. The process of spermatogenesis follows an endocrine-regulated developmental program that features the transformation of an undifferentiated diploid stem cell into highly differentiated haploid spermatozoa.

In mammals, spermatogenesis is composed of three stages; the mitotic proliferation of spermatogonia, meiotic division of spermatocytes, and morphogenetic processes converting haploid spermatids to spermatozoa (Abe, 1987). Spermiation and sperm maturation occur during the final stage of spermatogenesis and are critical step for successful fertilization (Callard, 1991; Zirkin, 1993).

The mitotic proliferation of spermatogonia includes the germinal stem cells and other mitotic germinal cells produced from the stem cells (Grimes, 2004) and starts with the self-renewal and differentiation of a small population of spermatogonial stem cell. Spermatogonial stem cells are found in the basal part of the seminiferous epithelium, in contact with the basement membrane. They are also in close association with the nursing Sertoli cells, which produce the growth factors necessary to induce self-renewal and differentiation (Braydich-Stolle *et al*., 2007).

For the second stage, meiotic division of spermatocytes, DNA replication does not occur in spermatocytes but DNA repair is critical during this time period. Many unique genes are involved in the process of genetic recombination for example unique genes encode SCP1 and COR1 proteins are components of the synaptonemal complex, protein involved in recombination and DNA repair, and the *Dmc1* gene are all expressed in spermatocytes (Grim, 2004).

In addition, a targeted mutation of 70-kDa heat-shock gene Hsp70-2, which is expressed in the meiotic phase of spermatocytes in mice, leads to infertility. Development is arrested in late pachytene spermatocytes at the G2/M phase of the meiotic cell cycle. Hsp70-2 may be molecular chaperones required for Cdc2 activation that may facilitate dimerization of Cdc2 with cyclinB1 to become the active kinase in male germ cells (Eddy, 1999).

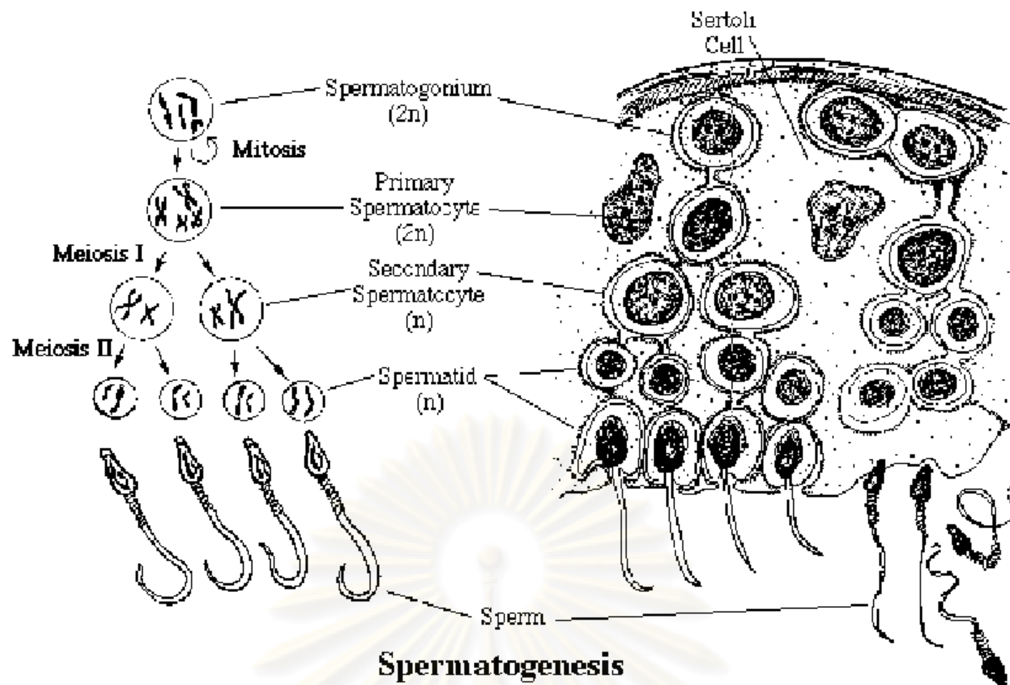


Figure 1.11 General diagram of spermatogenesis (http://www.luc.edu/faculty/wwasser/dev/spermeio.htm)

The third stage, called spermiogenesis, is morphogenetic processes converting haploid spermatids to mature spermatozoa. Many molecular events occur in spermatids that are required for completion of spermatogenesis. Significant progress has been made in understanding the unique chromatin remodeling and regulation of post-meiotic transcription in male germ cells that occurs during spermiogenesis (Sassone-Corsi, 2002). There is greatly increased transcriptional activity giving rise to several indispensable post-meiotic proteins in the early spermatids. For example, testis-specific isoforms of TATA-binding protein (TBP) are typically found (Sassone-Corsi, 2002). cAMP-responsive elements (CREs), members of the CREB family of transcription factors (Sassone-Corsi, 1998) are poorly expressed in testis, but another CREB family member, CREM, is present at high levels.

1.8 Histology of testes of penaeid shrimp

Histological studies of normal penaeid shrimp were reported (Bell and Lightner, 1988). A composition of male reproductive system is showed in Figures 1.12-1.14. As can be seen from Figure 1.12.1, the central core cells of the ventrally

directed vas deferens appear to have a highly granular, basophilic cytoplasm, with sparse nuclei (Nuc). The core cells are surrounded by a thin, fibrous, bi-layered sheath, with elongated nuclei (bold arrow); possibly representing the basal lamina and muscle layers found in older males. In addition, Figure 1.12.2 indicated that testes are paired organs straddling the median plane, and extend laterally around both sides of the hepatopancreas (Hep). In this view, a lateral lobe of the testes (Tes) is closely apposed by the hepatopancreas (Hep) proximally and muscle (Mus) peripherally. Its origin is near the posterior aorta (Aop).

From Figure 1.13.1, a distinct morphological evolution occurs in the vas deferens proceeding from the proximal to the distal region. The section depicts the vas deferens as having two separate lumens, a dorsal primary lumen and ventral secondary lumen. As can be seen Figure 1.13.2-1.13.4, a longitudinal septum (Sep) formed within the vas deferens dividing it into the primary (Lvp) and secondary lemen (Lvs). Formation of the longitudinal fold (Fld) increased the secondary channel surface area. The peripheral sheath (1.13.3; bold arrow) comprises fibrous and muscular components. The central region is the longitudinal septum (Sep), with the longitudinal fold (Sld) to the left and the primary lumen (Lvp) main wall to the right (1.13.4). The double layer of the longitudinal septum epithelial cells (Ept), and dividing connective tissue/hemal sinus layer (4; hollow arrow) might indicate an in-folding of the immature single-channelled vas deferens forming the septum.

From Figure 1.14.1, the spermatozoa, encased within a spermatophore (neither of which are present at this age), is transferred to the female during copulation from the terminal ampoule (Tam) to the genital pore (bold arrow) via the distal "terminal ampoule duct" (Tad). It appears that the "terminal ampoule duct" in juvenile males is merely an undeveloped portion of the terminal ampoule; the duct does not exist in adult males. The genital pore opens to the exterior on the coax of the fifth pereiopods (Pp5). Four distinct chambers have been reported from adult penaeids; they are discernable even in the small shrimp.

Morphology of shrimp with 2.0 cm in body length is slightly more complex than that noted in the younger shrimp (Fig. 1.14.1). This micrograph 1.14.2 shows only two of the four lumen chambers. The epithelium (Ept) typically consists of very tall columnar cells. A think, complex, fibrous and muscular wall (Wal) surrounds the

epithelium. The organ wall is composed of multiple layers if muscle (Msz), oriented in various directions, and connective tissue. The outer-most layer is reported to be a thin layer of squamous epithelium (Eps).

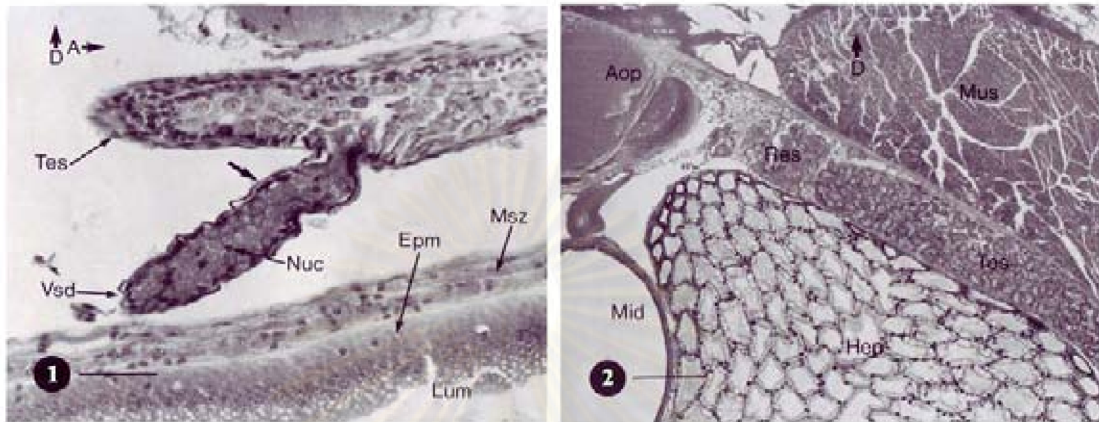


Figure 1.12 (1) Longitudinal section of a less than 1.0 gram juveniles showing the testes (Tes)/vas deferens (Vsd) junction. Msz: muscle layers, Epm: epithelium, and Lum: lumen. Longitudinal 4-5 μm paraffin section, H&E stain, Davidson's fixative, bar length = 50 μm. (2) Cross sectional orientation view of a laterally projecting testicular lobe (Tes) from a 2.0 gram juvenile. Transverse 4-5 μm paraffin section, H&E stain, Davidson's fixative, bar length = 400 μm. (Bell and Lightner, 1988)

In Fig. 1.14.3, the section illustrated the lumen (Lum) of only a single chamber. As in Fig. 1.14.2, the wall is composed of a single layer of very tall columnar eg., epithelial cells (Ept), and a thick layer of connective tissue and muscle (hollow arrow). In addition, the terminal ampoule had been excised from the shrimp prior to preparation (Fig. 1.14.4). The spermatophore, in this preparation, is in the very early development stages, as indicated by the small mass of spermatozoa (Sp4) located dorsally. This appears to depict multiple interconnecting chambers or lumina. These have been arbitrarily labeled #1 through #4. A cellular material (bold arrow) is noted within chamber #2. The large ventral mass (hollow arrow) contains a rather amorphous cellular material.

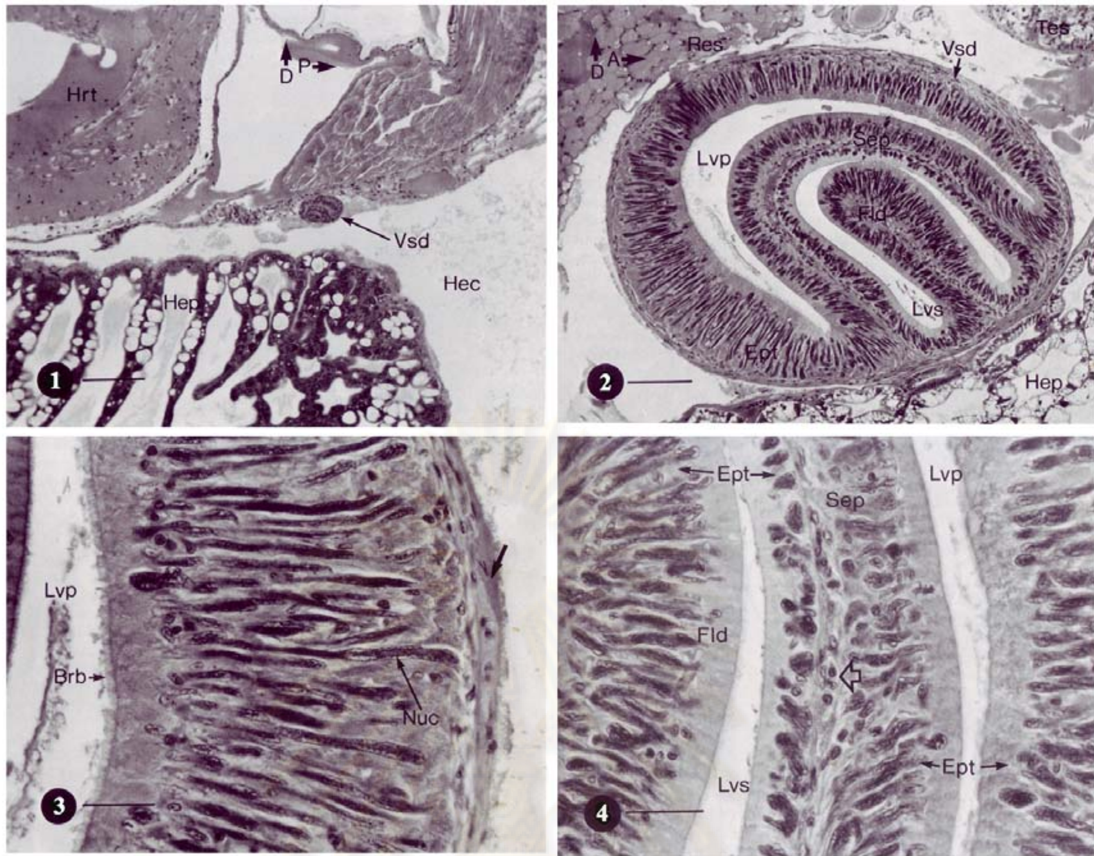


Figure 1.13 (1) Longitudinal orientation section from a less than 1.0 gram juvenile showing the proximal vas deferens (Vsd). Longitudinal 4-5 µm paraffin section, H&E stain, Davidson's fixative, bar length = 100 µm. (2-4) Longitudinal orientation section of the medial vas deferens (Vsd), from a 2.0 gram juvenile. Longitudinal 4-5 µm paraffin section, H&E stain, Davidson's fixative, bar length = 100 µm and 30 µm for (2) and (3-4), respectively. (Bell and Lightner, 1988)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

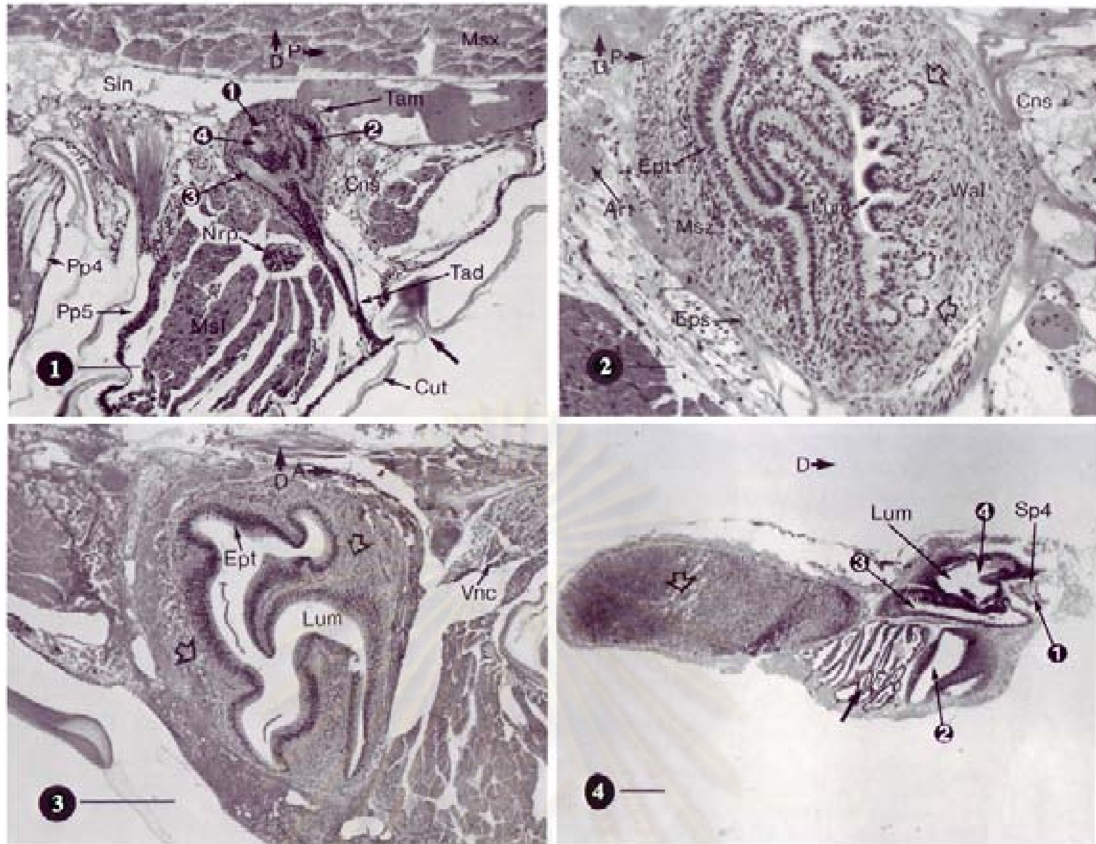


Figure 1.14 (1) Longitudinal orientation view of the enlarged distal vas deferens or terminal ampoule (Tam) from a less than 1.0 gram juvenile. Longitudinal 4-5 µm paraffin section, H&E stain, Davidson's fixative, bar length = 100 µm. (2) Longitudinal section of the terminal ampoule (Tam), (itself sectioned transversely or obliquely), from a less than 2.0 gram juvenile. Longitudinal 4-5 µm paraffin section, H&E stain, Davidson's fixative, bar length = 50 µm. (3) Longitudinal view of the terminal ampoule from a less than 2.0 gram juvenile. Longitudinal 4-5 µm paraffin section, H&E stain, Davidson's fixative, bar length = 400 µm. and (4) Orientation view of the terminal ampoule from a mature 35.0+ gram shrimp. Longitudinal 4-5 µm paraffin section, H&E stain, Davidson's fixative, bar length = 1 µm. (Bell and Lightner, 1988)

From Figure 1.15.1, each aluminal tubule or testicular cord (Tec) as they are more commonly called, like the lobe as a unit, is surrounded by a delicate fibrous connective tissue (Cnf) layer, of which a fibrocyte nucleus (Nuc) is noted. The cords are separated by hemal sinuses (Sin). Cellular differentiation has probably not yet began. A germinal layer of spermatogonia (Sp0) is typically localized on one side of

the cord's periphery. Spermatogonia contain a large granular nucleus with sparse cytoplasm. Spermatogonia are noted here to be dividing mitotically. Synchrony in cellular division is typical, note the large number of the mitotic figures (bold arrows). Nurse cells (Ncl) occupy a peripheral position.

From Fig. 1.15.2, the major differences between that of the shrimp sectioned in Fig. 1.15.1 are a) seminiferous tubules rather than the aluminal testicular cords are present, b) each tubule has a lumen and is significantly larger in size, c) each tubule contains a significantly higher number of cells, d) and the cells occupying the core of the tubule are primarily spermatids (Sp3), with possibly some primary (Sp1) and secondary spermatocytes (Sp2). As noted earlier in less mature shrimp, one side of the tubular periphery is composed of spermatogonia (Sp0), while the remainder of the periphery is composed primarily of nurse cells (Ncl).

Orientation view of mature spermatozoa (Fig. 1.15.3) arranged in bundles, each of which is presumably surrounded by an acellular "primary spermatophore layer". The bundles appear very similar to seminiferous tubules with spermatozoa (Sp4) in the central core and the periphery partially made up of presumed nurse cells (hollow arrows). The terminal ampoule has been excised from the shrimp prior to preparation (Fig. 1.15.4). The spermatophore, in this preparation, is in the very early developmental stage, as indicated by the small mass of spermatozoa (Sp4) located dorsally. This preparation appears to depict multiple interconnecting chambers or lumina. These have been arbitrarily labeled as #1 through #4. An acellular material (bold arrow) is noted within chamber #2. The large ventral mass (hollow arrow) contains a rather amorphous cellular material.

1.9 Candidate genes involved with testicular development and spermatogenesis of *P. monodon*

Several functional important genes possible related with growth, testicular development and spermatogenesis have been reported. An example of these genes is small ubiquitin-related modifier-1 (SUMO-1), a member of a ubiquitin-related protein family. Ubiquitin and its related proteins play important roles in diverse reproductive functions such as spermatogenesis and modulation of steroid receptor activity. SUMO-binding motif have been identified in several nuclear receptor including the

androgen receptor (AR), progesterone receptor (PR) and glucocorticoid receptor (GR) suggesting distinct roles of SUMO for growth and reproduction. Sumoylation is a posttranslational modification system that covalently attaches SUMO to target protein.

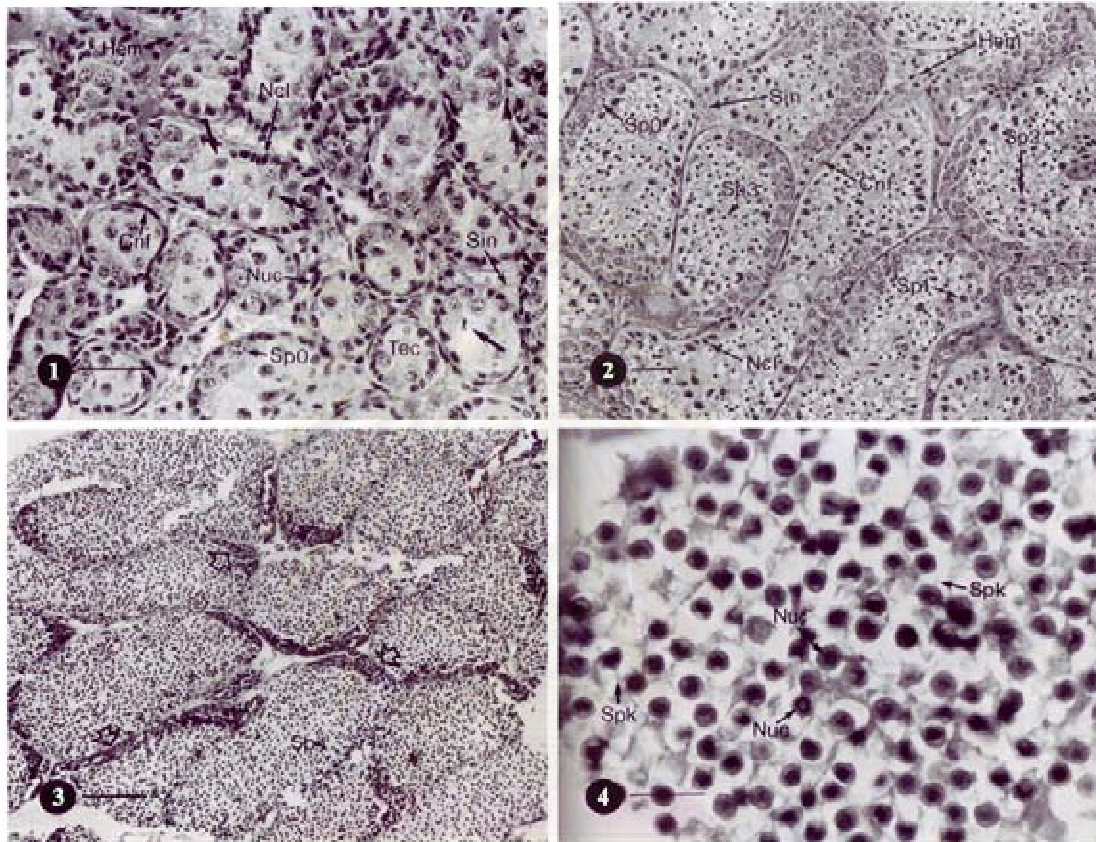


Figure 1.15 (1) Enlarged view of early testicular development stages from a 2.0 gram juvenile. Longitudinal 4-5 μm paraffin section, H&E stain, Davidson's fixative, bar length = 40 μm. (2) Enlarged view of mature testes from 15.0+ gram shrimp. Longitudinal 4-5 μm paraffin section, H&E stain, Davidson's fixative, bar length = 50 μm. (3) Orientation view of mature spermatozoa arranged in bundles, each of which is presumably surrounded by an acellular "primary spermatophore layer". The bundles appear very similar to seminiferous tubules with spermatozoa (Sp4) in the central core and the periphery partially made up of presumed nurse cells (hollow arrows). Longitudinal 4-5 μm paraffin section, H&E stain, Davidson's fixative, bar length = 100 μm. (4) Enlarged view of mature sperm of spermatozoa from terminal ampoule. Longitudinal 4-5 μm paraffin section, H&E stain, Davidson's fixative, bar length = 1 mm. (Bell and Lightner, 1988)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Meiosis is an indispensable process of sexual reproduction. However, detailed information on the regulatory mechanisms that initiate meiosis is not available. Progestins are important steroids regulating final maturation in male and female vertebrates. In male teleosts, it is known that progestin induces spermiation and sperm maturation. However, a role for progestin in early spermatogenesis or meiosis has not yet been described.

Recently, the functions of progestin on the initiation of meiosis in male Japanese eel were reported. A natural progestin in teleost fish 17α,20β-dihydroxy-4-pregnen-3-one (DHP) and its receptors were present in the testis at an early stage of spermatogenesis. By using an eel testicular culture system, DHP was shown to induce DNA replication in spermatogonia. Although 11-ketotestosterone, a known initiator of spermatogenesis, also stimulated DNA synthesis in spermatogonia, antibodies against DHP prevented DNA replication when added during the period in which meiosis was initiated. DHP treatment also induced the expression of meiosis-specific markers, such as Dmc1 and Spo11. Furthermore, Spo11 expression and synaptonemal complexes, specific features of the meiotic prophase, were detected in testicular fragments cultured with DHP in some germ cells that showed morphological characteristics of undifferentiated spermatogonia. Therefore, DHP is an essential factor for the initiation of meiosis (Miura *et al*., 2006).

Dmc1 (a RAC A-like recombinase) is known as a specific factor for meiotic recombination and has been identified as a gene product specifically expressed during the early meiotic prophase. Using the yeast two hybrid system, interaction between Ubc9 which is required as the enzyme for SUMO-1 conjugation to targets, and Dmc1 was illustrated in the basidiomycete *Coprinus cinereus* (Koshiyama, A. *et al*., 2006).

Recently, the A-type cyclins (cyclins A1 and A2) and Dmc1 cDNAs from the eel testis were cloned. Cyclin A1 mRNA was predominantly expressed in the livers, ovaries and testes of the eels. In contrast, a very high expression of cyclin A2 mRNA was observed in brains, livers, kidneys, spleens, ovaries, and testes of the eels. Dmc1 mRNA was predominantly expressed in the testes and ovaries; expression in the brain was also detected. In the eel testis, a few type-A spermatogonia incorporating 5-bromo-2-deoxyuridine (BrdU) were seen before the initiation of spermatogenesis by

hormonal induction. On day 1 after hormonal induction, the number of BrdU-labeled spermatogonia increased remarkably, and after 3 and 6 days, many labeled type-B spermatogonia were also observed. The expression of cyclin A2 increased 1 day after the induction of spermatogenesis and reached a plateau after 6 days, when many type-B spermatogonia with high proliferative activity were found. In contrast, the expression of cyclin A1 mRNA was detected after 9 days, coincident with the first appearance of spermatocytes. Cyclin A1 mRNA was localized in germ cells of all stages, from primary spermatocytes to round spermatids, whereas cyclin A2 mRNA was specifically localized in spermatogonia, secondary spermatocytes, round spermatids, and testicular somatic cells, including Sertoli cells. Dmc1 was localized only in the earlier stages of primary spermatocytes; before this stage, cyclin A1 mRNA was not detectable. Overall, cyclin A2, Dmc1, and cyclin A1 are expressed in spermatogenic cells sequentially before and during meiosis in the eel testis (Kajiura-Kobayashi *et al*., 2005).

Mitotic and meiosis cell divisions are fundamental processes of eukaryotes (Kobayashi et al., 1991). Progression through the cell cycle is regulated by association between cyclin-dependent kinases (Cdks) and their cyclin partners governing at different points in the cell division. Cyclin B-dependent cdc2 kinase activity has a key role in triggering the G2/M-phase transition during the mitotic and meiotic cell cycles.

The 70 kDa heat-shock proteins (HSP70) are molecular chaperones that assist other proteins in their folding, transport and assembly into complex. HSP70.2 is a testis-specific member of the HSP70 family, known to play a critical role in the completion of meiosis during male germ cells differentiation. Although abundantly present in post-meiotic cells, its function during spermiogenesis remained obscure. The HSP70-2 protein is synthesized during the meiotic phase of spermatogenesis and is abundant in pachytene spermatocytes. The knockout approach was used to determine whether HSP70-2 is a chaperone for proteins involved in meiosis. Male mice lacking HSP70-2 were infertile while females lacking HSP70-2 were fertile.

Spermatogenic cell development was arrested in prophase of meiosis I at the G2–M-phase transition and late pachytene spermatocytes were eliminated by apoptosis, resulting in an absence of spermatids. HSP70-2 is required for cdc2 to form a heterodimer with cyclin B1, suggesting that it is a chaperone necessary for the

progression of meiosis in the germ cells of male mice. HSP70-2 is also associated with the synaptonemal complex and desynapsis is disrupted in male mice lacking this protein. Homologues of HSP70-2 are present in the testes of many animals, suggesting that the role of this spermatogenic cell chaperone is conserved across phyla (Eddy, 1999).

A global proteomic approach was used to identify genome-organizing proteins in condensing spermatids and an unexpected role for HSPA2 (formerly HSP70.2), which acquires new functions and becomes tightly associated with major spermatid DNA-packaging proteins, Transition Proteins (TP) 1 and 2, were discovered. Therefore, HSPA2 is identified here as the first TP chaperone and these data provided the initial understanding on the yet totally unknown process of genome condensing structures assembly in spermatids (Govin *et al*., 2006).

Recently it has been indicated that estrogen, estradiol-17β (E2), is involved in regulating the renewal of spermatogonial stem cells in eels. Subtractive cDNA between testes of eel cultured with estradiol-17β (E+) and without (E-) were constructed to identify genes directly regulated this process. From northern blot analysis, eel spermatogenesis-related substance 34 (*eSRS34*) was expressed in testis cultured with E2 but was not expressed in that cultured with 11-ketotestosterone (11-KT, an androgen of teleosts, induces complete spermatogenesis, including spermatogonial prolifilation toward meiosis). Therefore, it was suggested that *eSRS34* was associated only with the regulation of spermatogonial stem cell renewal. The longest cDNA clone of *eSRS34* showed similarity with human platelet-derived endothelial cell growth factor (PD-ECGF). Recombinant protein of *eSRS34* was constructed and produced polyclonal antibody. Function of *eSRS34* was examined using several *in vitro* systems. Results showed that recombinant eSRS34 induced spermatogonial mitosis in testicular organ culture. Furthermore, the addition of an antibody specific for eSRS34 prevented spermatogonial mitosis induced by E2 stimulation in a germ cell/somatic cell co-culture system. Therefore, *eSRS34* concluded that is a spermatogonial stem cell renewal factor (Miura, 2007).

Prosaposin, the other reproductively related gene, is a precursor of four saposins, termed saposin A, B, C, and D, which activate glycosphingolipid hydrolysis. All four saposins contain six equally placed cysteines and a conserved *N*-

glycosylation site. The prosaposin gene contains 15 exons. It is transcribed into several mRNAs generated by alternative splicing of exon 8. Several functions of prosaposin were found such as in the nervous system, cancer development, and fertilization. Prosaposin was found to be important in development, maintenance and differentiation of male reproductive organs, spermatogenesis and fertilization such as human, mouse, rat, and chicken.

SGP-l/prosaposin can be secreted or targeted to the lysosomes where it is processed into smaller saposins (A, B, C, and D) required for the hydrolysis of glycosphingolipids. The deficiency of saposins B and C results in variant forms of metachromatic leukodystrophy and Gaucher's disease, respectively, which are characterized by lysosomal storage of undegraded glycosphingolipids. In the nervous system, prosaposin presents trophic activity.

A mouse model was recently developed by creating a null allele in embryonic stem cells through gene targeting to investigate the phenotypic diversity of prosaposin mutations and the involvement of this protein in lysosomal storage diseases, and for the development of therapeutic approaches. Mice homozygous mutants die at the age of 35–40 days and neurological disorders contribute to the early demise of the mutant mice. Male reproductive organs in homozygous mutants show several abnormalities, such as a decrease in testis size with reduced spermiogenesis and an involution of the prostate, seminal vesicles, and epididymis. In these animals, the blood levels of testosterone remain normal. In the prostate of homozygous mutants, only the basal epithelial cells appear to be present, while the secretory cells are absent. These findings suggest that prosaposin may be involved in the development and maintenance of the male reproductive organs, as well as, in cellular differentiation (Morales *et al*., 2000).

The level and cellular localization of fotillin-1, a lipid raft protein, was examined in the testis of rats during postnatal development and spermatogenesis. The testes of rats were sampled on postnatal days 7, 14, 21, 40, and 60, and analyzed by Western blot and immunohistochemistry. Western blot analysis detected flotillin-1 in the testes at days 7 and 14 after birth but the level decreased significantly at postnatal days 21, 40 and 60. At postnatal days 7, 14, 21, and 40, flotillin-1 immunolocalization was observed mainly in the Sertoli cells. However, there was little flotillin-1

immunolabeling in the spermatogenic cells from the seminiferous tubule of the testes. In the seminiferous tubule of the testes at postnatal day 60, flotillin-1 immunoreactivity in the Sertoli cells varied according to the stages of the spermatogenic cycle; intense immunoreactivity being observed in stages IX–III and less in stages IV–VIII. These results suggest that flotillin-1 participates in the developmental process of Sertoli cells and is involved in the regulation of spermatogenesis (Kim *et al*., 2008).

Different biotechnological approaches, for example; injection of vertebrate steroid hormones, neurotransmitters and ecdysteroids (Benzie, *et al*., 1998) and the use of specially formulated feed have been applied to induce the gonad maturation of penaeid shrimp but results are inconsistent owing to limited knowledge on genetic and hormonal control of penaeid species.

Biogenic amines (e.g serotonin or 5-HT, epinephrine and dopamine) and peptide neuroregulators are known to modulate the release of neuropeptide hormones from the sinus gland. Serotonin, epinephrine and dopamine cause hyperglycemic effects possibly due to the release of crustacean hyperglycemic hormone (CHH) from the sinus gland. Injections of Dopamine result in hyperglycemia in normal but not bilaterally eyestalk-ablated *P. monodon* (Kuo *et al*., 1995).

Dopamine injections (10^{-8}, 10^{-7}, or 10^{-6} mol/shrimp) affects levels of glucose, lactate, Na+, K+, Cl-, protein, and oxyhemocyanin in the hemolymph and its osmolality in juvenile *L. vanammei*. Elevations of hemolymph glucose and lactate was observed at between 2 and 4 hours whereas increases in hemolymph osmolality, Cl-, Na+, and total protein were found at 2 hours, and a reduction in hemolymph oxyhemocyanin was found at 4 hours after the dopamine injection. All physiological parameters except K+ had returned to the control values 8–16 hours after receiving the dopamine (Chiu *et al*., 2006).

Injections of serotonin and dopamine antagonist, spiperone (25 µg/g body weight + 1.5 or 5 µg/g body weight) induced ovarian maturation and spawning in wild *L. stylirostris* and pond reared *L. vannamei* (Alfaro *et al*., 2004). In addition, *in vivo* effects of dopamine, a dopaminergic antagonist (spiperone), and a dopaminergic agonist (ADTN) on maturation of the testes in the fiddler crab, *Uca pugilator*, were

determined. Dopamine inhibited testicular maturation dose-dependently. ADTN also inhibited maturation of the testes whereas spiperone induced testicular maturation. (Sarojini *et al*., 1999).

In newt (*Cynop pyrrhogaster*), apoptosis of spermatogonia are induced by prolactin. In spring to summer when the ambient temperature is mild, dopamine is released from hypothalamus to suppress prolactin release from the pituitary. The spermatogonia survive and undergo active spermatogenesis with the assistance of FSH. On the other hand, low temperature in fall to winter suppresses dopamine release resulting in prolactin release from the pituitary inducing spermatogonial death (Abe, 2004).

Several neurotransmitters and neuropeptides (e.g MCL1 and neuroparsin) that are found in heart are important for cardiac regulation, brain development, growth and reproduction in several species. The roles of neuropeptides for growth and reproductive regulation have been reported in gastropods (i.e. abalone) for examples; mutation of the MCL1 protein causes brain disorders (Schmitt *et al*, 2003), Neuroparsin A and B involved with brain development and reproductive physiology in gonads (Janssen *et al*, 2001).

1.10 Objectives

The objectives of this thesis were identification and characterization of genes involving testicular development and spermatogenesis of *P. monodon*. Expressed sequence tag (ESTs) analysis was carried out using cDNAs from a conventional testes library and from suppression subtractive hybridization libraries of *P. monodon*. The full length cDNA of various genes were isolated and characterized by RACE-PCR. In addition, the expression profiles of genes in juvenile and adult *P. monodon* were examined using RT-PCR, semi-quantitative RT-PCR, real-time PCR, or microarray analysis. Recombinant proteins of functionally important gene homologues were expressed *in vitro* and used for production of the polyclonal antibody.

CHAPTER II

MATERIALS AND METHODS

2.1 Experimental animals

Specimens used in this study were male broodstock of *P. monodon* collected from Satun (Andaman Sea, west). This group of samples was used for construction of conventional testes cDNA library and RACE-PCR. In addition, broodstock-sized of male and female *P. monodon* were also live-caught from Angsila, Chonburi (Gulf of Thailand, east) and juvenile *P. monodon* males and females (approximately 20 g body weight, 4-month-old) were purchased from local farms in Chachengsao, eastern Thailand and used for construction of suppression subtractive hybridization (SSH) cDNA libraries, RT-PCR or semiquantitative RT-PCR analyses. Male broodstock originated from the Andaman Sea were collected at different periods after molting. These specimens were used for semiquantitative PCR and real-time PCR analyses.

2.2 RNA extraction

Total RNA was extracted from ovaries and testes of each the shrimp using TRI REAGENT®. A piece of tissues was immediately placed in mortar containing liquid nitrogen and ground to the fine powder. The tissue powder was transferred to a microcentrifuge tube containing 500 µl of TRI REAGENT (1 ml/50-100 mg tissue) and homogenized. Additional 500 µl of TRI REAGENT were added. The homogenate and left for 5 min, before adding 0.2 ml of chloroform. The homogenate was vortexed for 15 s and left at room temperature for 2-15 min and centrifuged at 12000g for 15 min at 4 °C. The mixture was separated into the lower red phenol-chloroform phase, the interphase, and the colorless upper aqueous phase. The aqueous phase (inclusively containing RNA) was transferred to a new 1.5 ml microcentrifuge tube. RNA was precipitated by an addition of 0.5 ml of isopropanol and mixed thoroughly. The mixture were left at room temperature for 10-15 min and centrifuged at 12000g for 10 min at 4-25 °C. The supernatant was removed. The RNA pellet was washed with 1 ml of 75% ethanol and centrifuged at 12000g for 5 min at 4 °C. The ethanol was removed. The RNA pellet was air-dried for 5-10 min. RNA was dissolved in DEPC-

treated H_2O for immediately used. Alternatively, the RNA pellet was kept under absolute ethanol in a -80 °C freezer for long storage.

Total RNA was also extracted from other tissues including eyestalks, gills, heart, hemocytes, hepatopancreases, lymphoid organs, intestine, stomach, pleopods and thoracic ganglion of *P. monodon* using the same extraction procedure.

2.3 Measuring concentrations of extracted RNA by spectrophotometry and electrophoresis

The concentration of extracted RNA samples was estimated by measuring the optical density at 260 nanometer (OD_{260}). An OD_{260} of 1.0 corresponds to a concentration of 40 μg/ml single stranded RNA and 33 μg/ml single stranded DNA (Sambrook et al., 2001). Therefore, the concentration of RNA samples were estimated in μg/ml by using the following equation,

$$[RNA] = OD_{260} \text{ x dilution factor x } 40 \text{ (or 33 for single stranded DNA)}$$

The purity of DNA samples can be evaluated from a ratio of OD_{260} / OD_{280}. The ratios of appropriately purified RNA were 2.0 (Sambrook et al., 2001).

2.4 Construction of conventional cDNA libraries from heart and testes and suppression subtractive hybridization (SSH) cDNA libraries from different stages of testes of *P. monodon*

2.4.1 Purification of mRNA

Total RNA was extracted from testes of wild *P. monodon* caught from Satun (Andaman Sea, west) using TRI-REAGENT. Messenger (m) RNA was further purified using a QuickPrep *micro* mRNA Purification Kit (GE Healthcare).

Four hundred microliters of the extraction buffer were added to a microcentrifuge tube containing 25 μl of total RNA and mixed by pipetting. Two volume (0.8 ml) of the elution buffer was added and mixed thoroughly. The mixture was centrifuged at 16,000g for 1 min. Concurrently, the tube containing 1 ml of oligo(dT)-cellulose for each purification was centrifuged at the same speed for 1 min. The supernatant was removed. The homogenate was transferred into the

microcentrifuge tube containing the oligo(dT)-cellulose pellet. The tube was gently inverted to resuspend the oligo(dT)-cellulose for 3 min and centrifuged at 16000 g for 10 s at room temperature. The supernatant was carefully removed. The high salt buffer (1 ml) was added to a microcentrifuge tube and spun for 10 s at 16000g. The supernatant was carefully removed. The wash was repeated four more times, as described above. The low salt buffer (1 ml) was added to the oligo(dT)-cellulose pellet. The tube was inverted and spun at 16000g for 10 s. This wash was repeated once. The pellet from the final wash was resuspended in 0.3 ml of the low salt buffer. The slurry was transferred to a MicroSpin column and spun for 5 s. The flow-though solution was discarded. The low salt buffer (0.5 ml) was added and further spun for 5 s. This step was repeated twice. The column was then placed into a sterile 1.5 ml microcentrifuge tube and briefly centrifuged. The mRNA was eluted out by an addition of 0.2 ml of the prewarmed elution buffer (55°C) to the top of column and centrifuged at 16000g for 5 s. Additional 0.2 ml of the prewarmed elution buffer was added to the top of column to elute residual mRNA and centrifuged as described above.

2.4.2 Construction of the conventional testis and heart cDNA libraries of *P. monodon*

The typical cDNA libraries were carried out using a cDNA Synthesis Kit, ZAP-cDNA Synthesis Kit, and ZAP-cDNA Gigapack III Gold Cloning Kit (Clontech).

2.4.2.1 First and second strand cDNA synthesis

Five micrograms of testis or heart mRNA was combined with 5 µl of 10x first-strand buffer, 3 µl of first-strand methyl nucleotide mixture (10 mM dATP, dGTP, and dTTP supplemented with 5 mM 5-methyl dCTP), 2 µl of linker-primer (1.4 µg/µl), and 1 µl of (40 units/µl) RNase Block Ribonuclease Inhibitor. DEPC-treated water was added to a final volume of 50 µl. The reaction was gently vortexed and briefly centrifuged. The reaction was incubated at room temperature for 10 min and 1.5 µl of (50 units/µl) StrataScript RT was added to the first strand synthesis reaction. The reaction was gently vortexed and briefly centrifuged. Five microliters of the first-strand synthesis reaction was transferred to a separate tube and served as the first-

strand synthesis control reaction. The reaction was incubated at 42 °C for 1 hr. The tubes were placed on ice to terminate the first strand cDNA synthesis.

The components of second-strand synthesis reaction including 20 µl of 10x second-strand buffer, 6 µl of second-strand dNTP mixture (10 mM dATP, dGTP, and dTTP plus 26 mM dCTP), 116 µl of Sterile distilled water, 2 µl of (1.5 units/µl) RNase H, and 11 µl of (9.0 units/µl) DNA polymerase were added to the first-strand synthesis reaction on ice. The reaction was gently vortexed, briefly centrifuged, and incubated at 16 °C for 2.5 hr. After second-strand synthesis reaction, immediately placed the reaction tube on ice.

2.4.2.2 Bunting the cDNA termini

Twenty-three microliters of blunting dNTP mix and 2 µl of (2.5 units/µl) cloned *Pfu* DNA polymerase were added to the second-strand synthesis reaction. The reaction was gently vortexed, briefly centrifuged, and incubated at 72 °C for 30 min.

An equal volume (200 µl) of phenol-chloroform [1:1 (v/v)] was added. The mixture was vortexed and centrifuged at 14000g for 2 min at room temperature. The upper aqueous layer was transferred to a new tube. An equal volume of chloroform: isoamyl alcohol (24:1) was added and vortexed. The mixture was centrifuged at 14000g for 2 min at room temperature and the upper aqueous layer was transferred to a new tube. The cDNA was precipitated by adding 20 µl of 3M sodium acetate and 400 µl of absolute ethanol. The reaction was vortexed and kept overnight at -20 °C.

The synthesized cDNA was recovered by centrifugation at 14000g for 60 min at 4 °C. The cDNA pellet was gently washed by adding 500 µl of 75% (v/v) ethanol to the side of the tube away from the precipitate and centrifuged at 14000g for 2 min at room temperature. The pellet was air-dried, resuspended in 9 µl of *Eco*R I adapters, and incubated at 4 °C for at least 30 min.

2.4.2.3 Ligation of *Eco*R I adapters

One microliter of 10x ligase buffer, 1 µl of 10 mM rATP, and 1 µl of T4 DNA ligase (4 units/µl) were added to the tube containing the blunted cDNA and the *Eco*R I adapters. The reaction was centrifuged and incubated overnight at 8 °C or at 16 °C

for 2 days. After ligation reaction, the ligase activity was heat-inactivated at 70 °C for 30 min. The reaction was centrifuged for 2 min at room temperature and cooled at room temperature for 5 min.

2.4.2.4 Phosphorylation of *Eco*R I ends

The components including 1 µl of 10x ligase buffer, 2 µl of 10 mM rATP, 5 µl of sterile water, and 2 µl of T4 polynucleotide kinase (5 units/µl) were added to the reaction and incubated at 37 °C for 30 min. The kinase activity was heat-inactivated at 70 °C for 30 min. The reaction was centrifuged for 2 min at room temperature and cooled at room temperature for 5 min.

2.4.2.5 *Xho* I digestion

Twenty-eight microliters of *Xho* I buffer and 3 µl of *Xho* I (40 units/µl) were added. The reaction was incubated at 37 °C for 1.5 hr. Digested cDNA was precipitated by adding 5 µl of 10x STE buffer (1M NaCl, 200 mM Tris-HCl, pH 7.5, and 100 mM EDTA) and 125 µl of absolute ethanol and incubated overnight at -20°C. Following precipitation, the reaction was centrifuged at 14000g for 60 min at 4°C. The pellet was dried, resuspended in 14 µl of 1x STE buffer and 3.5 µl of the column loading dye (50% (v/v) glycerol, 10% (v/v) 10x STE buffer, and 40% (w/v) saturated bromophenol blue). The resuspended sample was ready to be size-fractionated through a drip column containing sepharose CL-2B gel filtration medium.

2.4.2.6 Size fractionation

2.4.2.6.1 Assembling and loading the drip column

A sterile 1-ml pipette was used as column. Small piece (about 8 millimeters) of the connecting tubing was used to connect the 1 ml pipette to the 10 ml syringe (Fig 2.1). Prior to loading the drip column, sepharose CL-2B gel filtration medium was gently mixed until the resin was uniformly suspended and 2 ml of 1x STE buffer (100 mM NaCl, 20 mM Tris-HCl pH 7.5, and 100 mM EDTA) and added to a column until the surface of the packed bed was approximately 0.25 inch below the "lip of the pipette" (the point where the pipet and the syringe were joined). The drip column was washed with 10 ml of 1x STE buffer. When about 50 µl of the STE buffer remained

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

above the surface of the resin, cDNA sample was immediately loaded using an automatic pipette.

2.4.2.6.2 Collecting the sample fractions

Sepharose CL-2B gel filtration medium separated molecules on the basis of size. Three drops per fraction were collected to a fresh microcentrifuge tube. The fractions began to collect when the leading edge of the dye reached the -.4 ml graduation of the pipette (Fig 2.1), and stopped when trailing edge of the dye reached the -.3 ml graduation. Before processing the fractions and recovering the size-fractionated cDNA, 8 μl of each collected fraction was electrophoresed in a 5% nondenaturing polyacrylamide gel to determine the appropriate fractions used for ligation.

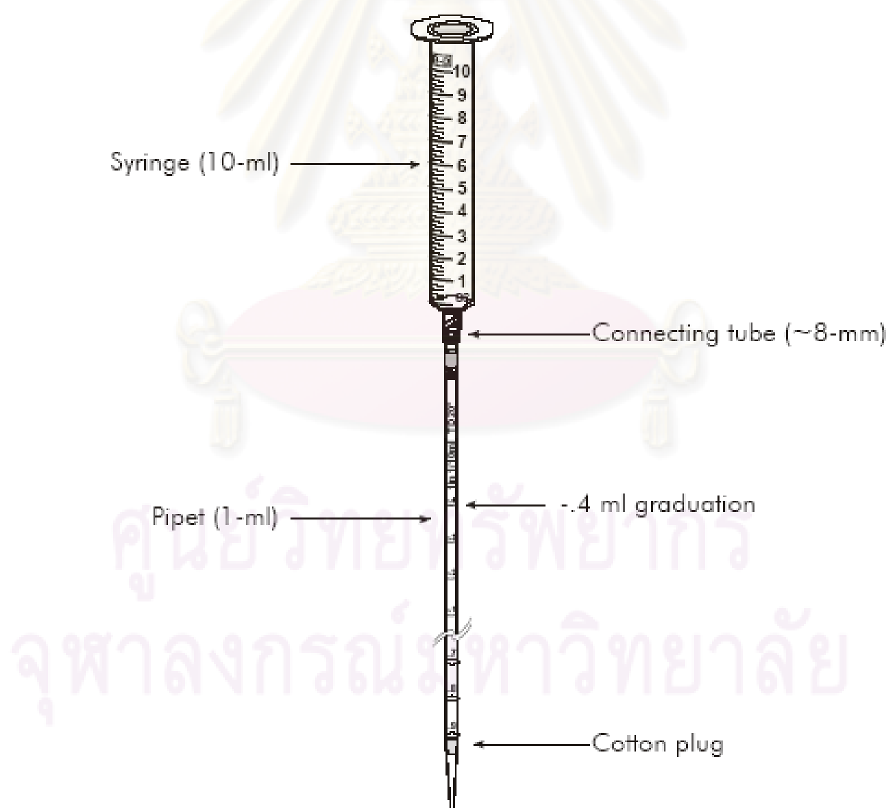


Figure 2.1 Assembly of the drip column for size-fractionation of synthesized cDNA

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.4.2.7 Processing the cDNA fraction

The collected fractions that did not contain linkers (>500 bp frgments) were selected and extracted with an equal volume of phenol-chloroform [1:1 (v/v)], vortexed, and spun in a microcentrifuge at the maximum speed for 2 min at room temperature. The upper aqueous layer was transferred to a fresh microcentrifuge tube. An equal volume of chloroform:isomyl alcohol (24:1) was added, vortexed, and spun in a microcentrifuge at the maximum speed for 2 min at room temperature. Two volume of absolute ethanol was added to extracted sample and cDNA was precipitated overnight at -20°C. The sample was centrifuged at the maximum speed for 60 min at 4°C. The pellet was carefully washed with 200 µl of 80% (v/v) ethanol. The sample was centrifuged at the maximum speed for 2 min at room temperature. The pellet was air-dried and resuspended in 3.5-5 µl of sterile water.

2.4.2.8 Ligation of the cDNA insert

Size-selected cDNAs (>500 bp) were ligated into dephosphorylated *Eco* RI/*Xho* I-digested UNI-ZAP XR. The ligation reaction contained 100 ng of resuspended cDNA, 0.5 µl of 10x ligase buffer, 0.5 µl of 10 mM rATP (pH 7.5), 1.0 µl of the predigested Uni-ZAP XR vector (1 µg), sterile water for a final volume of 4.5 µl, and 0.5 µl of T4 DNA ligase (4U/µl). The ligated reaction was incubated overnight at 12 °C or for up to 2 days at 4 °C.

2.4.2.9 Packaging and titering

ศูนย์วิทยทรัพยากร

2.4.2.9.1 Packaging

จุฬาลงกรณ์มหาวิทยาลัย

The packaging extract was removed from a -80 °C freezer and placed the on ice. Three microliters (containing 0.1-1.0 µg) of the ligation product was added to the packaging extract, gently stirred the tube with a pipette, and briefly centrifuged. The tube was incubated at room temperature for 2 hr and 500 µl of SM buffer and 20 µl of chloroform were added to the tube and mixed gently. The tube was centrifuged briefly and the supernatant was transferred to a newly sterile tube. The supernatant containing the phage was ready for tittering. The supernatant may be stored at 4 °C for up to 1 month.

2.4.2.9.2 Titering of the primary library

E. coli XL1-Blue MRF′ cells were cultured in LB broth with supplements (0.2% (w/v) maltose and 10 mM $MgSO_4$) with shaking at 37 °C for 4-6 hr (OD_{600} < 1.0), or overnight at 30 °C. The bacterial cells were centrifuged at 1000 g for 10 min, and the cell pellet was resuspended in 25 ml of sterile 10 mM $MgSO_4$. After OD determination, the culture was diluted to an OD_{600} of 0.5 with sterile 10 mM $MgSO_4$.

To determine the titer of the primary library, 1 μl of the final packaged reaction and 200 μl of *E. coli* XL1-Blue MRF′ cells at an OD_{600} of 0.5. In addition, 1 μl of a 1:10 dilution of the final packaged reaction was also combined with 200 μl of *E. coli* XL1-Blue MRF′ cells. The phage/bacteria mixture was incubated at 37 °C for 15 min to allow the phage to attach to the cells. The component was added into 3 ml of the melted NZY top agar that was pre-cooled to approximately 48°C, and plated immediately onto dry, the prewarmed NZY agar plate. The plate was incubated at 37°C about 6-8 hr. The plaques were counted and the titer of the library was estimated and expressed in plaque-forming units per milliliter (pfu/ml).

2.4.2.10 Amplification of the library

The *E. coli* XL1-Blue MRF′ cells were overnight cultured in LB broth with supplements at 30 °C with shaking. The cells were gently centrifuged and resuspended in 25 ml of 10 mM $MgSO_4$. The cell suspensions were measured at 600 mM and then diluted to an OD_{600} of 0.5 using 10 mM $MgSO_4$.

Aliquots of the packaged mixture containing about 5 x 10^4 pfu of bacteriophage (< 300 μl of phage) with 600 μl of *E. coli* XL1-Blue MRF′ cells at OD_{600} of 0.5 were combined in polypropylene tubes. To amplify 1 x 10^6 plaques, a total of 20 aliquots were combined (each aliquot contained 5 x 10^4 plaques/150-mm plate).

Each aliquot tube was incubated for 15 min at 37 °C and 6.5 ml of the NZY top agar, melted and cooled to 48 °C, was mixed with each aliquot of infected bacteria and evenly spread onto a 150-mm NZY agar plate. Plates were incubated at 37 °C for 6-8 hr. After that plates were overlaid with 5-10 ml of SM buffer and stored at 4 °C overnight.

The bacteriophage suspension from each plate was recovered and pooled into a sterile polypropylene container. Each plate was rinsed with additional 2 ml of SM buffer and pooled. Chloroform was added to a 5% (v/v) final concentration, mixed well, and incubated for 15 min at room temperature. The cell debris was removed by centrifugation for 10 min at 500 g. Chloroform was added to a final concentration of 0.3% (v/v). Aliquots of the amplified library were stored in 7% (v/v) DMSO at -80 °C

2.4.2.11 *In vivo* excision to convert the lambda library to the phagemid library

The lambda library was converted into the pBluescript library by *in vivo* excision. *E. coli* XL1-Blue MRF′ and SOLR cells were overnight cultured in LB broth with supplements at 30°C with shaking. The cells were gently centrifuged and resuspended in 25 ml of 10 mM $MgSO_4$. The OD_{600} of the cell suspensions was measured and then diluted to an OD_{600} of 1.0 (8 x 10^8 cells/ml) using 10 mM $MgSO_4$.

A portion of the amplified lambda bacteriophage library was combined with *E. coli* XL1-Blue MRF′ cells in a 50-ml conical tube at a MOI of 1:10 lambda phage-to-cell ratio. It is recommended to excise 10- to 100-fold more lambda phage than the size of the primary library to ensure statistical representation of the excised clones and ExAssist helper phage was added at a 10:1 helper phage-to-cell ratio to ensure that every cell was co-infected with lambda phage and helper phage.

A mixture containing 10^7 pfu of the lambda phage, 10^8 cells of *E. coli* XL1-Blue MRF′, and 10^9 pfu of ExAssist helper phage were combined in a conical tube and incubated at 37C° for 15 min. After that, 20 ml of LB broth with supplements was added and the conical tube was incubated at 37C° for 2.5-3.0 hr with shaking. The conical tube was heated at 65-70 °C for 20 min to lyse the lambda phage particles and the cells, centrifuged at 1000 g for 10 min, and the supernatant was transferred into a sterile conical tube. One microliter of the supernatant containing excised phagemids was combined with 200 µl of *E. coli* SOLR cells in a 1.5 ml microcentrifuge tube, and incubated at 37C° for 15 min. The cell mixtures were plated onto LB ampicillin agar plates (50 µg/ml) and incubated overnight at 37 C°.

2.4.2.12 Colony PCR

Recombinant clones were selected by a lacZ′ system following standard protocols (Sambrook and Russel, 2001). Colony PCR was performed to identify sizes of positive clones. Recombinant clones carrying insert sizes greater than 500 were extracted from testis and heart cDNA library.

Colony PCR was performed in a 25 ul reaction mixture containing 10 mM Tris-HCl, pH 8.8, 50 mM KCl, 0.1% Triton X - 100, 100 mM of each dNTP, 2 mM $MgCl_2$, 0.1 μM each of pUC1 (5′-CCG GCT CGT ATG TTG TGT GGA-3′) and pUC2 (5′-GTG GTG CAA GGC GAT TAA GTT GG-3′), 0.5 unit of Dynazyme™ DNA Polymerase (FINNZYMES, Finland). An interesting colony was picked by a pipette tip and served as the template in the reaction, PCR was carried out in a thermocycler consisting of predenaturation at 94 °C for 3 min followed by 35 cycles of denaturation at 94 °C for 30 s, annealing at 50 °C for 1 min and extension at 72 °C for 2 min. The final extension was carried out at the same temperature for 7 min. The colony PCR products were electrophoresed though 1.2% agarose gel and visualized after ethidium bromide staining.

2.4.2.13 Plasmid DNA extraction

Plasmid DNA was isolated using a HiYield™ Plasmid Mini Kit (RBC; Real Biotech Corporation). A recombinant clone was inoculated into 3 ml of LB broth (1% tryptone, 0.5% yeast extract, 1.0 % NaCl) containing 50 μg/ml of ampicillin and incubated at 37°C with constant shaking at 250 rpm overnight. The culture was transferred into 1.5 ml microcentrifuge tube and centrifuged at 14,000 rpm for 1 min. The supernatant was discarded. The bacterial cell pellet was collected and resuspended with 200 μl of the PD1 buffer containing RNaseA and thoroughly mixed by vortexed. The resuspended cells were lysed by the addition of 200 μl of the PD2 buffer and mixed gently by inverting the tube 10 times. The mixture was stood for 2 min at room temperature. After that, 300 μl of the buffer PD3 was added to neutralize the alkaline lysis step and mixed immediately by inverting the tube for 10 times. To separate the cell debris, the mixture was centrifuged at 14,000 rpm for 15 min. The supernatant was transferred into a new microcentrifuge tube and to the PD column and centrifuged at 6,000*g* (8,000 rpm) for 1 min. The flow-through was discarded.

The PD column was placed back in the collection tube. The column was washed by adding 400 µl of the W1 buffer and centrifuged at 6,000*g* (8,000 rpm) for 1 min. After discarding the flow-through, 600 µl of the ethanol-added Wash buffer was added and centrifuged as above. The flow-through was discarded. The spin tube was centrifuge for an additional 2 min at full speed (14,000 rpm) to remove the residual Wash buffer. The dried PD column was placed in a new 1.5 ml microcentrifuge tube and 30-50 µl of the Elution buffer or water was added at the center of the column to elute the extracted plasmid DNA. The column was left at room temperature for 2 min and centrifuge at 14,000 rpm for 2 min. The concentration of extracted plasmid DNA was spectrophotometrically measured.

2.4.2.14 Sequencing, EST clustering and assembly

A total of 896 from the testis cDNA library and 412 clones from the heart cDNA library were randomly selected and unidirectional sequenced from the 5' direction using a MegaBase 1000 automated DNA sequencer (GE Healthcare). Nucleotide sequences of ESTs were compared with those previously deposited in the GenBank using BLAST*N* and BLAST*X* (Altschul *et al*., 1990, available at http://www.ncbi.nlm.nih.gov). Significant matched nucleotides/proteins were considered when the E-value was $<10^{-4}$. Clustering and assembling of sequences were performed using TIGR Gene Indices Clustering Tools (TGICL) (Pertea *et al*., 2003) with CAP3 (Huang and Madan, 1999).

2.4.3 Construction of suppression subtractive hybridization (*SSH*) cDNA libraries

Forward and reverse subtractions (Diatchenko et al., 1996) from testes cDNAs of broodstock and juvenile of *P. monodon* were carried out using a PCR Select cDNA Subtraction Kit (Clontech, USA).

2.4.3.1 First strand cDNA Synthesis

One microgram of each tester and driver mRNA was combined with 1 µl of 10 µM of the cDNA synthesis primer in a sterile microcentrifuge tube. Sterile H_2O was added to a final volume of 5 µl. The reaction mixture was mixed by pipetting and centrifuged briefly. The reaction tube was incubated at 70 °C in a thermal cycler for 2

min. The tube was cooled on ice for 2 min and briefly centrifuged. The first strand cDNA synthesis was synthesized by adding 2 µl of 5x First-Strand Buffer (250 mM Tris-HCl, pH 8.5, 40 mM $MgCl_2$, 150 mM KCl and 5 mM dithiothreitol), 1 µl of dNTP Mix (10 mM each of dNTPs), 1 µl of sterile H_2O, and 1 µl of (20 units) AMV Reverse Transcriptase. The reaction was gently vortexed and briefly centrifuged. The reaction was incubated at 42 °C for 1.5 hr in an air incubator. The tubes were placed on ice to terminate the first strand cDNA synthesis.

2.4.3.2 Second strand cDNA synthesis and purification

The first strand synthesis reaction mixture was combined with 48.4 µl of Sterile H_2O, 16.0 µl of 5x second strand buffer (100 mM Tris-HCl, pH 7.5, 25 mM $MgCl_2$, 500 mM KCl, 50 mM ammonium sulfate, 0.75 mM β-NAD and 0.25 mg/ml BSA), 1.6 µl of dNTP mix (10 mM) and 4.0 µl of 20x second strand enzyme cocktail (DNA polymerase I, 6 units/µl; RNase H, 0.25 units/µl; and *E. coli* DNA ligase, 1.2 units/µl), mixed and briefly spun. The reaction was incubated at 16 °C for 2 hr. After that, 2 µl of T4 DNA Polymerase (3 units/µl) were added to the second strand reaction mixture and further incubated at 16 °C for 30 min. Then, 4 µl of 20x EDTA / Glycogen (0.2 mM EDTA; 1 mg/ml glycogen) was added to the reaction and mixed thoroughly. Then, 100 µl of phenol:chloroform:isoamyl alcohol (25:24:1) were added and vortexed. The mixture was centrifuged at 14,000 rpm for 10 min at room temperature. The top aqueous layer was carefully transferred to a sterile microcentrifuge tube. The phenol extraction was repeated. To precipitate the synthesized cDNA, 40 µl of 4 M NH_4OAc and 300 µl of absolute ethanol were added and thoroughly mixed by vortexing and centrifuged at 14,000 rpm for 20 min at room temperature. The supernatant was carefully removed. The pellet was overlaid with 500 µl of 80% ethanol and centrifuged at 14,000 rpm for 10 min. The pellet was air dried for 10 min to evaporate residual ethanol. The pellet was dissolved in 50 µl of sterile H_2O.

2.4.3.3 *Rsa* I Digestion

The restriction mixture containing 43.5 µl of double strand cDNA of tester and driver, 5.0 µl of 10x *Rsa* I restriction buffer (100 mM Bis Tris Propane-HCl (pH 7.0), 100mM $MgCl_2$ and 1 mM DTT) and 1.5 µl of *Rsa* I (10 units/µl) was set up. The

Table 2.1 Sequences of the PCR select cDNA synthesis primer, adaptors and PCR primers

Primer/Adaptor	Sequence
cDNA synthesis primer	5′- TTTTGTACAAGCTT$_{30}$N$_1$ n -3′
Adaptor 1	5′- CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT -3′ 3′- GGCCCGTCCA -5′
Adaptor 2R	5′-CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT-3′ 3′-GCCGGCTCCA-5′
PCR primer 1	5′- CTAATACGA CTCCATATAGGGC -3′
Nested PCR primer 1	5′- TCGAGCGGCCGCCCGGGCAGGT -3′
Nested PCR primer 2R	5′- AGCGTGGTCGCGGCCGAGGT -3′
Control Primers:	
G3PDH 5′ Primer	5′- ACCACAGTCCATGCCATCAC -3′
G3PDH 3′ Primer	5′- TCCACCACCCTGTTGCTGTA -3′

digestion was incubated at 37 °C for 1.5 hr. At the end of the incubation time, 5 µl of the digest were collected for analysis of *Rsa* I digestion efficiency by agarose gel electrophoresis. After that, 2.5 µl of 20x EDTA/glycogen mix was added to terminate the reaction. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was added and vortexed. The reaction tube was centrifuged at 14,000 rpm for 10 min at room temperature. The top aqueous layer was carefully removed and placed in a sterile microcentrifuge tube. The phenol/chloroform extraction was repeated. To recover the digested cDNA, 25 µl of 4 M NH_4OAc and 187.5 µl of absolute ethanol were added. The tube was thoroughly vortexed and centrifuged at 14,000 rpm for 20 min at room temperature. The supernatant was carefully removed. The pellet was briefly washed with 200 µl of 80% ethanol and centrifuged at 14,000 rpm for 5 min. The pellet was air dried for about 10 min and dissolved in 5.5 µl of sterile H_2O and stored at -20 °C. This digested cDNA served as the experimental driver cDNA. Subsequently, the sample was ligated with adapters to create tester cDNAs for forward and reverse subtraction.

2.4.3.4 Adaptor Ligation

One microliter of *Rsa* I-digested tester cDNA was diluted with 5 µl of sterile H_2O. A ligation master mix was prepared by combining 3 µl of sterile H_2O, 2µl of 5x ligation

buffer (250 mM Tris-HCl; pH 7.8, 50mM $MgCl_2$, 10 mM DTT and 0.25 mg/ml BSA) and 1 µl of T4 DNA ligase (400 units/µl). For each experimental tester cDNA, the reagents were combined in Table 2.7 in order,

Table 2.2 Ligation reactions of cDNA of testes (tester 1-1 and 1-2) and testes (tester 2-1 and 2-2)

Component	cDNA for ovaries		cDNA for testes	
	Tube 1 Tester 1-1 (µl)	Tube 2 Tester 1–2 (µl)	Tube 3 Tester 2–1 (µl)	Tube 4 Tester 2–2 (µl)
Diluted tester cDNA	2	2	2	2
Adaptor 1 (10 µM)	2	-	2	-
Adaptor 2R (10 µM)	-	2	-	2
Master Mix	6	6	6	6
Final volume	10	10	10	10

* Tester 1-1 and 1-2 and 2-1 and 2-2 were mixed and subsequently served as unsubtracted tester control 1-C and 2-C, respectively.

The tubes were briefly centrifuged and incubated at 16 °C overnight. The ligation reactions were stopped by adding 1 µl of the EDTA/glycogen mix, heated at 72 °C for 5 min to inactivate the ligase activity.

2.4.3.5 First Hybridization

For each experimental subtraction, the reaction mixture was prepared in order according to Table 2.8. A drop of mineral oil and centrifuged briefly. The samples were incubated in a thermal cycler at 98 °C for 1.5 min. The first hybridization

Table 2.3 Composition of the first hybridization reaction of each subtraction

Component*	Hybridization Sample 1 (µl)	Hybridization Sample 2 (µl)
Rsa I – digested driver cDNA	1.5	1.5
Adaptor 1 – ligated Tester 1 - 1	1.5	-
Adaptor 2R – ligated Tester 1 - 2	-	1.5
4x Hybridization Buffer	1.0	1.0
Final volume	4.0	4.0

* Tester 2-1 and 2-2 (testes) were identical

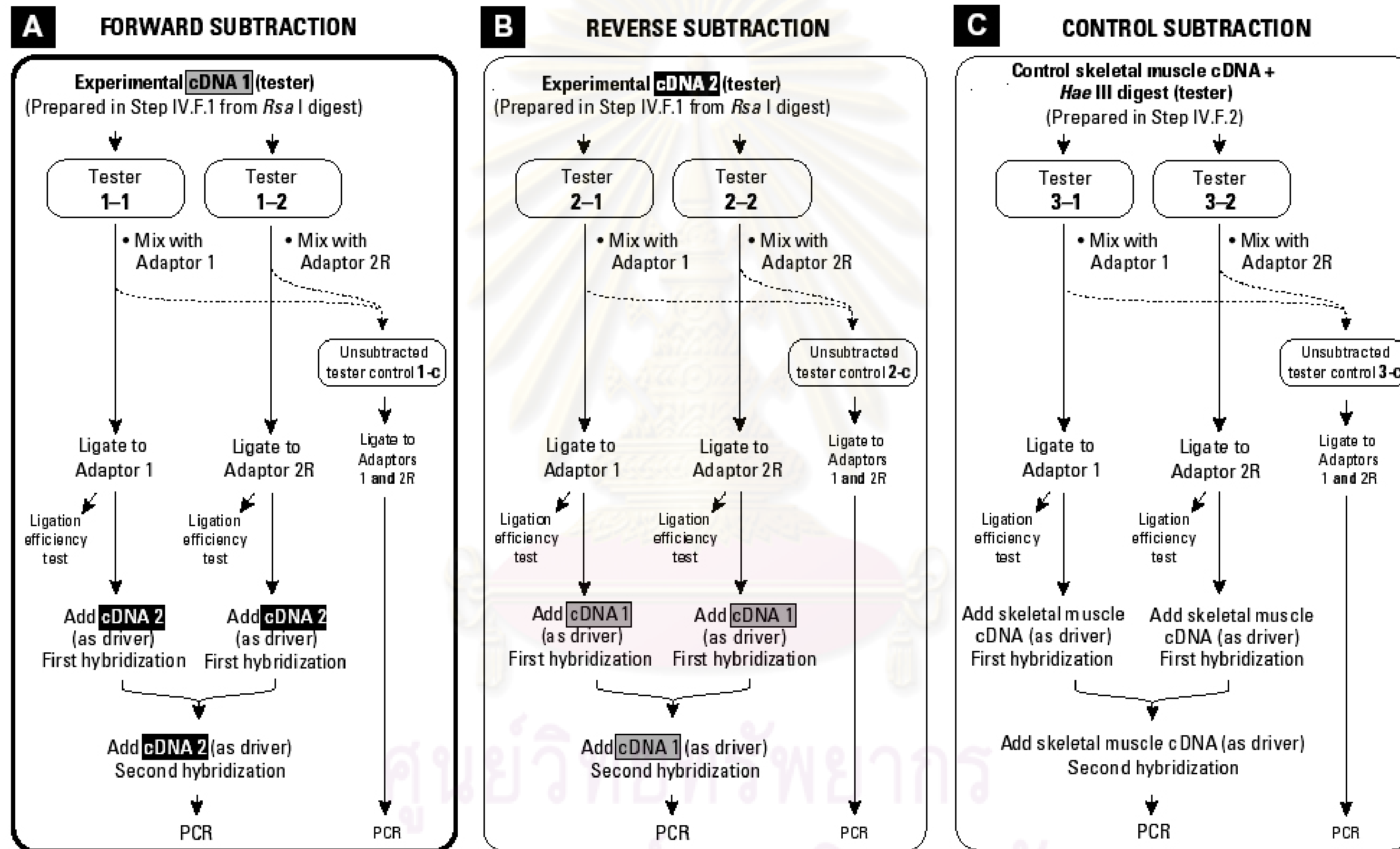


Figure 2.2 Preparation of adaptor-ligated tester cDNA for hybridization and PCR

was carried out at 68 °C for 8 hr and proceeded immediately to the second hybridization step.

2.4.3.6 Second Hybridization

To carry out the second hybridization: 1 µl of *Rsa* I – digested driver cDNA, 1 µl of 4x Hybridization Buffer and 2µl of sterile H_2O were added to a 0.5 ml microcentrifuge tube. One microliter of mixture was placed in another 0.5 ml microcentrifuge tube and overlaid with a drop of mineral oil. The mixture was incubated in a thermal cycler at 98 °C for 1.5 min and simultaneously mixed with the two hybridization samples (hybridization samples 1 and 2 from the first hybridization, Table 2.8) by pipetting. The reaction was incubated at 68 °C overnight. After that 200 µl of dilution buffer (20mM HEPES, pH 6.6, 20 mM NaCl and 0.2 mM EDTA, pH 8.0) was added and mixed by pipetting. The reaction mixture was heated in a thermal cycler at 68 °C for 7 min and stored at -20 °C.

2.4.3.7 PCR Amplification

The PCR templates were prepared by aliquot 1 µl of each subtracted cDNA samples or unsubtracted testes control into an appropriately labeled tube. A master mix enough for all of the primary PCR tubes was prepared (Table 2.4).

Table 2.4 Preparation of the primary PCR master mix

Reagent	Amount per reaction (µl)
Sterile H_2O	19.5
10x PCR reaction buffer	2.5
dNTP Mix (10 mM)	0.5
PCR primer 1 (10 mM)	1.0
50x Advantage cDNA Polymerase Mix	0.5
Total volume	24.0

The reaction was mixed well by vortexing, and briefly centrifuged. An aliquot of 24 µl of the master mix was dispensed to each tube and overlaid with 50 µl of mineral oil. The reaction was incubated in a thermal cycler at 75 °C for 5 min to

extend the adaptors. The amplification reaction was carried out for 27 cycles composing of a 94 °C for 30 s, 66 °C for 30 s and a 72 °C for 1.5 min. After amplification, 8 µl from each tube were electrophoretically through a 2.0% agarose gel.

A ten-fold dilution was performed using 3 µl of each primary PCR mixture. One microliter of each diluted primary PCR product mixture was added into an appropriately labeled tube. A master mix for secondary PCR was prepared (Table 2.5).

Table 2.5 Preparation of the secondary PCR master mix

Reagent	Amount per reaction (µl)
Sterile H_2O	18.5
10x PCR reaction buffer	2.5
Nested PCR primer 1 (10 µM)	1.0
Nested PCR primer 2R (10 µM)	1.0
dNTP Mix (10 mM)	0.5
50x Advantage cDNA Polymerase Mix	0.5
Total volume	24.0

The reaction was mixed well and briefly centrifuged. An aliquot of 24 µl of the second master mix was dispensed to each of tube and overlaid with a drop of mineral oil. The reaction was carried out for 12 cycles composing of a 94 °C for 30 s, 68 °C for 30 s and a 72 °C for 1.5 min. After PCR, 8 µl from each tube were size-fractionated through a 2.0% agarose gel.

2.4.3.8 Ligation of PCR products to the pGEM®–T easy vector

The resulting products from the forward subtraction (cDNA from ovaries as the tester) and reverse reaction (cDNA from testes as the tester) were separately ligated to the pGEM – T Easy vector (Promega, USA) in a 10 µl reaction volume containing 5 µl of 2x Rapid Ligation Buffer (60 mM Tris-HCl, pH 7.8, 20 mM $MgCl_2$, 20 mM DTT, 2 mM ATP and 10% polyethylene glycol; MW 8000), 3 weiss units of T4 DNA ligase, 25 ng of $pGEM^R$ –T easy vector and 50 ng of DNA insert. The reaction mixture was incubated overnight at 4 °C before transformed to *E. coli* JM109.

2.4.3.9 Transformation of ligation products to *E. coli* host cells

The commercial *E. coli* JM109 competent cells (Stratagene) were thawed on ice for 5 min. Two to four microliters of the ligation mixture were added and gently mixed by pipetting and left on ice for 30 min. The transformation reaction was heat-shocked in a 42 °C water bath for exactly 45 s without shocking. The reaction tube was immediately placed in ice for 2-3 min. The mixture were removed from the tubes and added to a new tube containing 1 ml of prewarmed SOC (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose). The cell suspension was incubated with shaking at 37 °C for 90 min. The mixture were centrifuged for 20 s at room temperature, and gently resuspended in 100 µl of SOC medium and spread onto a selective LB agar plates containing 50 µg / ml of ampicillin, 25 µg / ml of IPTG and 20 µg / ml of X-gal and further incubated at 37 °C overnight (Sambrook et al., 2001). The recombinant clones containing inserted DNA are white whereas those without inserted DNA are blue.

2.4.3.10 Colony PCR of the recombinant clones

Recombinant clones were selected by a lacZ′ system following standard protocols (Sambrook and Russel, 2001). Colony PCR was performed to identify sizes of positive clones. Recombinant clones carrying insert sizes greater than 250 bp was extracted from broth SSH libraries (approximately 200 clones from each SSH libraries). Colony PCR was performed with condition and cycles following 2.4.2.11. The colony PCR products were electrophoresed though 1.2% agarose gel and visualized after ethidium bromide staining.

2.4.3.11 Extraction of recombinant plasmid DNA

A colony was inoculated into a sterile tube containing 3 ml of LB broth supplemented with 50 µg/ml of ampicillin and incubated with shaking (200 rpm) at 37 °C overnight. Plasmid DNA was isolated using HiYield™ Plasmid Mini Kit (RBC; Real Biotech Corporation) following 2.4.2.12. The concentration of extracted plasmid DNA was spectrophotometrically measured.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.4.3.12 Sequencing, EST clustering and assembly

Three hundred and sixty-seven clones (178 and 189 clones for subtractive cDNA libraries of testis from broodstock and juvenile, respectively) were randomly selected and unidirectional sequenced from the 5’ direction using a MegaBase 1000 automated DNA sequencer (GE Healthcare). Nucleotide sequences of ESTs were compared with those previously deposited in the GenBank using BLAST*N* and BLAST*X* (Altschul *et al*., 1990, available at http://www.ncbi.nlm.nih.gov). Significant matched nucleotides/proteins were considered when the E-value was $<10^{-4}$. Clustering and assembling of sequences were performed using TIGR Gene Indices Clustering Tools (TGICL) (Pertea *et al*., 2003) with CAP3 (Huang and Madan, 1999).

2.5 Isolation and characterization of the full length cDNA of functionally important gene homologues of *P. monodon* using Rapid Amplification of cDNA Ends-Polymerase Chain Reaction (RACE-PCR)

2.5.1 Preparation of the 5′ and 3′ RACE template

Full length cDNAs of interesting gene homologues were further characterized using a SMART RACE cDNA Amplification Kit (Clontech). Total RNA was extracted from testis of *P. monodon* using TRI-REAGENT. Messenger (m) RNA was purified using a QuickPrep *micro* mRNA Purification Kit (GE Healthcare) as described previously. RACE-Ready cDNA was prepared by combining 1 µg of ovarian mRNA with 1 µl of 5′-CDS primer and 1 µl of 10 µm SMART II A oligonucleotide for 5′- RACE-PCR and 1 µg of ovarian mRNA with 1 µl of 3′-CDS primer A for 5′- RACE-PCR. The components were mixed and spun briefly. The reaction was incubated at 70°C for 2 min and snap-cooled on ice for 2 min. The reaction tube was spun briefly. After that, 2 µl of 5x First–Strand buffer, 1 µl of 20 mM DTT, 1 µl of dNTP Mix (10 mM) and 1 µl of PowerScript Reverse Transcriptase were added. The reactions were mixed by gently pipetting and centrifuged briefly to collect the contents at the bottom. The tubes were incubated at 42°C for 1.5 hr in a thermocycler. The first strand reaction products were diluted with 250 µl of Tricine-EDTA (or TE) buffer and heated at 72°C for 7 min. The first strand cDNA template can be stored at -20 °C for up to three months.

2.5.2 Primer design

Gene-specific primers (GSPs) were designed from interesting transcripts obtained from testis and heart cDNA libraries and the subtractive testis cDNA libraries of *P. monodon*. Antisense and/or sense primers were designed for 5′RACE-PCR and 3′RACE-PCR, respectively (Table 2.6-2.7). Internal forward and/or reverse primers were also designed for further sequencing of the internal regions of large RACE-PCR fragments (Table 2.8).

2.5.3 RACE-PCR

The master mix for 5′RACE-PCR and 3′RACE-PCR was prepared. For each 25 µl amplification reaction, 16.0 µl of PCR-grade water, 2.5 µl of 10x Advantage 2 PCR buffer, 0.5 µl of 10 µM dNTP mix and 1 µl of 50x Advantage 2 polymerase mix were combined. 5′RACE-PCR and 3′RACE-PCR were set up according to Table 2.9 and 2.10 respectively.

Table 2.6 Primer sequences for the first strand cDNA synthesis and RACE - PCR

Primer	Sequence
SMART II A Oligonucleotide	5′-AAGCAGTGGTATCAACGCAGAGTACGC GGG-3′
3′-RACE CDS Primer A	5′-AAGCAGTGGTATCAACGCAGAGTAC$(T)_{30}$ N_{-1} N-3′
	(N = A, C, G or T; N_{-1} = A, G or C)
5′-RACE CDS Primer	5′-$(T)_{25}$ N_{-1} N-3′
	(N = A, C, G or T; N_{-1} = A, G or C)
10X Universal Primer A Mix (UPM)	Long : 5′-CTAATACGACTCACTATAGGGCAA GCAGTGGTATCAACGCAGAGT-3′
	Short : 5′-CTAATACGACTCACTATAGGGC-3′
Nested Universal Primer A (NUP)	5′– AAGCAGTGGTATCAACGCAGAGT-3′

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

Table 2.7 Gene-specific primers (GSPs) used for further characterization of the full length cDNA of functionally important gene homologues in *P. monodon*

Primer	Sequence	Tm (°C)
Cell division protein kinase 7 (CDK7)		
3′-RACE	5′-GGCAGAAAAACCGTCTCTCAAGCGAAAG-3′	84
Flotillin 2		
3′-RACE	5′-GAGGAGAGCGTGATGGTTGGTGGAGGA-3′	86
Growth factor receptor-bound protein		
3′-RACE	5′-GGGGCTCTTTCCTGCCACCTACG-3′	76
Innexin 1		
5′-RACE	5′-CAGGGAGTGTGAAGAGGGCAACTCGG-3′	84
Low molecular weight neurofilament protein XNF-L		
5′-RACE	5′-GTATTCAAGGACGACCTCCCACGGAC-3′	82
3′-RACE	5′-AGAAGGGGAAAGCGAGGCGGTACAAGA-3′	84
Multiple inositol polyphosphate phosphatase 2 (TT 0004)		
5′-RACE	5′-TGAACAGCCCCAGTAAGGTAATGAACG-3′	80
Multiple inositol polyphosphate phosphatase 2 (TT 0678)		
5′-RACE	5′-CATGGTGCGGAGTTTGGTTAAGTTGAC-3′	80
3′-RACE	5′-TTGGCGGAAGCGAAAAATGGGCTGAGCA-3′	86
Prohibitin 2 (a repressor of estrogen receptor activity)		
3′-RACE	5′-TACAGCCATCCCTAACATCCACCAGA-3′	78
Rac GTPase activating protein 1		
5′-RACE	5′-CGCAATACTGGTGTGACAAGATTGTGTG-3′	82
Transformer-2		
5′-RACE	5′-CAGGCACTTAGATGGCTCGGGATTGTC-3′	84
Innexin2		
5′-RACE	5′-GCTACGATGATGAACCAGAACCACAGGA-3′	84
3′-RACE	5′-TCGGGTAGCCTGGAGAAGCACGACGGAC-3′	92
Meiotic recombination protein DMC1/LIM15 homolog isoform 1		
5′-RACE	5′-GCCGCCAATCGGTTTCTTCGGGTCC-3′	82
3′-RACE	5′-CCTTCGCTATCACAGCAGGAGGCATTG-3′	84
Progestin receptor membrane component 1		
5′-RACE	5′-TGTCGTTTCATCTTGGGCACAGGAGGTT-3′	84
3′-RACE	5′-GCAAAGGACACCAAAGCGAAGACGGATG-3′	86
Dihydrolipoamide dehydrogenase		
5′-RACE	5′-AGGGATGCCTGGGAAGGGAGTAACC-3′	80
3′-RACE	5′-GGTCGCATTCCTGTCAACTCTCGCTTCC-3′	88
ERO-1		
5′-RACE	5′-TCAAGTATCCAAGGTCATTCTCTCCATC-3′	80
3′-RACE	5′-ATTGATGACGACAGTTCTGAGGAT-3′	68

Table 2.7 (cont.)

Primer	Sequence	Tm (°C)
TroponinT		
3′-RACE	5′-CGACGCACAGACAGGAGGACCTATGACG-3′	90
Saposin		
5′-RACE-1	5′-CTGGCACCAATAAGAAGGACCCCAAGTG-3′	86
5′-RACE-2	5′-CAGATGGGCAGAGATGCAGCATAGGACA-3′	86

Table 2.8 Internal primers used for primer walking sequencing of the full length cDNA of functionally important gene homologues in *P. monodon*

Primer	Sequence	Tm (°C)
BUB3 budding uninhibited by benzimidazoles 3 homolog (mitotic checkpoint)		
TT153seq-F1	5′-CTGAGGAACATGGGCTTTGC-3′	62
Mitoticseq-F2	5′-AGACCAAGTATTTGACCAGCAC-3′	64
TT153Seq-F3	5′-ATGATATGAACAGTGTAAGTGAC-3′	62
TT153Seq-R1	5′-GGTCAGAAAAACCTTTGAGGAAG-3′	66
Cell division cycle 2 (Cdc-2)		
cdc-2seq-F1	5′-CATTTACCTTGTGTTCCGAGTTCC-3′	70
cdc-2seq-F2	5′-GAGACCAAACATCAACAGGACAGG-3′	72
Innexin 1		
Innexin1seq-F	5′-TGACCCTCTTATCTACCGTGACATC-3′	74
Flotillin 2		
Flotillin2-F	5′-CGAGAATGTAGAAACCTCCCTGG-3′	70
Flotillin2-R	5′-TACTCTCTGCGTCTGCCTTACCC-3′	70
***Rac GTPase activating protein 1* (For 3′ end)**		
TT1036seq-F	5′-TCCTACAGCCTTGGTGGTCC-3′	64
***Transformer*-2 (For 3′ end)**		
Tra2seq-F1	5′-CACTCCAACTCCAGGAAT ATAC-3′	64
Tra2seq-F2	5′-TTTCGACCGTTATACAGTTACA-3′	64
Tra2seq-R	5′-ATGAGAAGGTGTTTGAGTGAC-3′	60
Translationally controlled tumor protein (TCTP)		
TT449seq-F	5′-CTGATGGCATGGTTGTTCTC-3′	60
Ubiquitin conjugating enzyme 2 CG6720-PA, isoform A isoform 1		
TT872seq-F	5′-ATTCAAACCACCTAAGGTCACG-3′	64
Molecular weight neurofilament protein XNF-L		
3′NEUseq-F	5′-GGCATTGGAGAATCATTGGG-3′	60
3′NEUseq-R	5′-CACCAATTTGAACTCAGCACAG-3′	64

Table 2.8 (cont.)

Primer	Sequence	Tm (°C)
Progestin receptor membrane component 1		
Prog800-F	5′-CACAGCACCTACCAAACTTATCC-3′	68
Prog800-R	5′-TGAAGAACAGTTAACAAGGCAGC-3′	66
ProgSeq-F	5′-GACAACAGGATCATAATGCTGC-3′	64
3Prog1-5Seq-F	5′-GTGCTGGTGGTGTAAGGTCC-3′	64
Saposin		
5′Prosa2Seq-R	5′-TGCCAGTCTGACGACACAAGC-3′	66
ERO-1		
ERO-1seqF	5′-ATTGATGACGACAGTTCTGAGGAT-3′	68

Table 2.9 Composition of 5′RACE–PCR

Component	5′-RACE Sample	UPM only (-Control)	GSP1 only (-Control)
5′-RACE-ReadycDNA	1.5μl	1.5μl	1.5μl
UPM (10x)	2.5μl	2.5μl	-
GSP1 (10μM)	1.0μl	-	1.0μl
GSP2 (10μM)	-	-	-
H_2O	-	1.0μl	2.5μl
Master Mix	20.0μl	20.0μl	20.0μl
Final volume	25μl	25μl	25μl

Table 2.10 Composition of 3′RACE–PCR

Component	3′-RACE Sample	UPM only (Control)	GSP2 only (Control)
3′-RACE-Ready cDNA	1.5μl	1.5μl	1.5μl
UPM (10x)	2.5μl	2.5μl	-
GSP1 (10μM)	-	-	-
GSP2 (10μM)	1.0μl	-	1.0μl
H_2O	-	1.0μl	2.5μl
Master Mix	20.0μl	20.0μl	20.0μl
Final volume	25μl	25μl	25μl

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

If Tm of GSP > 70 °C, the reaction was carried out for 5 cycles composing of a 94 °C for 30 s, and 72 °C for 2 min, 5 cycles composing of a 94 °C for 30 s, 70 °C for 1 min and 72 °C for 2 min, 20 cycles composing of a 94 °C for 30 s, 68 °C for 1 min and 72 °C for 2 min. The final extension was carried out at 72 °C for 7 min.

If Tm of GSP < 70 °C, the reaction was carried out for 25 cycles composing of a 94 °C for 30 s, 65-68 °C for 1 min and 72 °C for 2 min. The final extension was carried out at 72 °C for 7 min. Five microliters of RACE product was electrophoretically analyzed.

2.5.4 Agarose gel electrophoresis

An appropriate amount of agarose was weighed out and mixed with an appropriate volume of 1 x TBE buffer (89 mM Tris-HCl, 89 mM boric acid and 2 mM EDTA, pH 8.3). The gel slurry was boiled in a microwave oven to complete solubilization, and allowed to approximately 60 °C before poured into the gel mold. A comb was inserted. The gel was left to solidified. When needed, the comb was carefully removed. The agarose gel was submerged in a chamber containing an enough amount of 1 x TBE buffer covering the gel for approximately 0.5 cm.

Appropriate volumes of PCR products were mixed with the one-fourth volume of the 10x loading dye (0.25% bromophenol blue and 25% Ficoll in water) and loaded into the well. A 100 bp DNA ladder was used as the standard DNA marker. Electrophoresis was carried out at 5–6 volts/cm until bromophenol blue moved to approximately one-half of the gel**.** The electrophoresed gel was stained with an ethidium bromide solution (2.5 μg/ml) for 5 min and destained in running tap water to remove unbound ethidium bromide from the gel. DNA fragments were visualized under a UV transilluminator and photographed through a red filter using Fomapan Classic 100 film. The exposure time was 10 – 18 s.

2.5.5 Elution of DNA fragments from agarose gels

After electrophoresis, the desired DNA fragment was excised from the agarose gel (200-300 mg) using a sterile scalpel and placed in a preweighed microcentrifuge tube. DNA was isolated using HiYield^{TM} Gel/PCR DNA Extraction Kit (RBC; Real

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Biotech Corporation). 0.5 ml of DF buffer was added and mixed by vortexing. The mixture was incubated at 55 °C for 10-15 min or until the gel slice was completely dissolved. The mixture was transferred into a DF column inserted in a collection tube and centrifuged at 8000 rpm for 30 s. The flow-though solution was discarded. After this step, 0.5 ml of wash buffer was added to the DF column and centrifuged as above. The flow-though solution was discarded. The column was recentrifuged to remove the trace amount of the washing solution. The DF column was then placed into a sterile 1.5 ml microcentrifuge tube. DNA was eluted out by an addition of 15 μl of the Elution buffer or H_2O to the center of the DF membrane and left for 1 min, before centrifuged at 12000 rpm for 2 min.

2.5.6 Ligation of PCR products to the pGEM®–T easy vector

DNA fragments were ligated to the pGEM – T Easy vector (Promega, USA) in a 10 μl reaction volume as described previously. The reaction mixture was incubated overnight at 4 °C.

2.5.7 Transformation of ligation products to *E. coli* host cells

2.5.7.1 Preparation of competent cells

A single colony of *E. coli* JM109 was inoculated in 10 ml of LB broth (1% tryptone, 0.5% yeast extract and 0.5% NaCl, pH 7.0) with vigorous shaking at 37 °C overnight. The starting culture was then inoculated into 50 ml of LB broth and continued culture at 37 °C with vigorous shaking to an OD_{600} of 0.5 to 0.8. The cells were briefly chilled on ice for 10 min, and recovered by centrifugation at 2700*g* for 10 min at 4°C. The pellets were resuspended in 30 ml of ice-cold $MgCl_2/CaCl_2$ solution (80 mM $MgCl_2$ and 20 mM $CaCl_2$) and centrifuged as above. After resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$, the concentrated cell suspension was divided to 200 μl aliquots. These competent cells was either used immediately or stored at -80 °C for subsequently used.

2.5.7.2 Transformation

The competent cells were thawed on ice for 5 min. Two to four microliters of the ligation mixture were added and gently mixed by pipetting and left on ice for 30

min. The transformation reaction was heat-shocked in a 42°C water bath for exactly 45 s without shocking. The reaction tube was immediately placed in ice for 2-3 min. The mixture were removed from the tubes and added to a new tube containing 1 ml of prewarmed SOC (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose). The cell suspension was incubated with shaking at 37 °C for 90 min. The mixture were centrifuged for 20 s at room temperature, and gently resuspended in 100 µl of SOC medium and spread onto a selective LB agar plates containing 50 µg / ml of ampicillin, 25 µg / ml of IPTG and 20 µg/ml of X-gal and further incubated at 37°C overnight (Sambrook and Russell, 2001). The recombinant clones containing inserted DNA are white whereas those without inserted DNA are blue.

2.5.8 DNA sequencing and assembly sequence

The recombinant clones were unidirectional sequenced using the M13 forward and reverse primer on an automatic sequencer at Macrogen (Korea). Nucleotide sequence from RCAE product was assembled with EST sequence using clustalW. Nucleotide sequences were blasted against data in the GenBank using Blast*N* (nucleotide similarity) and Blast*X* (translated protein similarity). Significant similarity was considered when the probability (E) value was $<10^{-4}$.

2.5.9 Phylogenetic analysis of *CyclophilinA, SUMO, Tra-2, PGRMC1,* and *Dmc1*

Protein sequences of *Dmc1, PGMRC1, cyclophilin A*, *SUMO-1* and *Tra-2* from various species were retrieved from GenBank and compared with those of *P. monodon*. Multiple alignments were carried out using ClustalW (Thompson et al., 1994). Sequences were bootstrapped 1000 times using Aeqboot. The divergence between pairs of protein sequences was estimated using Prodist. A bootstrapped neighbor-joining tree (Saitou and Nei, 1987) was constructed to illustrate phylogenetic relationships among sequences using Neighbor and Consense. All phylogenetic programs described were routine in PHYLIP (Felsenstein, 1993).

2.6 Examination of expression patterns of genes related with testicular development by RT-PCR and tissue distribution analysis

2.6.1 Primer design

Forward and reverse primers of each gene were designed from nucleotide sequence obtained from the conventional cDNA libraries of heart and testis and the subtractive cDNA libraries of *P. monodon* testis using Primer Premier 5 program (Table 2.11).

2.6.2 First stranded cDNA synthesis

One and a half micrograms of total RNA from various tissues of *P. monodon* were reverse-transcribed to the first strand cDNA using an ImProm- II^{TM} Reverse Transcription System Kit (Promega). Total RNA was combined with 0.5 µg of oligo dT_{12-18} and appropriate DEPC-treated H_2O in final volume of 5 µl. The reaction was incubated at 70 °C for 5 minutes and immediately placed on ice for 5 minutes. Then 5x reaction buffer, $MgCl_2$, dNTP Mix, RNasin were added to final concentrations of 1x, 2.25 mM, 0.5 mM and 20 units, respectively. Finally, 1 µl of ImProm- II^{TM} Reverse transcriptase was add and gently mixed by pipetting. The reaction mixture was incubated at 25 °C for 5 minutes and at 42°C for 90 minutes. The reaction mixture was incubated at 70 °C for 15 minutes to terminate the reverse transcriptase activity. Concentration and rough quality of the newly synthesized first strand cDNA was spectrophotometrically examined (OD_{260}/OD_{280}) and electrophoretically analyzed by 1.0% agarose gels.

2.6.3 RT-PCR analysis

Basically, amplification reactions were performed in a 25 µl reaction volume containing 10 mM Tris-HCl, pH 8.8 at 25 °C, 50 mM KCl and 0.1% Triton X-100, 1.5-2.0 mM $MgCl_2$, 100-200 µM each of dATP, dCTP, dGTP and dTTP, 0.2-0.4 µM of each primer, 1 unit of Dynazyme™ DNA Polymerase (FINNZYMES, Finland) and 2 µl of a 10 fold-diluted first strand cDNA (about 200 ng). RT-PCR was initially performed by predenaturation at 94 °C for 3 min followed by 25 cycles of a 94 °C denaturation for 30 s, a 55 °C annealing for 45 s and a 72 °C extension for 45 s. The final extension was carried out at 72 °C for 7 min. Eight microliters of the amplification products were electrophoretically analyzed though 1.0-2.0% agarose gel.

Table 2.11 Sequence, melting temperature and expected product sizes from primers designed from the conventional testis cDNA library and SSH cDNA libraries of *P. monodon*

Gene/Primer	Sequence	Tm (°C)	Size (bp)
1. ***Actin-binding protein anillin***			
F:	5′-TGTTTGAGGATGTTGGGGGCT-3′	64	199
R:	5′-AACTGGAAGGTATGCTGACGGG-3′	68	
2. ***Acyl-CoA oxidase (ACX3)***			
F:	5′-CACTGCGTATCTTTGCTGCTC-3′	64	250
R:	5′-TCATCCGTTTCCTTGTCGTAA-3′	60	
3. ***Adapter-related protein complex 1 beta 1 subunit***			
F:	5′-TTCCCTGATGTGGTCAACTGTATG-3′	70	290
R:	5′-TGGTCTACGCAGATAGCAGCAG-3′	68	
4. ***Arginyl-tRNA-protein transferase 1 (Arginyltransferase 1)***			
F:	5′-GGGATGGAGACGGAGTGGAA-3′	64	292
R:	5′-TGGCATCTGGAGGGATACACC-3′	66	
5. ***B-cell receptor-associated protein 37 (Prohibitin2)***			
F:	5′-CGTATGGCATCTCGCAGTCC-3′	64	563
R:	5′-CCTCTCCTGTCTCGCTCTCTCG-3′	72	
6. ***BCS-2***			
F:	5′-TGAAGTGTAAAGTGCTGTGGGG-3′	66	372
R:	5′-TGGGCGGTGAACTCCGTGGT-3′	66	
7. ***Budding uninhibited by benzimidazoles 3 homolog (mitotic checkpoint)***			
F:	5′-CGAGTCTGAAGTCGGCAAAATG-3′	66	257
R:	5′-GCTGACCATCTAAGCCTCCACT-3′	68	
8. ***Carbonyl reductase 1***			
F:	5′-GCTTCGCTCCTCGTTTTCATCA-3′	66	457
R:	5′-ATACGCCAACTTTGCTCTACCAC-3′	68	
9. ***Cell division cycle 2 (Cdc2)***			
F:	5′-AAGAACCGCAAAAGTGGGAAG-3′	66	510
R:	5′-GCCAAGAGACCAAACATCAACAG-3′	68	

Table 2.11 (cont.)

Gene/Primer	Sequence	Tm (°C)	Size (bp)
10. ***Cell division protein kinase 7 (CDK7)***			
F:	5′-CGGAAGACAGGATGGAAGTAGAA-3′	68	381
R:	5′-ATGTTGGATGGGCGTGAGGATG-3′	68	
11. ***COP9 constitutive photomorphogenic homolog subunit 5 isoform 1***			
F:	5′-CGGTCTGGAGGCACACTTGAG-3′	68	190
R:	5′-CATTTTCCTGGCGACCAACCT-3′	62	
12. ***Cyclin dependent kinase 2 (CDK2)***			
F:	5′-CGACTGCTGGATGTGGCGTA-3′	64	348
R:	5′-CCGAGGAGAATCTGTGGGGC-3′	66	
13. ***Cyclophilin A***			
F:	5′-ATGGGCAACCCCAAAGTCTTTTTCGA-3′	76	495
R:	5′-TTACAGCTGGCCGCAGTTGGCG-3′	72	
14. ***Cystathionine gamma-lyase***			
F:	5′-CCCAGCAGATTTTAAGGCATTTGA-3′	68	145
R:	5′-GTGCGTGATGGTGGTTGTCG-3′	64	
15. ***Degradation in endoplasmic reticulum protein 1 (DER1)***			
F:	5′-CACGACGACCAGCAGACTACC-3′	68	206
R:	5′-GAGAAGCACCCACGGGAGAT-3′	64	
16. ***Dynectin 5***			
F:	5′-AAGTTATCGTGCTTAGTGGTGTG-3′	66	300
R:	5′-TCTCGGTGAGGCAAGACTGTTT-3′	66	
17. ***Dynein light intermediate chain***			
F:	5′-GCAAGTCTGTTCTCGTCCTGG-3′	66	324
R:	5′-TGTCTATGTGGTCTTGGAGAGTGG-3′	72	
18. ***E1B-55kDa-associated protein 5 isoform 5***			
F:	5′-CCTTGGTGCTCTCCATTGACTG-3′	64	196
R:	5′-CTGTTGATGAGGCTGGGCTG-3′	64	
19. ***Eukaryotic translation initiation factor 4 gamma, 2***			
F:	5′-CAACCTCCTAACCAGAAGCCAACA-3′	60	337
R:	5′-CAACATTCATAAGTCCTTCACTGCGA-3′	74	

Table 2.11 (cont.)

Gene/Primer		Sequence	Tm (°C)	Size (bp)
20. ***Flotillin2***				
	F:	5′-CGAGAATGTAGAAACCTCCCTGG-3′	70	854
	R:	5′-TACTCTCTGCGTCTGCCTTACCC-3′	72	
21. ***Growth factor receptor-bound protein***				
	F:	5′-GCTGGACGGCAGAGAAGGAC-3′	66	299
	R:	5′-ACACAGACGCTGACCGATGG-3′	64	
22. ***Heat shock-related 70 kDa protein 2 (Testis-specific heat shock protein-related)***				
	F:	5′-CCGCACAACGCCTTCCTACG-3′	66	321
	R:	5′-GCATCTTTGACGACTCCACCCAG-3′	62	
23. ***Histone H2AV (H2A.F/Z) (Old BlastX: SPARC)***				
	F:	5′-AGCGTATCACCCCTCGCCATCT-3′	70	329
	R:	5′-TCTCAAACACCTCCCTACTCCATCA-3′	74	
24. ***Importin 7***				
	F:	5′-AGTGCCTCCCATCCATTATCC-3′	64	297
	R:	5′-GGACTCCCAAGCGTTCAGG-3′	62	
25. ***Inhibitor of apoptosis protein***				
	F:	5′-CCTGAAGAGTTAGCAGCAGATGG-3′	70	238
	R:	5′-TACTTGCTTTTGGAGGATTGTCAC-3′	68	
26. ***Innexin1***				
	F:	5′-GGAGGAGGTTGAAGAAAGGTGG-3′	68	422
	R:	5′-TGGTCATTCGTGGGAAGACATAG-3′	66	
27. ***Karyopherin alpha 2 (Importin-α1)***				
	F:	5′-GATGACCCAACTTCACCCCTT-3′	64	295
	R:	5′-CTGCCTGTGTCTGCTCTGATGT-3′	68	
28. ***Laminin beta chain***				
	F:	5′-ATGCGATGAGTGTGCCCGAG-3′	64	363
	R:	5′-GCCAAAGAAATGCGTTGTGTAGTG-3′	70	
29. ***Low molecular weight neurofilament protein (PMTT1)***				
	F:	5′-GGGAGGAAGAACACCCCAATG-3′	66	374
	R:	5′-TACACGCTGAGCAACGAGAACG-3′	68	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 2.11 (cont.)

Gene/Primer	Sequence	Tm (°C)	Size (bp)
30. ***Metaxin2***			
F:	5′-AGATACTGCTCCCTGATAATGCCCA-3′	74	212
R:	5′-GCCGTCTGTCAAGGTCACTCCC-3′	72	
31. ***Multiple inosital polyphosphate (MIPP)*** *(TT 0004)*			
F:	5′-TATTCCAAGGACAACCCAGGCT-3′	66	174
R:	5′-TTTACATCCCCTCGTCCCGCTT-3′	68	
32. ***Multiple inosital polyphosphate 2 (MIPP2)*** *(TT 0678)*			
F:	5′-CCTTCGGATACACCACGACCAC-3′	70	396
R:	5′-AACCTCTCCCAGGGGCAACC-3′	66	
33. ***Multiprotein bridging factor 1***			
F:	5′-TGCCACCACCTTCAACACAG-3′	62	235
R:	5′-CATCCCAATAGCCTTCTCAATC-3′	64	
34. ***Nudix-type motif 9 isoform a (NUDT9)***			
F:	5′-CGCACTGATGATAAGACTCCTCG-3′	70	341
R:	5′-CCAGCATCCATTTTCCACCG-3′	62	
35. ***Oncoprotein nm23***			
F:	5′-CTGACAAGCCCTTCTACCCTGG-3′	70	229
R:	5′-AGAGCAATCTCCTTGTTGGCAG-3′	66	
36. ***PCTAIRE protein kinase 2***			
F:	5′-CGAGACCTCAAGCCTCAGAACC-3′	70	250
R:	5′-CTACTGTGGCACCTGGGAAGAG-3′	70	
37. ***Polyadenylate binding protein II***			
F:	5′-CCCTCTAGCCTCGCTCTATGTG-3′	70	187
R:	5′-GTCTAGGGCTCGTTCAGCATCA-3′	68	
38. ***Profilin***			
F:	5′-CTGGCTATGTTTCTAAGGCGGT-3′	66	259
R:	5′-ATGGCTATCAGGACTGCTTGC-3′	64	
39. ***Programmed cell death protein 7***			
F:	5′-CCCTGACAGCCCTGCGACA-3′	64	181
R:	5′-GCACTTTCACCCATCATAACCCG-3′	70	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 2.11 (cont.)

Gene/Primer	Sequence	Tm (°C)	Size (bp)
40. ***Proteasome (prosome, macropain) 26S subunit, ATPase2***			
F:	5′-AACTCTCCAGAATGAGCAGCCA-3′	66	187
R:	5′-TACTTATTACGATGCCACACCCAC-3′	70	
41. ***Proteasome alpha 3 subunit***			
F:	5′-AAAGATGGTGTTGTGTTTGCTGTAG-3′	70	250
R:	5′-AGAGGGTATAGGCATGAAGGTAGG-3′	60	
42. ***26S Proteasome non-ATPase regulatory subunit 3 (Diphenol oxidase A2 component)***			
F:	5′-CGCCTGGTTGAACGCAGCATTG-3′	70	140
R:	5′-TTCTCTGGTCGCCCATAGTAAGT-3′	68	
43. ***Rac GTPase activating protein 1 isoform 1***			
F:	5′-GAAACAGGGCACAGGTCAGATG-3′	68	248
R:	5′-CACAAGGAAGCGGCACAGAT-3′	62	
44. ***Serine palmitoyl transferase LCB2 subunit***			
F:	5′-ATCCAACAACTGTCTCGCAATG-3′	68	178
R:	5′-ATCCAACCCCTACGCCAGCCAC-3′	72	
45. ***Serine/threonine-protein kinase 23 (Muscle-specific serine kinase 1) (MSSK-1)***			
F:	5′-CACTGTTTGGCTGTGTTGGG-3′	62	272
R:	5′-CCTGGTAATTGGAGCGGATG-3′	62	
46. ***Small ubiquitin-like modifier (SUMO-1)***			
F:	5′-GGAAGGGAACGAATACATCAAA-3′	62	362
R:	5′-GCCTGGTCTGTGCGAAAATCTC-3′	68	
47. ***Spermidine synthase***			
F:	5′-GAGGTAGCCAAACACCCAGCA-3′	66	223
R:	5′-GGTCATCATCAGCAACAACAGGA-3′	68	
48. ***Synaptobrevin like protein 1B***			
F:	5′-CTCCTCTACAGTGTCATTGCTCGG-3′	74	291
R:	5′-AAGTGCGGTGTGAACTCGGCT-3′	66	
49. ***Bromodomain containing 8 (Thyroid hormone receptor coactivating protein 120 kDa, (TrCP120)***			
F:	5′-CGCAGTGGCGACCAGAAC-3′	60	392
R:	5′-CACTGTTGGTGCTTCCTTTCCT-3′	66	

Table 2.11 (cont.)

Gene/Primer	Sequence	Tm (°C)	Size (bp)
50. ***Thyroid hormone receptor-associated protein complex 240 kDa component (Trap240)***			
F:	5′-TAGGTAGGCTTGGTAGAATGGGC-3′	70	335
R:	5′-GGAATCTCTGCTGTGCTGACTGA-3′	70	
51. ***Transformer-2***			
F:	5′-CACCAGAGACAATCCCGAGC-3′	64	229
R:	5′-GTCAATCTCCATCCCAGAACACT-3′	68	
52. ***Ubiquitin specific protease 14***			
F:	5′-ATACTGCCGAACCCCCAATG-3′	62	240
R:	5′-CAGTGTTCCTCTCAGCTTCAGTCA-3′	60	
53. ***Ubiquitin carboxyl-terminal hydrolase 5 (old name: Ubiquitin isopeptidase T)***			
F:	5′-CAAGTTGGCTGCCCCTGAAG-3′	64	528
R:	5′-GTTGCCTGCTCTCGTGTGAATC-3′	68	
54. ***Ubiquitin-conjugating enzyme E2***			
F:	5′-TCTGCCTCGCCTGCTGGT-3′	60	232
R:	5′-TGGTGCTGAGTGCCTTTGACAT-3′	66	
55. ***USO-1***			
F:	5′-GCTGACCTATTCCTGCGTCTTTG-3′	70	314
R:	5′-TCGTGTTTCTTGGCGACCCTTTG-3′	70	
56. ***Meiotic recombination protein DMC1/LIM15 homolog isoform 1 (Dmc1)***			
F:	5′-ATGGAAGATCAGGCTTTAGATGC-3′	66	425
R:	5′-GTGACGCAGAGAGTGTGGGAG-3′	68	
57. ***Innexin2***			
F:	5′-AAGATGTGGGAAGGAGGCAAGA-3′	66	391
R:	5′-TGAGCGGGAGAACGCAGAGT-3′	64	
58. ***Progestin receptor membrane component 1 (PGRMC1)***			
F:	5′-GCCCAAGATGAAACGACAGG-3′	62	122
R:	5′-TGGAGCCTCGGGTGACATC-3′	62	
59. ***Prosaposin isoform 3***			
F:	5′-GCTATGGTTCAGGTTGATGACTTGC-3′	74	614
R:	5′-ACTCCCTTCCACACCTTCGTTTC-3′	70	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table2.12 Sequence, melting temperature and expected product sizes from primers designed from the heart cDNA library of *P. monodon*

Gene/Primer	Sequence	Tm (°C)	Size (bp)
1. *ERO-1*			
F:	5′-AAGACCCCAAACGGAGACAC-3′	62	501
R:	5′-CTCAGATTCAGCAATAGATCAACAT-3′	68	
2. *Fasciclin-like protein*			
F:	5′-ACTGTCTTCGCCCCAAGCAAC-3′	66	375
R:	5′-GCAATCTGTTCGTCCAAATCCA-3′	64	
3. *GAPDH*			
F:	5′-CGCCGCCCAGAACATTATCC-3′	64	230
R:	5′-CCTCGGTTGTATCCCAGCACG-3′	68	
4. *High mobility group 20A (HMG20A)*			
F:	5′-GAATGAGACCAGCCAGCAACG-3′	66	
R:	5′-AGACTCTGGTATTGTTGTGGGTTG-3′	70	
5. *Dihydrolipoamide dehydrogenase*			
F:	5′-GCTGGTGGCATTGCTCATCTC-3′	62	456
R:	5′-CCTTGCTTTGCTGGCACTCAT-3′	64	
6. *MLC1 protein*			
F:	5′-GTGGCAAGGAGAAGAGGAAAGA-3′	66	195
R:	5′-GACGCTCACCAAGGGACAGAAG-3′	70	
7. *TroponinT*			
F:	5′-GCACCATGTCTGACGACGAAT-3′	64	422
R:	5′-GAACTTGTCACCACCGCTCTT-3′	64	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

2.6.4 Tissue distribution analysis

Total RNA was extracted from eyestalks, gills, heart, hemocytes, hepatopancreas, lymphoid organs, intestine, pleopods, stomach, testes and thoracic ganglion of a male and ovaries of a female broodstock of *P. monodon* and subjected to the first strand cDNA synthesis.

Two microliters (approximately 200 ng) of the first strand cDNA was used as the template in a 25 µl reaction volume containing 10 mM Tris-HCl, pH 8.8 at 25°C, 50 mM KCl and 0.1% Triton X–100, 2 mM $MgCl_2$, 100 µM each of dATP, dCTP, dGTP and dTTP, 0.2 µM of each primer and 1 unit of Dynazyme™ DNA Polymerase (FINNZYMES, Finland). The reaction was predenaturation at 94 °C for 3 min followed by 25 cycles composing of a 94 °C denaturation step for 30 s, a 55 °C annealing step for 45 s and a 72 °C extension step for 45 s. The final extension was carried out at 72 °C for 7 min. Eight microliters of the amplification product was electrophoretically analyzed though 1.8% agarose gels.

2.7 Semiquantitative RT-PCR of functionally important gene homologues in testes and ovaries of broodstock and juvenile *P. monodon*

2.7.1 Experiment animals

Six sample groups composed of testes from different periods (1-5, 6-9 and 10-16 days after molting) of broodstock-sized shrimp after molting, wild broodstock-sized shrimp, domesticated broodstock-sized shrimp, and juvenile males and ovaries from wild broodstock-sized shrimp and juvenile females were used for semiquantitative RT-PCR analysis.

2.7.2 Primers

Expression levels of 6 specific primers, *low molecular weight neurofilament protein XNF-L (PMTT1)*, *transformer-2* (*Tra-2*), *prohibitin2*, *thyroid hormone receptor-associated protein complex 240 kDa component (Trap240), multiple inositol polyphosphate phosphatase 2* (*MIPP2*) and *heat shock-related protein 2* (*HSP70-2*), were determined. *EF-1α* gene was used as an internal control. Primers used for RT-PCR were also used for semiquantitative RT-PCR except the forward primer for *Tra-2* was replaced by a new forward primer (5′-CAGTCTCACACCTCGTTCGC-3′) and the expected sized was 499 bp.

2.7.3 Total RNA extraction and the first strand cDNA synthesis

Total RNA was extracted from testes from molted shrimp broodstock (N = 18), wild broodstock-sized shrimp (N = 8), domesticated broodstock-sized shrimp (N = 9),

and juvenile males (N = 5) and ovaries from wild broodstock-sized shrimp (N = 4) and juvenile female *P. monodon* (N = 4) using TRI REAGENT. The first strand cDNA synthesis was carried out as described previously.

2.7.4 Optimization of semiquantitative RT-PCR conditions

Amplification was performed in a 25 μl reaction volume containing 2 μl of the first strand cDNA template diluted ten-fold, 1 x PCR buffer (10 mM Tris-HCl pH 8.8, , 50 mM KCl and 0.1% Triton X – 100) , 200 μM each of dATP, dCTP, dGTP and dTTP and 1 unit of Dynazyme™ DNA Polymerase (FINNZYMES, Finland). PCR was carried out using the conditions described in the standard RT-PCR.

2.7.4.1 Primer concentration

The optimal primer concentration for each primer pair (between 0.1-0.4 μM) was examined using the standard PCR conditions. The resulting product was electrophoretically analyzed. The primer concentration that gave product specificity and clear results were selected for further optimization of PCR conditions.

2.7.4.2 $MgCl_2$ concentration

The optimal $MgCl_2$ concentration of each primer pair (between 1-4 mM $MgCl_2$) was examined using the standard PCR conditions and the optimized primer concentration. The concentration that gave the highest specificity was chosen.

2.7.4.3 Cycle numbers

The PCR amplifications were carried out at different cycles (e.g. 20, 25, 30 and 35 cycles) using the optimized concentration of primers $MgCl_2$ and analyzed by gel electrophoresis. The number of cycles that still provided the PCR product in the exponential rage and did not reach a plateau level of amplification was chosen.

2.7.4.4 Gel electrophoresis, quantitative, and data analysis

Eight microliters of the PCR products were combined with 2 μl of the loading dye before loaded to 1.8% agarose gel and electrophoresed at 5-6 volts/cm. The gel was stained with 2.5 μg / ml EtBr for 5 min and destained in the running tap water for

15 min. The intensity of target and control bands was quantified from photographs of the gels using the Gel Pro program.

The expression level of each transcript was normalized by that of *EF-1α*. Significantly different expression levels between different groups of *P. monodon* were tested using one way analysis of variance (ANOVA) followed by Duncan multiple range test ($P < 0.05$).

2.8 Examination of expression levels of interesting genes in testes of *P. monodon* by real-time PCR

Expression levels of several transcripts related with testicular development, for examples; *SUMO-1, cyclophilin A (CYA), Dmc1*, *progestin receptor membrane component 1 (PGRMC1)*, *saposin*, and *spermatogonial stemcell renewal factor* were examined using quantitative real-time PCR analysis.

2.8.1 Effects of dopamine on expression of genes in testes of juvenile *P. monodon*

Female juveniles of *P. monodon* (with the body weight about 20-25 g) were purchased from a local farm in Chonburi, eastern Thailand. Shrimp were acclimatized for 7 days at the laboratory conditions (25 ppt of seawater and the ambient temperature) and fed daily at approximately 5% of the body weight. Eight shrimp were placed in a glass aquarium (150 L) for each treatment and were not fed at 24 hr before the treatment.

Shrimp were intramuscularly injected individually with normal saline (control, N = 5) and dopamine hydrochloride to obtain the doses of 10^{-6}, 10^{-7} and 10^{-8} mol/shrimp. Testes of the negative control (no injection), normal saline-injected and treated shrimp at 0, 3, 6, 12 and 24 hr post treatment (N = 5 for each treatment) were sampling and immediately placed in liquid N_2. The samples were stored at -80 °C prior to RNA extraction and first-stand cDNA synthesis.

2.8.2 Expression of genes in testes in different groups of *P. monodon*

Shrimp at the intermolt stage were used in the experiments. However, stages of molted male broodstock were accounted as days after molting (DAM). Of these, juvenile males and females (4 months old with a body weight of approximately 25-30

g) were purchased from a commercial farm in Chachoengsao (eastern Thailand). Broodstock-sized domesticated shrimp (F1 generation, an average body weight, ABW = 88.66 ± 9.02 g; the gonadosomatic index, GSI = 0.34 ± 0.15%, N = 9) were obtained from the broodstock management center located at Burapha University (Chanthaburi, eastern Thailand). Wild male (ABW = 81.69 ± 15.63 g, GSI = 0.84 ± 0.32%, N = 8) and wild female (ABW = 142.98 ± 28.37 g, GSI = 2.98 ± 2.02%, N = 4) broodstock were also collected. Shrimp were acclimated under the laboratory conditions (28-30°C ambient temperature and 30 ppt salinity) for 3-5 days before their gonads were dissected out. The other group of wild shrimp were maintained in the farm and testes of each shrimp were collected at 1-5 (ABW = 104.79 ± 12.40 g, GSI = 0.40 ± 0.13%, N = 6), 6-9 (ABW = 99.19 ± 14.94 g, GSI = 0.41 ± 0.29%, N = 4) and 10-16 days (ABW = 102.73 ± 19.23 g, GSI = 0.48 ± 0.16%, N = 8) after molting.

2.8.3 Primers and construction of the standard curve

Primers for *Dmc1*, *spermatogonial stem-cell renewal factor*, *saposin* and *cyclophilin A* (Table 2.11) previously designed and ORF primers of *PGRMLC1* and *SUMO-1* (Table 2.13) were applied for real-time PCR analysis. For construction of the tandard curve of each gene, the PCR product was amplified, electrophoresed through agarose gel and eluted out. The gel eluted product was cloned into pGEM-TEasy vector and transformed into *E. coli* JM109. Plasmid DNA were extracted and used as the template for construction of the standard curve. Templates of each gene homologues were ten fold diluted to cover $10^3 - 10^9$ copy numbers. Real-time RT-PCR assay was carried out in a 96 well plate and each standard point was run in triplicate.

2.8.4 Quantitative real-time RT-PCR

The first strand cDNA was reverse-transcribed. The target transcript (*PGMRC1* and *Dmc1*) and the internal control (*EF-1α*) of each shrimp treated with 10^{-6} mol/shrimp of dopamine at the different time-point were amplified in a reaction volume of 25 µl containing 12.5 µl of 2x SYBR Green Master Mix (QIAGEN). The specific primer pairs were used at a final concentration of 0.4 µM. The thermal profile for SYBR Green RT-PCR was 95 °C for 15 min followed by 40 cycles of 95°C for 15

s, 55°C for 30 s and at 72°C for 45s. Continually, cycles for the melting curve analysis was carried out at 95°C for 15 s, 60°C for 1 min and at 95°C for 15 s.

In addition, *SUMO-1, cyclophilin A, spermatogonial stem-cell renewal factor, Dmc1* and *saposin* and the internal control (*EF-1α*) of shrimp possessing different stages of testes and ovaries were amplified in a reaction volume of 15 µl containing 7.5 µl of 2x LightCycler 480 SYBR Green I Master (Roche). The specific primer pairs were used at a final concentration of 0.4 µM (except *Saposin* where 0.2 µM primers were used). The thermal profile for SYBR Green RT-PCR was 95 °C for 10 min followed by 40 cycles of 95°C for 15 s, 58°C for 30 s and at 72°C for 30 s. Continually, cycles for melting curve analysis was 95°C for 15 s, 65°C for 1 min and at 97°C for continue and cooling was 40°C for 10 s. Real-time RT-PCR assay was carried out in a 96 well plate and each sample was run in duplicate. Relative expression levels of different groups of samples were statistically tested by ANOVA followed by Duncan's new multiple range test or Tukey test ($P < 0.05$).

2.9 Large scale examination of gene expression of *P. monodon* by microarray analysis

2.9.1 Total RNA extraction

Total RNA was extracted from testes of normal juvenile and broodstock of *P. monodon* and those of juvenile shrimp treated with dopamine. For treated *P. monodon*, juveniles treated with 10^{-6} mol/shrimp of dopamine solution at 3 and 6 hr post injection were further used.

2.9.2 Synthesis of target cDNA

Twenty micrograms of total RNA were fluorescently labeled with Aminoallyl-dUTP using a LabelStar Array Kit, cDNA Labeling Module (QIAGEN). After incubation at 65 °C for 5 min, total RNA was immediately placed on ice and briefly centrifuge immediately and centrifuge briefly. The denatured template was added into 30 µl of the master mix containing 1X RT-Buffer RT, dNTP mix (0.5 mM for dATP, dCTP, dGTP and Aminoallyl-dUTP), 2 µM Oligo(dT) primer, 20 units of RNase Inhibitor, and LabelStar Reverse Transcriptase and incubate at 37 °C for 2 hr

Table 2.13 Nucleotide sequences of primers used for real-time PCR analysis of *Dmc1, PGRMLC1, SUMO-1*, *cyclophilinA*, *spermatogonial stem-cell renewal factor* and *saposin* in *P. monodon*

Gene/Primer	Sequence	Tm (°C)	Size (bp)
Dmc1			
Real-time	F: 5′-ATGTGCGAGAAGCGAAGGC-3′	60	150
	R: 5′-GCAGAGAGTGTGGGAGATTTGTG-3′	70	
PGRMLC1			
Standard	F: 5′-ATGGCGGACGAGGGAGCG-3′	62	573
	R: 5′-CTAATCATCCGTCTTCGCTTTGGT-3′	70	
Real-time	F: 5′-GCCCAAGATGAAACGACAGG-3′	62	122
	R: 5′-TGGAGCCTCGGGTGACATC-3′	62	
SUMO			
Standard	F: 5′-ATGTCTGATAACACTGACGCCAAGC-3′	74	282
	R: 5′-TCAATGGCCGCCGGTCTG-3′	60	
Real-time	F: 5′-AGAAGGGGAAGGGAACGAATACA-3′	68	148
	R: 5′-ACGCAGCGATGCTACAGGGA-3′	64	
CyclophilinA			
	F: 5′-GGGCGGCAAGTCCATCTACG-3′	66	160
	R: 5′-GTGCTTGTTGTCCAGCCAGGG-3′	78	
Spermatogonial stem-cell renewal factor			
	F: 5′-ATCTGTGCCGCTCATCGTCTCGT-3′	72	121
	R: 5′- CCCTGCCTCCTCCTGGGTCTTC-3′	74	
Saposin			
	F: 5′-CCATAAAGTTCTGCCCCCACCAC-3′	68	145
	R: 5′-CCCTTCCACACCTTCGTTTCACA-3′	70	
EF-1α			
Standard	F: 5′-GCTCTTACCGAGGCTGTCCC-3′	66	434
	R: 5′-GTGGGTGTAATCAAGGAGGTCAA-3′	68	
Real-time	F: 5′-TTCCGACTCCAAGAACGACC-3′	62	122
	R: 5′-GAGCAGTGTGGCAATCAAGC-3′	62	

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

At the end of the incubation period, 2 µl of the stop solution was added and mixed thoroughly. Labeled samples were purified using a QIAquick PCR Purification Kit. Subsequently, cDNA was labeled with Cy3 to represent the control sample and Cy5 to represent the tested sample at room temperature for 1 hr., Cy3/Cy5 labeled cDNA were purified prior to hybridization using a LabelStar Array Kit, cDNA cleanup Module (QIAGEN).

2.9.3 Hybridization and detection

The hybridization buffer (5X SSC, 4X Denhardt's solution and 0.5% SDS) was heated at 95°C for 2 min and placed on ice for 1 min. Poly A and formamide (10% v/v) were added and pre-hybridized at 42 °C for 15 min with a microarray chip (5776 spots representing 2036 ESTs from hemocyte cDNA library of *P. monodon* and *M. japonicus*). Cy3/Cy5 labeled cDNA was added and further hybridized at 42°C for 16 - 20 hr. The microarray chip was post-hybridization washed once with 5X SSC/0.1% SDS at 30°C for 10 min, twice with 0.5X SSC at 30°C for 2 min and once with 0.5X SSC/0.01% Tween20 for 2 min. The microarray chip was air-dried and scanned with GenePix 4000B Array Scanner (Axon Instruments, Inc.). The signal intensities were converted to approximate measurement of absolute expression by subtracting the background signal levels and normalized with the signal levels of the positive control (*EF-1α* and/or *β-actin*) across different arrays.

2.10 *In vitro* expression of the full length cDNA using the bacterial expression system

2.10.1 Designation of primers

A primer pair was designed to amplify the mature full length cDNA of *cyclophilin A, SUMO, spermatogonial stem-cell renewal factor*, *PGMRC1* and *Dmc1*. The forward and reverse primers containing a *Nde* I site and an *Eco* RI site and six His encoded nucleotides were shown in Table 2.14, respectively.

2.10.2 Construction of recombinant plasmid in cloning and expression vectors

The mature cDNA of *cyclophilin A, SUMO-1, spermatogonial stem-cell renewal factor*, *PGMRC1*, and *Dmc1* were amplified, ligated, cloned into pGEM-T easy vector and transformed into *E. coli* JM109. Plasmid DNA of the positive clones was sequenced to confirm the orientation of recombinant clones and used as the template for amplification using the forward and reverse primers for cloning into the pET32a expression vector.

Table 2.14 Nucleotide sequences of primers used for *in vitro* expression of *cyclophilin A, SUMO-1, spermatogonial stem-cell renewal factor*, *PGRMC1* and *Dmc1* in *P. monodon*

Primer	Sequence
ORF	
Cyclophilin A-ORF	F: 5′-ATGGGCAACCCCAAAGTCTTTTTCGA-3′
	R: 5′-TTACAGCTGGCCGCAGTTGGCG-3′
SUMO-ORF	F: 5′-ATGTCTGATAACACTGACGCCAAGC-3′
	R: 5′- TCAATGGCCGCCGGTCTG-3′
Spermatogonial-ORF	F:5′-ATGAGCGCTGCACAGACCTCTCA-3′
	R:5′-CTAGACAACGCGAGCGGCAAC-3′
PGRMC1-ORF	F:5′-ATGGCGGACGAGGGAGCG-3′
	R:5′-CTACTAATCATCCGTCTTCGCTTTGGT-3′
DMC1-ORF	F:5′-ATGGAAGATCAGGCTTTAGATGC-3′
	R:5′-TTACTCCTTAGCATCAGCAATGC-3′
ORF+HIStag+Restrictionsites	
CycA-*Nde*I	F:5′-TTT CAT ATG GGC AAC CCC AAA GTC TTT TTC GA-3′
CycA-*Eco*RI+His	R:5′-AAA GAA TTC TTA **ATG ATG ATG ATG ATG GTG** CAG CTG GCC GCA GTT GGC G-3′
SUMO-*Nde*I	F:5′-TTT CAT ATG TCT GAT AAC ACT GAC GCC AAG C-3′
SUMO-*Eco*RI+His	R:5′-AAA GAATTC TCA **ATG ATG ATG ATG ATG GTG** ATG GCC GCC GGT CTG-3′
Spermatogonial-*Nde*I	F:5′-TTT CAT ATG AGC GCT GCA CAG ACC TCT CA-3′
Spermatogonial-*Eco*RI+His	R:5′-AAA GAA TTC CTA **ATG ATG ATG ATG ATG GTG** GAC AAC GCG AGC GGC AAC-3′
PGRMC1-*Nde*I	F:5′-TTT CAT ATG GCG GAC GAG GGA GCG GAC-3′
PGRMC1-*Eco*RI+His	R:5′-AAA GAA TTC CTA **ATG ATG ATG ATG ATG GTG** ATC ATC CGT CTT CGC-3′
DMC1-*Nde*I	F:5′-TTT CAT ATG GAA GAT CAG GCT TTA GAT GC-3′
DMC1-*Eco*RI+His	R:5′-AAA GAA TTC TTA **ATG ATG ATG ATG ATG GTG** CTC CTT AGC ATC AGC AAT GC-3′

The amplification product was digested with *Nde* I and *Eco* RI and analyzed by agarose gel electrophoresis. The gel-eluted product was ligated into pET32a and transformed into *E. coli* JM109. Plasmid DNA of the positive clones was subsequently transformed into *E. coli* BL21(DE3)codon+ RIPL.

2.10.3 Expression of recombinant proteins

A single colony of recombinant *E. coli* BL21(DE3) codon+ RIPL carrying desired recombinant plasmid was inoculated into 3 ml of LB medium, containing 50 μg/ml ampicillin and 34 μg/ml chloramphinical at 37 °C and 50 μl of the overnight culture was transferred to 50 ml of LB medium containing 50 μg/ml ampicillin 34 μg/ml chloramphinical and further incubated to an OD_{600} of 0.4-0.6. After IPTG induction (1.0 mM final concentration), appropriate volume of the culture corresponding to the OD of 1.0 was time-interval taken (1, 2, 3, 6, 12 and 24 hr) and centrifuged at 12000 *g* for 1 min. The pellet was resuspended and examined by 15% SDS-PAGE (Laemmli, 1970). In addition, aliquots of 1 ml of the IPTG induced-culture (3 or 6 hr) were centrifuged, resuspended in 1X PBS and repeated freeze-thaw 4 times in liquid nitrogen. The protein concentration of both soluble and insoluble portions was measured using a dye-binding assay (Bradford, 1972). Overexpression of the recombinant protein was analyzed by 15% SDS-PAGE.

2.10.4 Detection of recombinant proteins

Recombinant protein was analyzed in 15% SDS-PAGE. The electrophoresed proteins were transferred to a PVDF membrane (Hybond P; GE Healthcare) (Towbin, 1979). The membrane was washed three times with 1X TBST (100 ml of 10X TBS, 900 μl of H_2O and 500 μl of Tween20) for 5 min, blocked with 20 ml of a blocking buffer (1.0 g of BSA in 20 ml of 1X TBST) and incubate for 1 hr at room temperature with gentle shaking. The membrane was washed three times in 1xTBST and incubated with diluted Anti-His (GE Healthcare; 1:5,000) in the blocking buffer for 1 hr. The membrane was incubated with diluted Anti-mouse-IgG-AP Conjugate (Promega; 1:10,000) in the blocking buffer for 1 hr. The AP activity was detected by adding BCIP/NBT. The membrane was incubated in the dark place for 2-15 min.

2.10.5 Mass spectrometric analysis of recombinant proteins

The rPMCYA and rPMSUMO-1 were resolved in 15% SDS-PAGE and the protein bands at the expected molecular mass of 18.86 kDa and 8.78 kDa were excised from the gel. In-gel trypsin digestion was performed as described elsewhere (Shevchenko et al., 1996). The digests were injected into the mass spectrometer through a 100×0.18 mm, 5µm BioBasic C18 Kappa column (Thermo Electron) flowing at 800 nL/min. Peptides were eluted from the column using an acetonitrile/0.1% formic acid gradient (2-65% acetonitrile over 40 min). The peptide spectra were measured by a liner ion trap mass spectrometer (Finnigan LTQ) equipped with electrospray ionization and searched against the non-redundant protein database (http://www.ncbi.nlm.nih.gov) using SEQUEST.

2.10.6 Purification of recombinant proteins

Recombinant protein was purified by using a His GraviTrap kit (GE Healthcare). Initially 250 ml of IPTG-induced culture were harvested by centrifugation at 5000 rpm for 15 min. The pellet was resuspended in the binding buffer (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4), sonicated and centrifuged at 14000 rpm for 30 min. The soluble and insoluble fractions were separated. Soluble or insoluble fraction composed of the recombinant protein was loaded into column. The column was washed with 10 ml of binding buffer containing 20 mM imidazole (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4) and 5-10 ml of the binding buffer containing 80 mM imidazole (20 mM sodium phosphate, 500 mM NaCl, 80 mM imidazole, pH 7.4). After that the recombinant protein was eluted with 6 ml of the elution buffer (20 mM sodium phosphate, 500 mM NaCl, 500 mM imidazole, pH 7.4). Each fraction of the washing and eluting step were analyzed by SDS-PAGE and western blotting. The recombinant proteins in the insoluble fraction were purified under denaturing conditions. The purified proteins were stored at -20 ^{o}C.

2.10.7 Polyclonal antibody production

Polyclonal antibody was commercially produced from the purified recombinant protein in rabbit by Faculty of Associated Medical Sciences, Changmai University.

CHAPTER III

RESULTS

3.1 Construction of a conventional testis cDNA library of *P. monodon*

Genes related with testicular development of *P. monodon* were identified by an EST analysis. A conventional testis cDNA library of *P. monodon* broodstock was established. Initially, mRNA was purified from pooled total RNA extracted from testis of *P. monodon* broodstock (Fig. 3.1). Testis mRNA (7.5 µg) were used for cDNA library construction. First and second strand cDNA were synthesized and manipulated before cNDA was sized-fractionated by gel filtration (Sepharose CL-2B). Fractions were collected and a 5 µl aliquot of each fraction was examined in 5% native polyacrylamide gels. Fractions 4 - 6 of testes cDNA which were free from adaptors (Fig. 3.2) were selected, pooled and subjected to ligation of the cDNA insert and preparation of a lambda library, respectively.

Inserted sizes of randomly selected clones were examined by colony PCR. Clones carrying the inserted sizes > 500 bp (excluding approximately 340 bp of the vector) were further analyzed (Fig. 3.3). Plasmid DNA is extracted and

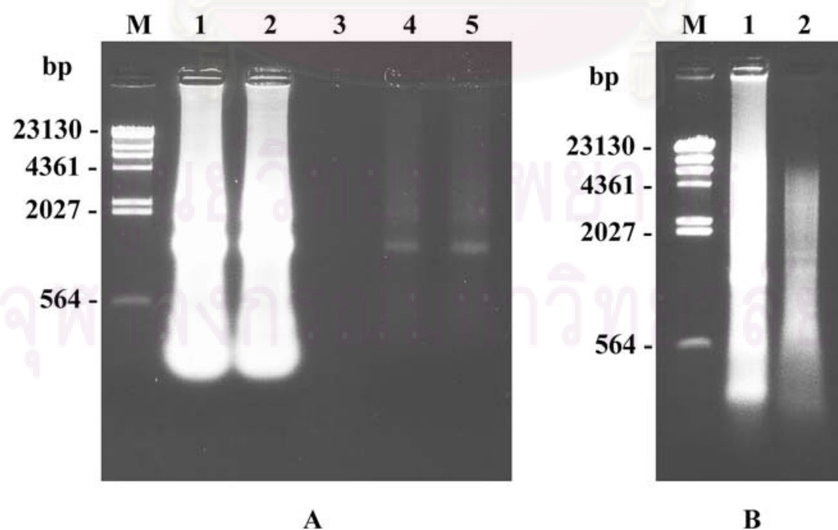


Figure 3.1 Total RNA (A, lanes 1-2) and mRNA (A, lanes 4-5) of *P. monodon* testes were used to construct testis cDNA library. First (B, lane 1) and second (B, lanes 2) strand cDNA were synthesized from purified mRNA.

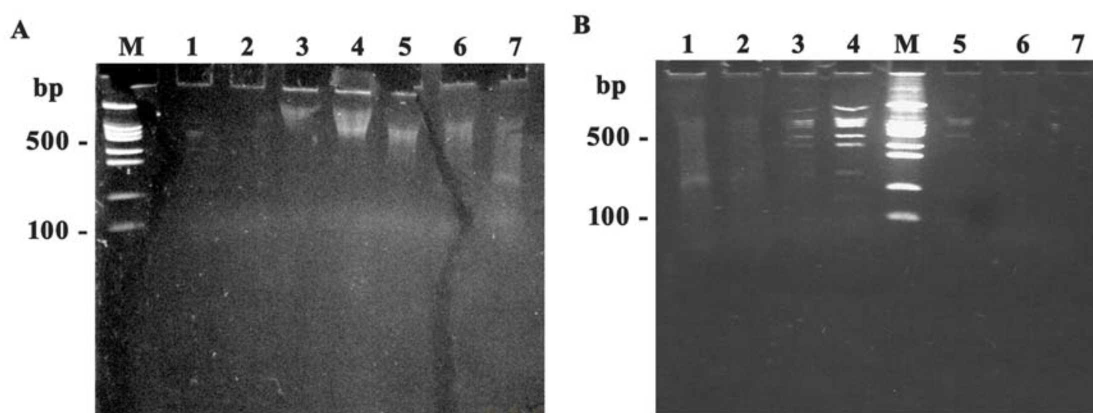


Figure 3.2 Size-fractionation of testes cDNA of *P. monodon*. Lanes 1-7 (A) and 1-4 and 5-7 (B) correspond to fractions 2-8, 9-12 and 13-15, respectively. Lanes M = a 100 bp marker.

unidirectionally sequenced from the 5′ direction. Nucleotide sequences were compared with those previously deposited in the GenBank using BLAST*N* and BLAST*X*.

The primary titer of a testis cDNA library of *P. monodon* was 5.7 x 10^6 pfu/ml. The percentage of positive clones was 94.0% (1366/1453 clones) and 73.6% of which carrying the inserts greater than 500 bp in size. From 896 recombinant clones sequenced, 606 ESTs (67.6%) corresponded to known sequences (E-value < 1e-04) whereas 290 sequences were regarded as novel (unknown) transcripts (32.4%, E-value > 1e-04). The relative discovery rate of new genes was approximately 30% for every 100 ESTs examined (Fig. 3.4). Six hundred and one transcripts (109 contigs and 492 singletons with the average length of 823 and 578 nucleotides) were obtained after sequence assembling.

Six hundred and six known transcripts from the conventional testis cDNA library of *P. monodon* were functionally categorized to 10 groups (Table 3.1). Disregarding ribosomal (16.8%) and hypothetical (newly unidentified transcripts, 11.3%) protein homologues, matched ESTs categorized as members of gene expression and protein synthesis (9.6%) predominated followed by those classified as

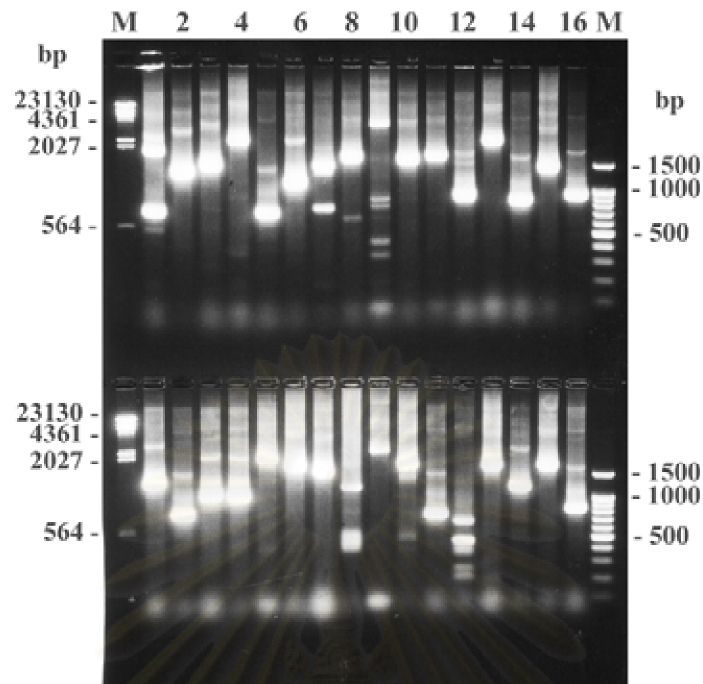


Figure 3.3 Colony PCR for determining sizes of inserts of positive clones from a conventional testis cDNA library of *P. monodon*.

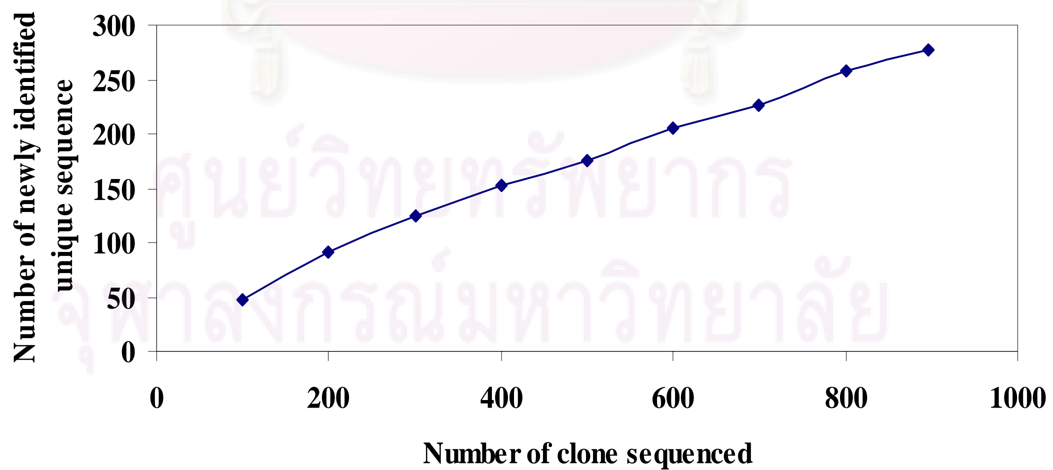


Figure 3.4 The relative discovery rate of new genes from a conventional testis cDNA library of *P. monodon*.

members of miscellaneous function (6.3%), metabolism (6.7%), internal/external structure (4.5%), and defense and homeostasis (3.3%). The remaining ESTs allocated to other functional categories (cell division/DNA synthesis, transport or mitochondrial proteins) were accounted for less than 3.2% of the characterized ESTs in this library.

Gene homologues functionally categorized to members of cell division/DNA synthesis, repair and replication, defense and homeostasis, gene expression, regulation and protein synthesis, internal/external structure and motility, metabolism**,** miscellaneous function**,** mitochondrial protein, transport, ribosomal and rRNA, newly unidentified (hypothetical) proteins, and unknown proteins are illustrated in Appendix A.

Transcripts involved in sex determination and sex differentiation and testicular development were found in this library such as *transformer-2*, *serine/threonine protein kinase* family*, B-cell receptor-associated protein 37* (*prohibitin-2* or *repressor of estrogen receptor activity*). In addition, those associated with flagellar architecture of sperm including *dynactin subunit 5* and *dynein* were also found.

Table 3.1 Functional categories of ESTs significantly matched with data in GenBank of testis cDNA library

	Functional category	No. of clones (%)
1	Gene expression, regulation and protein synthesis	86(9.6)
2	Internal / external structure and motility	40(4.5)
3	Metabolism	60(6.7)
4	Defense and homeostasis	30(3.3)
5	Cell division / DNA synthesis, repair and replication	29(3.2)
6	Ribosomal and rRNA	151(16.8)
7	Mitochondrial protein	25(2.8)
8	Transport	28(3.1)
9	Miscellaneous function	56(6.3)
10	Unidentified (hypothetical) – similar to other cDNA/DNA	101(11.3)
11	Unknown	290(32.4)
Total		**896(100)**

Table 3.2 Functionally important gene homologues found in a conventional testis cDNA library of *P. monodon*

Clone No.	Transcripts	Closest species	E-value
TT-N-S01-0483-W	*26S protease regulatory subunit*	*Aedes aegypti*	1e-119
TT-N-S01-0385-W	*26S Proteasome non-ATPase regulatory subunit 3 (Diphenol oxidase A2 component)*	*Apis mellifera*	6e-69
TT-N-S01-0283-W	*26S proteasome regulatory complex subunit p48B*	*Drosophila melanogaster*	2e-50
TT-N-S01-0638-W	*26S proteasome subunit P45 family protein*	*Tetrahymena thermophila SB210*	5e-67
TT-N-S01-0373-W	*2-Cys thioredoxin peroxidase*	*Aedes aegypti*	1e-35
TT-N-S01-0813-W			5e-63
TT-N-S01-0076-W	*Actin-binding protein anillin, contractile ring component anillin*	*Xenopus laevis*	3e-56
TT-N-S01-0447-W	*Acyl-CoA oxidase (ACX3)*	*Tetrahymena thermophila SB210*	7e-66
TT-N-S01-0177-W	*Adaptor-related protein complex 1, beta 1 subunit, isoform CRA_c*	*Homo sapiens*	2e-38
TT-N-S01-0917-W	*Arginyl-tRNA--protein transferase 1 (Arginyltransferase 1) (Arginine-tRNA--protein transferase 1)*	*Apis mellifera*	5e-43
TT-N-S01-0090-W	*B-cell receptor-associated protein 37 (Prohibitin-2)(Repressor of estrogen receptor activity*	*Tribolium castaneum*	3e-84
TT-N-S01-0900-W	*Bromodomain containing 8 (Skeletal muscle abundant protein, SMAP, SMAP2, Thyroid hormone receptor coactivating protein 120kDa, TrCP120)*	*Apis mellifera*	3e-36
TT-N-S01-0153-W	*BUB3 budding uninhibited by benzimidazoles 3 homolog*	*Xenopus tropicalis*	2e-72
TT-N-S01-0008-W	*Carbonyl reductase 1-like*	*Tribolium castaneum*	2e-68
TT-N-S01-0190-W	*Cell division cycle 2*	*Danio rerio*	2e-31
TT-N-S01-0169-W	*Cell division protein kinase 7 (CDK-activating kinase)*	*Mus musculus*	8e-85
TT-N-S01-0752-W	*checkpoint kinase 1 (Serine/threonine-protein kinase)*	*Strongylocentrotus purpuratus*	9e-76
TT-N-S01-0525-W	*CHK1 checkpoint homolog (Serine/threonine-protein kinase)*	*Xenopus tropicalis*	3e-06
TT-N-S01-0567-W	*Chromobox protein homolog 1 (Heterochromatin protein 1 homolog beta) (HP1 beta) (Modifier 1 protein) (M31) (Heterochromatin protein p25)*	*Apis mellifera*	1e-58
TT-N-S01-0475-W	*COP9 constitutive photomorphogenic homolog subunit 5 isoform 1*	*Apis mellifera*	5e-73
TT-N-S01-0695-W	*Cyclin dependent kinase 2*	*Sphaerechinus granularis*	8e-82
TT-N-S01-0020-W	*Cyclophilin A*	*Chlamys farreri*	7e-74
TT-N-S01-0217-W	*Cystathionine gamma-lyase*	*Rattus norvegicus*	1e-52

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.2 (cont.)

Clone No.	Transcripts	Closest species	E-value
TT-N-S01-0933-W	*Der1-like domain family member 1 (Degradation in endoplasmic reticulum protein 1, DER1)*	*Bombyx mori*	1e-59
TT-N-S01-0067-W	*Dynactin 5 (p25)*	*Strongylocentrotus purpuratus*	2e-63
TT-N-S01-0880-W	*Dynein light intermediate chain*	*Aedes aegypti*	6e-80
TT-N-S01-0444-W	*E1B-55kDa-associated protein 5 isoform 5*	*Pan troglodytes*	1e-51
TT-N-S01-1055-W	*Flotillin 2 CG32593-PB, isoform B*	*Drosophila melanogaster*	4e-48
TT-N-S01-0189-W	*Growth factor receptor-bound protein*	*Aedes aegypti*	7e-55
TT-N-S01-0679-W	*Heat shock 70kDa protein 8 isoform 2*	*Homo sapiens*	6e-66
TT-N-S01-0499-W	*Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 9*	*synthetic construct*	3e-55
TT-N-S01-0060-W	*Importin 7*	*Aedes aegypti*	1e-67
TT-N-S01-0586-W	*Inhibitor of apoptosis protein*	*Bombyx mori*	3e-19
TT-N-S01-0121-W	*Innexin inx1*	*Schistocerca americana*	3e-72
TT-N-S01-0246-W			3e-17
TT-N-S01-1038-W			5e-11
TT-N-S01-1001-W	*Karyopherin (importin) alpha 2*	*Ictalurus punctatus*	3e-65
TT-N-S01-0134-W	*Laminin beta chain*	*Schistocerca gregaria*	1e-43
TT-N-S01-0572-W	*Long chain acyl-CoA synthetase*	*Oryza sativa (japonica cultivar-group)*	1e-33
TT-N-S01-0071-W	*Low molecular weight neurofilament protein (PMTT1)*	*Xenopus laevis*	8e-04
TT-N-S01-0437-W	*Metaxin 2*	*Tribolium castaneum*	1e-79
TT-N-S01-0773-W	*Microtubule-associated protein 1 light chain 3 alpha*	*Xenopus tropicalis*	1e-20
TT-N-S01-0730-W	*Multiple inositol polyphosphate phosphatase*	*Aedes aegypti*	8e-22
TT-N-S01-0004-W	*Multiple inositol polyphosphate phosphatase 2*	*Apis mellifera*	2e-09
TT-N-S01-0495-W			8e-09
TT-N-S01-0678-W	*Multiple inositol polyphosphate phosphatase 2; MIPP2*	*Drosophila melanogaster*	1e-07
TT-N-S01-1007-W	*Multiprotein bridging factor 1*	*Bombyx mori*	4e-50
TT-N-S01-1026-W			4e-50
TT-N-S01-0371-W	*Novel protein similar to vertebrate PCTAIRE protein kinase 2 (PCTK2)*	*Danio rerio*	2e-95
TT-N-S01-0733-W	*Peptidyl-prolyl cis-trans isomerase (Cyclophilin 1)*	*Bombyx mori*	6e-72
TT-N-S01-0052-W	*Polyadenylate binding protein II*	*Apis mellifera*	6e-74
TT-N-S01-0119-W	*Profilin (Chickadee protein)*	*Tribolium castaneum*	1e-37
TT-N-S01-0710-W			1e-37
TT-N-S01-0030-W	*Programmed cell death protein*	*Aedes aegypti*	7e-63

Table 3.2 (cont.)

Clone No.	Transcripts	Closest species	E-value
TT-N-S01-1067-W	*Proteasome (prosome, macropain) 26S subunit, non-ATPase, 13*	*Apis mellifera*	6e-34
TT-N-S01-0554-W	*Proteasome (prosome, macropain) subunit, alpha type, 1, isoform CRA_a*	*Homo sapiens*	3e-07
TT-N-S01-0287-W	*Proteasome alpha 3 subunit*	*Bombyx mori*	2e-73
TT-N-S01-0405-W	*Proteasome subunit alpha type*	*Aedes aegypti*	8e-99
TT-N-S01-0812-W			3e-65
TT-N-S01-0949-W			2e-80
TT-N-S01-0063-W	*Proteasome subunit alpha type 1 (Proteasome component C2) (Macropain subunit C2) (Multicatalytic endopeptidase complex subunit C2)*	*Canis familiaris*	6e-06
TT-N-S01-0369-W	*Proteasome subunit beta type 1 (Proteasome 26 kDa subunit)*	*Tribolium castaneum*	5e-26
TT-N-S01-0916-W	*Proteasome subunit beta type 2 (Proteasome component C7-I) (Macropain subunit C7-I)*	*Tribolium castaneum*	6e-64
TT-N-S01-0879-W	*Proteasome subunit, alpha type, 5*	*Apis mellifera*	1e-20
TT-N-S01-0087-W	*Protein mago nashi (mago-nashi homolog, proliferation-associated) MAGOH*	*Apis mellifera*	3e-76
TT-N-S01-0974-W			4e-60
TT-N-S01-0863-W	*Protein serine/threonine kinase*	*Dictyostelium discoideum AX4*	3e-22
TT-N-S01-0957-W	*RAB, member of RAS oncogene family-like 3*	*Apis mellifera*	7e-64
TT-N-S01-1036-W	*Rac GTPase activating protein 1 isoform 1*	*Canis familiaris*	2e-27
TT-N-S01-1040-W			4e-28
TT-N-S01-1048-W			2e-23
TT-N-S01-0556-W	*RAD1 homolog isoform 1*	*Rattus norvegicus*	5e-13
TT-N-S01-0220-W	*RAS protein*	*Bombyx mori*	1e-56
TT-N-S01-0203-W	*RAS-related GTP binding protein*	*Bombyx mori*	4e-62
TT-N-S01-0118-W	*Receptor for activated protein kinase C-like*	*Blattella germanica*	1e-122
TT-N-S01-0039-W	*Ring finger protein 20*	*Gallus gallus*	8e-06
TT-N-S01-1009-W	*Ring finger protein 20 isoform 3*	*Macaca mulatta*	1e-26
TT-N-S01-0905-W	*Ring finger protein 44 isoform 3*	*Pan troglodytes*	7e-44
TT-N-S01-0847-W	*Ruvbl2-prov protein (RuvB-like DNA helicase reptin)*	*Xenopus laevis*	1e-115
TT-N-S01-0150-W	*Serine palmitoyl transferase LCB2 subunit*	*Drosophila melanogaster*	4e-34
TT-N-S01-0427-W	*Serine/arginine repetitive matrix 1*	*Gallus gallus*	4e-05
TT-N-S01-0667-W	*Serine/threonine protein kinase Pto (Pto-like serine/threonine kinase)*	*Lycopersicon esculentum*	3e-07
TT-N-S01-0903-W	*Serine/threonine-protein kinase 23 (Muscle-specific serine kinase 1) (MSSK-1)*	*Apis mellifera*	4e-90
TT-N-S01-0843-W	*Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform (PP2A-beta)*	*Tribolium castaneum*	1e-109

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.2 (cont.)

Clone No.	Transcripts	Closest species	E-value
TT-N-S01-0430-W	*IMP4, U3 small nucleolar ribonucleoprotein, homolog*	*Apis mellifera*	2e-48
TT-N-S01-0206-W	*LSM4 homolog, U6 small nuclear RNA associated*	*Danio rerio*	7e-46
TT-N-S01-0205-W	*U2 small nuclear ribonucleoprotein auxiliary factor 2 isoform 1*	*Bombyx mori*	2e-62
TT-N-S01-0967-W	*Small nuclear ribonucleoprotein D2-like protein*	*Toxoptera citricida*	5e-41
TT-N-S01-0673-W TT-N-S01-0915-W	*Small nuclear ribonucleoprotein E*	*Bombyx mori*	1e-31 1e-15
TT-N-S01-0873-W	*Small nuclear ribonucleoprotein polypeptide G*	*Homo sapiens*	2e-23
TT-N-S01-0259-W	*Spermidine synthase*	*Gallus gallus*	2e-39
TT-N-S01-0144-W TT-N-S01-0175-W TT-N-S01-0426-W	*SUMO, small ubiquitin-like modifier, SUMO, small ubiquitin-like modifier SMO-1 (10.2 kD)*	*Apis mellifera*	5e-38 8e-38 5e-38
TT-N-S01-0626-W	*Synaptobrevin-like protein 1*	*Canis familiaris*	1e-48
TT-N-S01-0414-W	*T-complex protein 1, alpha subunit(TCP-1-alpha)(CCT-alpha)*	*Delia antique*	5e-73
TT-N-S01-0702-W	*Tetratricopeptide repeat domain 9C (TTC9)*	*Rattus norvegicus*	3e-23
TT-N-S01-0897-W	*Thioredoxin 1*	*Litopenaeus vannamei*	1e-51
TT-N-S01-0232-W	*Thyroid hormone receptor-associated protein complex 240 kDa component (Trap240) (Thyroid hormone receptor associated protein 1)*	*Canis familiaris*	2e-57
TT-N-S01-0407-W	*TPA_inf: troponin I isoform a2*	*Drosophila pseudoobscura*	1e-26
TT-N-S01-0941-W	*Transformer-2 protein A*	*Bombyx mori*	3e-34
TT-N-S01-0985-W	*Transformer-2 protein B*	*Bombyx mori*	6e-57
TT-N-S01-0223-W	*Ubiquitin carboxyl-terminal hydrolase 14 (Ubiquitin thiolesterase 14) (Ubiquitin-specific processing protease 14)*	*Apis mellifera*	2e-51
TT-N-S01-0972-W	*Ubiquitin carboxyl-terminal hydrolase 5 (Ubiquitin thiolesterase 5) (Ubiquitin-specific-processing protease 5)(Isopeptidase T) isoform 2*	*Rattus norvegicus*	4e-65
TT-N-S01-0872-W	*Ubiquitin-conjugating enzyme E2*	*Aedes aegypti*	1e-58
TT-N-S01-0657-W	*WD repeat domain 61 (Meiotic recombination REC14 protein homolog)(WDR61)*	*Bombyx mori*	3e-64
TT-N-S01-0883-W	*Zinc finger protein 420*	*Danio rerio*	8e-44
TT-N-S01-0874-W	*Zinc finger protein 420 isoform 1*	*Mus musculus*	2e-33
TT-N-S01-0573-W	*Zinc finger protein 501*	*Pongo pygmaeus*	4e-45
TT-N-S01-0578-W	*Zinc finger, ZZ domain containing 3*	*Apis mellifera*	3e-48

Moreover, transcripts involving meiotic cell division such as *cell division cycle 2* (*cdc2* also called *cdk1*), *cell division protein kinase 7* (*cdk7*), and *cyclin dependent kinase 2* (*cdk2*) and those involving apoptosis, for example, *programmed cell death protein 7* and *inhibitor of apoptosis protein* were also found (Table 3.2). Examples of other functionally important transcripts found in this library were *serine/threonine protein kinase* family, *ring finger protein*, *small nucleolar ribonucleoprotein,* and several *proteasome subunits.*

3.2 Construction of suppression subtractive hybridization (SSH) cDNA libraries from testes of juvenile and broodstock of *P. monodon*

Forward and reverse subtractions between testis cDNAs of broodstock and juvenile *P. monodon* (cDNAs of broodstock as the tester whereas those of juveniles as the driver and *vice versa*) were carried out. Initially, first and second strand cDNA were separately synthesized from mRNA of broodstock and juveniles *P. monodon*. Double strand cDNA of tester and driver was digested with *Rsa* I (Fig. 3.5). Only digested testers were ligated with the adapter. Tester and driver were hybridized twice. Each subtracted cDNA was sequentially amplified twice using suppression PCR (Fig. 3.6). The resulting products are ligated to pGEM-T Easy vector and transformed to *E. coli* JM109. Recombinant clones of both forward and reverse SSH library were identified by colony PCR (Figs. 3.7 and 3.8). Plasmid DNA is extracted and unidirectionally sequenced.

A total of 367 ESTs of testes (178 and 189 clones from the forward and reverse SSH libraries, respectively) of *P. monodon* were unidirectionally sequenced. The positive recombinant clones of both libraries were 95.1 and 96.4%. Of which 82.4 and 86.2% had insert sizes > 250 bp in length. The percentage of ESTs significantly matched known genes in respective libraries was 37.1 (67 ESTs) and 54.0% (104 ESTs). Unknown transcripts predominated in both forward and reverse SSH libraries. They were 112 and 87 ESTs accounting for 62.9 and 46.0%, respectively (E-value > 1e-04) (Table 3.3).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

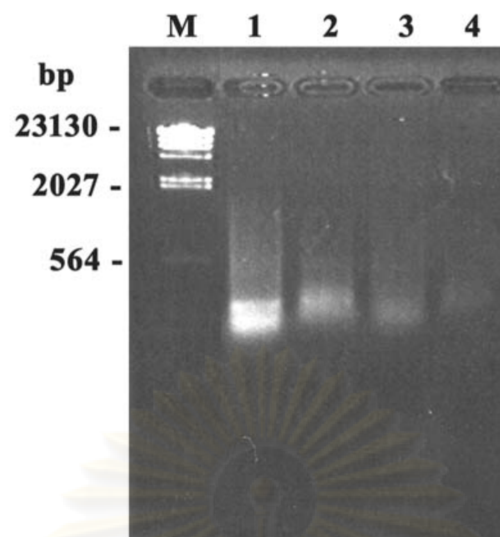


Figure 3.5 Agarose gel electrophoresis illustrating digested cDNA of *P. monodon* with *Rsa* I. Lanes 1 and 3 = undigested testis cDNA of broodstock-sized and juvenile *P. monodon*, respectively. Lanes 2 and 4 = *Rsa* I-digested testis cDNA of broodstock-sized and juvenile *P. monodon*, respectively.

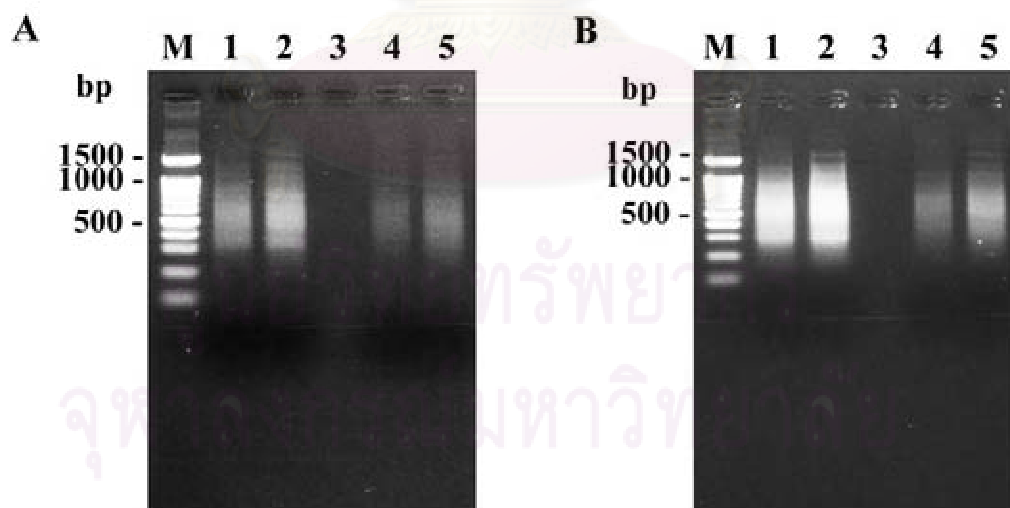


Figure 3.6 Primary (A; lanes 1-2) and secondary (B; lanes 4-5) suppression PCR of subtracted (lanes 1) and nonsubtracted (lanes 2) of forward (A) and reverse (B) SSH of testis cDNA of *P. monodon*.

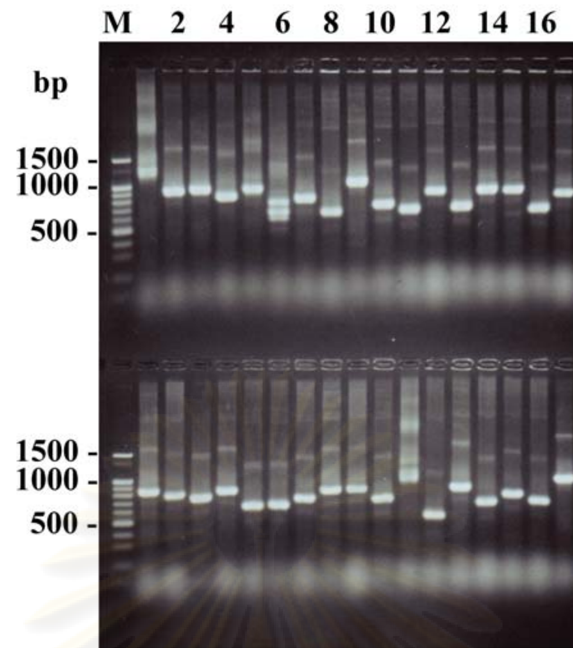


Figure 3.7 Colony PCR for determining sizes of inserts of positive cones clones from the forward SSH testis library of *P. monodon*.

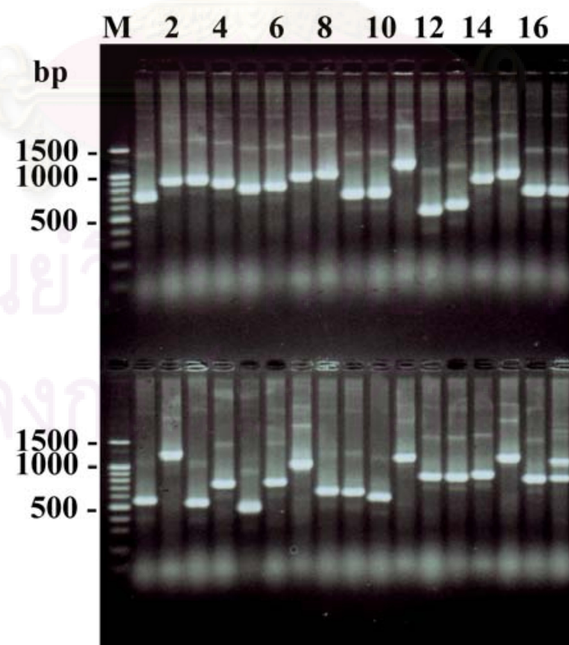


Figure 3.8 Colony PCR for determining sizes of inserts of positive cones clones from the reverse SSH testis library of *P. monodon*.

Table 3.3 Functional categories of ESTs in testes of *P. monodon* identified by SSH analysis

Category	Forward SSH (%)	Reverse SSH (%)	Both libraries
Matched EST	**67(37.1)**	**104(54.0)**	**171**
Sex related genes	1(0.6)	1(0.5)	2
Stress response and cell defense protein	4(2.2)	13(6.9)	17
Protein synthesis and DNA replication	14(7.9)	22(11.6)	36
Internal / external structure, motility, and Transport	2(1.1)	7(3.7)	9
Metabolism	18(10.1)	13(6.9)	31
Ribosomal proteins	18(10.1)	36(19.1)	54
Unidentified functions	9(5.1)	10(5.3)	19
Unmatched EST	**112(62.9)**	**87(46.0)**	**199**
Total EST	**178(100)**	**189(100)**	**367**

Seven known transcripts: *allergen Pen m 2* (*tyrosine kinase*), *COI*, *EF-1α*, *GTP-binding protein, 26S proteasome non-ATPase subunit 12, receptor for activated protein kinase C* (*RACK*) and *myelodysplasia/myeloid leukemia factor*, were found with low frequencies in both libraries suggesting that the cDNA subtraction was successful (Tables 3.4 and 3.5).

ESTs significantly matched known genes from both libraries were functionally categorized to 7 groups (Table 3.3). Disregarding ribosomal and hypothetical (functionally unidentified) protein homologues, ESTs categorized as members of gene expression and protein synthesis (14 and 22 ESTs accounting for 7.9 and 11.6% in the forward and reverse SSH libraries, respectively) and metabolism (18 and 13 ESTs accounting for 10.1 and 6.9%) predominated among known transcripts in both SSH libraries. Numbers of transcripts belonged to internal/external structure and stress response/cell defense groups of the forward SSH library (2 and 4 ESTs accounting for 1.1 and 2.2%) were lower than those of the reverse SSH library (7 and 13 ESTs accounting for 3.7 and 6.9%).

Highly redundant ESTs were not observed and a relatively large number of known gene homologues were found in these libraries (Tables 3.4 and 3.5). The diversity of genes found in these libraries demonstrated the promise in discovery of genes with functional importance by a SSH approach.

Sequence assembly revealed that 112 unknown transcripts of the forward SSH library were composed of 12 contigs and 55 singletons, whereas 87 unknown transcripts of the reverse SSH library were clustered to 12 contigs and 60 singletons.

Sex-related transcripts including *meiotic recombination protein DMC1/LIM15 homolog isoform 1*, *progestin receptor membrane component 1*, and *innexin 2* were found in both SSH libraries. *Meiotic recombination protein DMC1/LIM15 homolog isoform 1* (*Dmc1*) is involved in meiotic recombination occurred during the meiotic prophase (Ozaki *et al*., 2006). *Progestins* are sex steroid hormones that play important roles in gametogenesis. In fish, *progestin* also plays an important role in spermiation and sperm maturation (Miura *et al*., 2006). Two totally distinct classes of putative membrane-bound progestin receptors have been reported in vertebrates: progestin membrane receptor component (PGMRC; subtypes 1 and 2) and membrane progestin receptors (mPR; subtypes α, β, γ). Both have never been studied in any crustacean (Mourot *et al*., 2006).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.4 Known transcripts found in the forward SSH library of testes of *P. monodon*

Clone No.	Transcripts	Closest Species	E-value
TT-N-ST01-0155-W	*26S proteasome regulatory complex ATPase RPT4*	*Aedes aegypti*	2e-81
TT-N-ST01-0077-W	*Actin-depolymerizing factor 1*	*Bombyx mori*	8e-21
TT-N-ST01-0116-W	*Activated protein kinase C receptor*	*Toxoptera citricida*	4e-70
TT-N-ST01-0148-W	*Allergen Pen m 2*	*Penaeus monodon*	5e-22
TT-N-ST01-0150-W	*Antimicrobial peptide*	*Fenneropenaeus chinensis*	2e-22
TT-N-ST01-0156-W	*ATP-dependent RNA helicase*	*Aedes aegypti*	3e-76
TT-N-ST01-0085-W	*Basic leucine zipper and W2 domain-containing protein 2*	*Danio rerio*	2e-09
TT-N-ST01-0109-W	*Bmsqd-2*	*Apis mellifera*	1e-102
TT-N-ST01-0149-W	*C-1-tetrahydrofolate synthase, cytoplasmic (C1-THF synthase)*	*Pongo pygmaeus*	7e-85
TT-N-ST01-0147-W	*C2 domain containing protein*	*Tetrahymena thermophila SB210*	2e-19
TT-N-ST01-0053-W	*Cytochrome c oxidase subunit 6a polypeptide 1*	*Xenopus tropicalis*	2e-08
TT-N-ST01-0007-W	*Cytochrome c oxidase subunit I*	*Fenneropenaeus chinensis*	4e-76
TT-N-ST01-0137-W		*Marsupenaeus japonicus*	1e-112
TT-N-ST01-0030-W	*Cytochrome c oxidase subunit III*	*Penaeus monodon*	1e-70
TT-N-ST01-0119-W	*Cytosolic manganese superoxide dismutase*	*Penaeus monodon*	2e-12
TT-N-ST01-0141-W			6e-13
TT-N-ST01-0131-W	*DEAD (Asp-Glu-Ala-Asp) box polypeptide 54 isoform 3*	*Pan troglodytes*	1e-19
TT-N-ST01-0125-W	*Dolichyl-diphosphooligosaccharide--proteinglycotransferase*	*Branchiostoma belcheri tsingtaunese*	4e-19
TT-N-ST01-0059-W	*elongation factor-1 alpha*	*Penaeus monodon*	7e-62
TT-N-ST01-0099-W		*Armadillidium vulgare*	2e-95
TT-N-ST01-0039-W	*Eukaryotic translation initiation factor 2 subunit 2*	*Bombyx mori*	7e-36
TT-N-ST01-0182-W	*GTP-binding protein*	*Bombyx mori*	5e-49
TT-N-ST01-0010-W	*Malate dehydrogenase 1, NAD (soluble), isoform CRA_d*	*Homo sapiens*	4e-43
TT-N-ST01-0019-W	*Meiotic recombination protein DMC1/LIM15 homolog isoform 1*	*Canis familiaris*	1e-24

Table 3.4 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-ST01-0069-W	*Myelodysplasia/myeloid leukemia factor CG8295-PD, isoform D*	*Drosophila melanogaster*	1e-33
TT-N-ST01-0034-W	*Myosin*	*Dictyostelium discoideum AX4*	6e-23
TT-N-ST01-0093-W	*NTF2-related export protein (p15)*	*Tribolium castaneum*	4e-19
TT-N-ST01-0088-W	*Oncoprotein nm23*	*Ictalurus punctatus*	6e-34
TT-N-ST01-0108-W	*Proteasome (prosome, macropain) 26S subunit, ATPase, 5, isoform CRA_a*	*Homo sapiens*	3e-12
TT-N-ST01-0161-W	*Proteasome 26S non-ATPase subunit 12*	*Tribolium castaneum*	4e-18
TT-N-ST01-0074-W	*Proteasome subunit alpha type 2 (Proteasome component C3) (Macropain subunit C3) (Multicatalytic endopeptidase complex subunit C3)*	*Strongylocentrotus purpuratus*	4e-40
TT-N-ST01-0056-W	*Proteasome subunit, alpha type, 5*	*Apis mellifera*	4e-23
TT-N-ST01-0143-W	*Ras-related nuclear protein*	*Marsupenaeus japonicus*	4e-50
TT-N-ST01-0035-W TT-N-ST01-0058-W	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	1e-113 7e-17
TT-N-ST01-0178-W	*Sensitized chromosome inheritance modifier 19 CG9241-PA*	*Drosophila melanogaster*	3e-15
TT-N-ST01-0103-W	*Signal peptidase complex subunit 2 homolog*	*Tribolium castaneum*	3e-28
TT-N-ST01-0138-W	*signal sequence receptor*	*Bombyx mori*	9e-14
TT-N-ST01-0162-W	*Transmembrane protein*	*Pan troglodytes*	2e-64

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.5 Known transcripts found in the reverse SSH library of testes of *P. monodon*

Clone No.	Transcripts	Closest Species	E-value
TT-N-ST02-0155-LF	*ABC transporter ATP-binding protein*	*Flavobacteriales bacterium HTCC2170*	5e-73
TT-N-ST02-0029-LF TT-N-ST02-0156-LF TT-N-ST02-0176-LF	*Alcohol dehydrogenase*	*Bombyx mori*	3e-35 2e-90 2e-90
TT-N-ST02-0024-LF TT-N-ST02-0188-LF	*Allergen Pen m 2*	*Penaeus monodon*	2e-21 5e-22
TT-N-ST02-0097-LF	*Calcitonin gene-related peptide-receptor component protein isoform a*	*Homo sapiens*	8e-21
TT-N-ST02-0001-LF	*Calcium-dependent chloride channel-1*	*Homo sapiens*	4e-11
TT-N-ST02-0010-LF	*Cathepsin B*	*Hippoglossus hippoglossus*	1e-26
TT-N-ST02-0063-LF TT-N-ST02-0106-LF TT-N-ST02-0136-LF	*Cement precursor protein 3B variant 2*	*Phragmatopoma californica*	3e-08 2e-11 3e-07
TT-N-ST02-0038-LF	*Centromere/kinetochore protein zw10 homolog*	*Apis mellifera*	1e-20
TT-N-ST02-0141-LF	*Cytochrome b*	*Penaeus monodon*	3e-80
TT-N-ST02-0007-LF TT-N-ST02-0020-LF TT-N-ST02-0117-LF	*Cytochrome c oxidase subunit I*	*Fenneropenaeus chinensis*	5e-68 1e-66 7e-68
TT-N-ST02-0131-LF	*Drosophila melanogaster eEF1delta*	*Drosophila yakuba*	8e-16
TT-N-ST02-0004-LF TT-N-ST02-0033-LF	*Elongation factor-1 alpha*	*Armadillidium vulgare*	1e-113 1e-113
TT-N-ST02-0008-LF TT-N-ST02-0133-LF	*Eukaryotic translation initiation factor 3 subunit 4*	*Danio rerio*	1e-55 1e-55
TT-N-ST02-0092-LF	*F-box only protein 22*	*Gallus gallus*	6e-08
TT-N-ST02-0142-LF	*Ferric reductase-like protein*	*Aedes aegypti*	1e-28
TT-N-ST02-0138-LF TT-N-ST02-0166-LF	*Gelsolin, cytoplasmic (Actin-depolymerizing factor) (ADF)*	*Homarus americanus*	3e-05 3e-05
TT-N-ST02-0075-LF	*GTP binding protein*	*Bombyx mori*	2e-70
TT-N-ST02-0054-LF	*Heat shock protein gp96*	*Strongylocentrotus purpuratus*	1e-21
TT-N-ST02-0087-LF	*Helicase, lymphoid-specific isoform 2*	*Danio rerio*	1e-43
TT-N-ST02-0009-LF	*Innexin 2*	*Penaeus monodon*	1e-62
TT-N-ST02-0177-LF	*Intracellular fatty acid binding protein*	*Penaeus monodon*	1e-156
TT-N-ST02-0071-LF	*Karyopherin (importin) alpha 4*	*Rattus norvegicus*	8e-08

Table 3.5 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-ST02-0193-LF	*Kinesin heavy chain*	*Loligo pealei*	1e-20
TT-N-ST02-0053-LF	*Mcm3-prov protein (minichromosome maintenance protein 3)*	*Xenopus laevis*	3e-09
TT-N-ST02-0039-LF	*Myeloid leukemia factor 2 (Myelodysplasia-myeloid leukemia factor 2)*	*Danio rerio*	2e-06
TT-N-ST02-0078-LF TT-N-ST02-0108-LF	*Niemann-Pick disease type C2*	*Oreochromis mossambicus*	5e-06 5e-06
TT-N-ST02-0013-LF	*Nop56 CG13849-PA, isoform A (nucleolar KKE/D repeat protein; DmNOP56)*	*Drosophila melanogaster*	2e-49
TT-N-ST02-0017-LF	*Nucleolin*	*Xenopus laevis*	3e-04
TT-N-ST02-0047-LF	*Peptidylprolyl isomerase A*	*Ictalurus punctatus*	8e-15
TT-N-ST02-0022-LF	*Progestin receptor membrane component 1*	*Oryzias latipes*	1e-47
TT-N-ST02-0040-LF	*Proteasome (prosome, macropain) 26S subunit, non-ATPase, 13*	*Tribolium castaneum*	2e-78
TT-N-ST02-0101-LF TT-N-ST02-0130-LF	*Proteasome 26S non-ATPase subunit 12*	*Tribolium castaneum*	1e-71 1e-71
TT-N-ST02-0049-LF	*Proteasome 26S subunit subunit 4 ATPase CG5289-PA*	*Drosophila melanogaster*	1e-89
TT-N-ST02-0165-LF	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	1e-107
TT-N-ST02-0084-LF	*Ribosomal RNA methyltransferase*	*Aedes aegypti*	6e-10
TT-N-ST02-0015-LF	*RNA polymerase 1-1*	*Mus musculus*	2e-25
TT-N-ST02-0028-LF	*Small nuclear ribonucleoprotein D2 polypeptide 16.5kDa, isoform CRA_b*	*Homo sapiens*	2e-34
TT-N-ST02-0170-LF	*Small nuclear ribonucleoprotein E (snRNP-E) (Sm protein E) (Sm-E)*	*Drosophila melanogaster*	2e-14
TT-N-ST02-0065-LF TT-N-ST02-0070-LF	*Small optic lobes CG1391-PB, isoform B (Calpain)*	*Apis mellifera*	7e-81 7e-81
TT-N-ST02-0185-LF	*Tetraspanin 3, isoform CRA_a*	*Homo sapiens*	4e-10
TT-N-ST02-0056-LF	*Tetraspanin 96F CG6120-PA*	*Drosophila melanogaster*	5e-14
TT-N-ST02-0151-LF	*Transcription initiation factor TFIID subunit 12 (Transcription initiation factor TFIID 20/15 kDa subunits)*	*Xenopus laevis*	1e-21
TT-N-ST02-0016-LF	*Variable surface lipoprotein*	*Mycoplasma bovis*	1e-08

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.3 Construction of a conventional heart cDNA library of *P. monodon*

In addition, a conventional cDNA library of heart of *P. monodon* was also constructed. Initially, mRNA was purified from total RNA extracted from heart (Fig 3.9) of juvenile shrimp. Heart mRNA (4.3 µg) were used for a cDNA library construction. First and second strand cDNA were synthesized and manipulated before cDNA was sized-fractionated through column containing Sepharose CL-2B gel filtration medium. Fractions were collected and a 5 µl aliquot of each fraction was examined by polyacrylamide gel electrophoresis. Fractions 4 - 7 of the heart cDNA which were free from adaptors (Fig. 3.10) were selected, pooled and subjected to ligation of the cDNA insert and preparation of a lambda library, respectively.

The primary titer of the heart cDNA library was 1.55 x 10^6 pfu/ml. The percentage of positive clones examined by colony PCR was 93.4% (605/648 clones) and 98.7% of which carrying the inserts greater than 500 bp in size (Fig. 3.11) and subjected to similarity search using Blast *N* and Blast *X*. A total of 413 recombinant clones were sequenced, 283 ESTs (68.5%) corresponded to known sequences (E-value < 1e-04) whereas 130 sequences were regarded as novel (unknown) transcripts (31.5%, E-value > 1e-04). Genes encoding mitochondrial proteins such as *cytochrome b,* several subunits of *cytochrome c oxidase* and *NADH dehydrogenase* were highly redundant. Functionally important gene homologues such as *Myosin light chain 1* (*MLC1*), *Profilin, ERO1-like, Troponin T*, and *Thyroid hormone receptor interactor 12 isoform 7* were also found (Tables 3.6 and 3.7).

Two hundred and eighty-three known ESTs from the heart cDNA library were allocated to 10 functional categories (Table 3.6). Mitochondrial protein (25.2%) and ribosomal protein (16.9%) were abundant in this library. Disregarding hypothetical protein (9.7%) homologues, matched ESTs categorized as members of metabolism (6.5%) predominated followed by those classified as members of miscellaneous function (3.9%), gene expression and protein synthesis (3.2%), defense and homeostasis (1.5%), and internal/external structure (1.2%). A single EST allocated to cell division/DNA synthesis (0.2%) and transport (0.2%) was *high mobility group protein DSP1* (*Dorsal switch protein 1*) which significantly matched that of *Tribolium*

castaneum (4e-56) and *electron-transfer-flavoprotein beta polypeptide* which significantly matched that of *Bombyx mori* (6e-50), respectively.

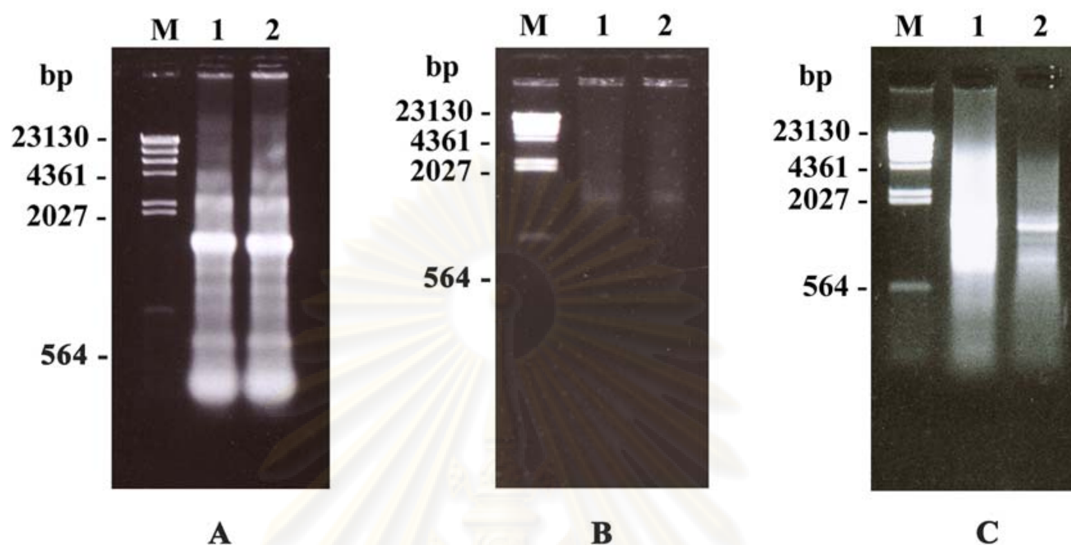


Figure 3.9 Total RNA (A) and mRNA (B) of heart of *P. monodon* used to construct a conventional cDNA library. The first (C, lane 1) and second strand (C, lane 2) cDNA were synthesized from purified mRNA of heart of juvenile *P. monodon*.

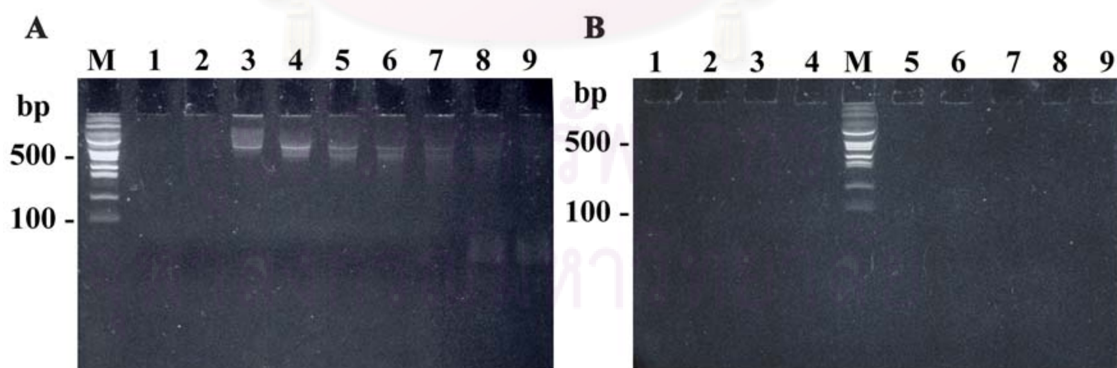


Figure 3.10 Size-fractionation of heart cDNA of *P. monodon*. Lanes 1-9 (A) and 1-4 and 5-9 (B) correspond to fractions 2-10, 11-14 and 15-19, respectively. Lanes M = a 100 bp marker.

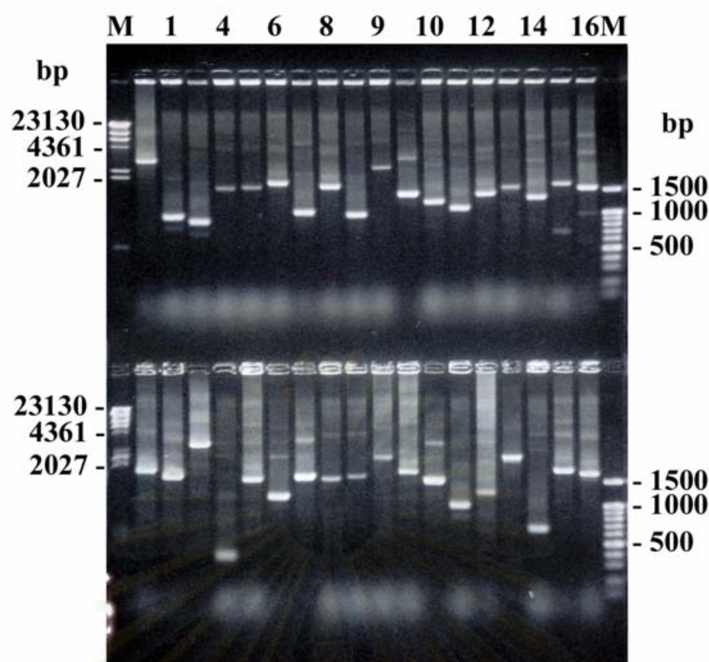


Figure 3.11 Colony PCR for determining sizes of inserts of positive cones clones from a conventional heart cDNA library of *P. monodon*.

Table 3.6 Functional categories of ESTs of the conventional heart cDNA library of *P. monodon* significantly matched previously deposited sequences in GenBank

	Functional category	No. of clones (%)
1	Gene expression, regulation and protein synthesis	13(3.2)
2	Internal / external structure and motility	5(1.2)
3	Metabolism	27(6.5)
4	Defense and homeostasis	6(1.5)
5	Cell division / DNA synthesis	1(0.2)
6	Ribosomal protein	70(16.9)
7	Mitochondrial protein	104(25.2)
8	Transport	1(0.2)
9	Miscellaneous function	16(3.9)
10	Unidentified (hypothetical)-similar to other cDNA/DNA	40(9.7)
11	Unknown	130(31.5)
Total		**413(100)**

Table 3.7 Functionally important transcripts from a conventional heart cDNA library of *P. monodon*

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0012-LF	*Hydroxyproline-rich protein*	*Micrococcus sp. 28*	4e-05
HT-N-S01-0016-LF	*Electron-transfer-flavoprotein beta polypeptide*	*Bombyx mori*	6e-50
HT-N-S01-0023-LF	*Myosin light chain 1*	*Aedes aegypti*	2e-38
HT-N-S01-0032-LF	*Fast tropomyosin isoform*	*Homarus americanus*	3e-44
HT-N-S01-0089-LF	*High mobility group 20A*	*Mus musculus*	1e-43
HT-N-S01-0093-LF	*Calponin*	*Aedes aegypti*	4e-34
HT-N-S01-0119-LF	*ADP/ATP translocase*	*Bombyx mori*	7e-73
HT-N-S01-0128-LF	*Troponin T*	*Libellula pulchella*	2e-12
HT-N-S01-0144-LF	*Serine proteinase inhibitor*	*Pacifastacus leniusculus*	4e-42
HT-N-S01-0154-LF HT-N-S01-0297-LF	*Fructose 1,6-bisphosphate aldolase*	*Oncometopia nigricans*	1e-16 1e-37
HT-N-S01-0158-LF	*Profilin (Chickadee protein)*	*Tribolium castaneum*	7e-38
HT-N-S01-0162-LF	*NADH-ubiquinone oxidoreductase Fe-S protein 7*	*Bombyx mori*	3e-62
HT-N-S01-0165-LF	*C-type lectin*	*Penaeus monodon*	7e-06
HT-N-S01-0178-LF	*Dehydrogenase/3-ketoacyl-Coenzyme A thiolase*	*Danio rerio*	5e-09
HT-N-S01-0185-LF	*Latency associated nuclear antigen*	*Saimiriine herpesvirus 2*	9e-12
HT-N-S01-0198-LF	*GRN protein (granulin)*	*Xenopus tropicalis*	1e-33
HT-N-S01-0200-LF	*Troponin T-1*	Drosophila melanogaster	1e-12
HT-N-S01-0204-LF	*Tetratricopeptide repeat domain 35*	*Danio rerio*	2e-47
HT-N-S01-0218-LF	*Peroxisomal 3,2-trans-enoyl-CoA isomerase (Dodecenoyl-CoA delta-isomerase) isoform 1*	*Strongylocentrotus purpuratus*	5e-53
HT-N-S01-0219-LF	*Plasminogen*	*Sus scrofa*	2e-13
HT-N-S01-0221-LF	*High mobility group protein DSP1 (Dorsal switch protein 1)*	*Tribolium castaneum*	4e-56
HT-N-S01-0222-LF	*Papilin*	*Aedes aegypti*	3e-36
HT-N-S01-0231-LF	*gcdh protein (Glutaryl-Coenzyme A dehydrogenase)*	*Danio rerio*	4e-70
HT-N-S01-0241-LF	*Neuroparsin A precursor*	*Locusta migratoria*	9e-11
HT-N-S01-0243-LF	*Myosin 1 light chain*	*Lonomia obliqua*	1e-48
HT-N-S01-0244-LF	*ERO1-like*	*Gallus gallus*	2e-41
HT-N-S01-0247-LF	*ATP synthase*	*Penaeus monodon*	6e-74
HT-N-S01-0289-LF	*Neuroparsin A precursor*	*Locusta migratoria*	5e-06
HT-N-S01-0291-LF	*Glutathione S-transferase*	*Anopheles gambiae*	7e-35

Table 3.7 (cont.)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0294-LF	*Trifunctional enzyme beta subunit (tp-beta)*	*Aedes aegypti*	3e-48
HT-N-S01-0296-LF	*ATP synthase*	*Penaeus monodon*	9e-93
HT-N-S01-0306-LF	*Serpin 3*	*Plutella xylostella*	9e-09
HT-N-S01-0317-LF	*Glyceraldehyde-3-phosphate dehydrogenase*	*Procambarus clarkii*	6e-76
HT-N-S01-0350-LF	*ATP lipid-binding protein like protein*	*Marsupenaeus japonicus*	1e-30
HT-N-S01-0358-LF	*Thyroid hormone receptor interactor 12 isoform 7*	*Canis familiaris*	7e-35
HT-N-S01-0382-LF	*Lipoamide dehydrogenase*	*Sus scrofa*	3e-51
HT-N-S01-0384-LF	*Supervillin, isoform CRA_a*	*Homo sapiens*	8e-16
HT-N-S01-0397-LF	*Phosphoglycerate kinase*	*Aedes aegypti*	1e-14
HT-N-S01-0407-LF	*Fasciclin-like protein*	*Aplysia californica*	3e-27
HT-N-S01-0417-LF	*Reticulon 4-L2*	*Takifugu rubripes*	6e-42
HT-N-S01-0419-LF	*Muscle lim protein*	*Aedes aegypti*	2e-34
HT-N-S01-0421-LF	*Tyrosine-protein phosphatase non-receptor type 13 (Protein-tyrosine phosphatase 1E) (PTP-E1) (hPTPE1) (PTP-BAS)*	*Tribolium castaneum*	8e-50
HT-N-S01-0427-LF	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	1e-90
HT-N-S01-0433-LF	*Macrophage migration inhibitory factor*	*Bombyx mori*	5e-25
HT-N-S01-0441-LF	*Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain*	*Danio rerio*	4e-73
HT-N-S01-0450-LF	*Cathepsin D*	*Aedes aegypti*	7e-86
HT-N-S01-0452-LF	*Arginine kinase*	*Fenneropenaeus chinensis*	2e-40

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.4 Isolation and characterization of the full length cDNA of functionally important gene homologues of *P. monodon*

The full length cDNA of 11 functionally important genes was discovered from testis (10 transcripts) and heart (1 transcript) cDNA libraries (Table 3.8). There were *small ubiquitin-like modifier* (*SUMO*, 1304 bp in length with an ORF of 282 bp corresponding to 93 aa, Fig. 3.12), *peptidyl-prolyl cis-trans isomerase* (*cyclophilin A*, 929 bp; ORF of 495 bp, 164 aa, Fig. 3.13)**,** *translationally controlled tumor protein* (*TCTP,* 730 bp; ORF 507 bp, 168 aa, Fig. 3.14), *small nuclear ribonucleoprotein polypeptide G* (676 bp; ORF 234 bp, 77 aa, Fig. 3.15), *dynactin subunit 5* (1436 bp; ORF 594 bp, 197 aa, Fig. 3.16), *ubiquitin conjugating enzyme 2* (1416 bp; ORF 675 bp, 224 aa, Fig. 3.17), *BUB3 budding uninhibited by benzimidazoles 3 homolog* (*mitotic checkpoint*, 3554 bp; ORF 981 bp, 326 aa, Fig. 3.18)*, cell division control protein 2 homolog* (*Cdc2*, 1696 bp; ORF 900 bp, 299 aa, Fig. 3.19), *thioredoxin 1* (731 bp; ORF 318 bp, 105 aa, Fig. 3.20), *multiprotein bridging factor 1* (715 bp; ORF 450 bp, 149 aa, Fig. 3.21), and *myosin 1 light chain* (*MLC1*, 1414 bp; ORF 465 bp, 154 aa, Fig. 3.22). All full length cDNA except *thioredoxin 1* and *multiprotein bridging factor 1* contained the ORF with 3′UTR and the poly A tail.

Sequence analysis revealed that *SUMO* of *P. monodon* (Fig. 3.12) contained an ubiquitin domain (17th – 88th of the deduced protein, E-value = 6.2e-15). The expected MW and p*I* of this deduced protein were 10.58 kDa and 4.99, respectively.

Cyclophilin A of *P. monodon* (Fig. 3.13) contains a pro-isomerase domain (also called cyclophilin type peptidyl-prolyl cis-trans isomerase/CLD domain, positions 5th - 164th, E-value = 2.2e-116) commonly found in the cyclophillin (also called peptidyl-prolyl isomerase) protein family. The expected MW and p*I* of this gene product were 18.86 kDa and 8.78, respectively.

TCTP (Fig. 3.14) contained a TCTP domain (positions 1th - 165th, E-value = 1.40e-58). Mammalian translationally controlled tumor protein (TCTP) (or P23) is a protein which has been found to be preferentially synthesized in cells during the early growth phase of some types of tumor, but which is also expressed in normal cells. The expected MW and p*I* of this gene product were 19.22 kDa and 4.39, respectively.

```
AAGACCACGCGTTCCTCCAACAAATTCATCAATAATGTCTGATAACACTGACGCCAAGCC 60
                                  M  S  D  N  T  D  A  K  P  9
AGAAGGGGAAGGGAACGAATACATCAAACTTAAAGTTGTAGGACAGGACTCCAATGAGAT 120
 E  G  E  G  N  E  Y  I  K  L  K  V  V  G  Q  D  S  N  E  I  29
CCACTTCCGAGTGAAGATGACCACACAGATGGGCAAGTTAAAGAAGTCATACAGTGAGCG 180
 H  F  R  V  K  M  T  T  Q  M  G  K  L  K  K  S  Y  S  E  R  49
GGTGGGAGTCCCTGTAGCATCGCTGCGTTTCCTCTTTGACGGACGACGCATTAATGACGA 240
 V  G  V  P  V  A  S  L  R  F  L  F  D  G  R  R  I  N  D  E  69
AGAAACGCCCAAAGCTCTGGAAATGGAGAATGATGACGTAATTGAAGTGTACCAGGAGCA 300
 E  T  P  K  A  L  E  M  E  N  D  D  V  I  E  V  Y  Q  E  Q  89
GACCGGCGGCCATTGATGCAACACATTCCCGCGACCATAGGAATAAGACATCGTTAGGTT 360
 T  G  G  H  *                                               93
AAGGAAGTTTATTTTTCGCCACACAGTGTACCTTTATTTTCTGGCTGAGATTTTCGCACA 420
GACCAGGCAATGTGCGCAGACCTTTTTAGATGGAATTTCTGCGAGTCTCGTACAATGTAT 480
AATCACGCAAGAGTCTGAAAAATTATTAAATTTTCTCTTTTTCTTTTTATATATGTATGT 540
TATTTGCCCAAGGATGGTTTCCTAGAGCAAATTGTGCAGCAGAAGTGTGCGACTTCACCA 600
GGCCTATCTCAAGACCAGGCATGAGAGAACTTAAGTTTCTGCATAGCTTTTGAGATTTAG 660
GTGTTAACGCATCTAATATGTTGTTAACCAAAAGAATGAAGATTTCCCTCTCCTCTTTTT 720
ATTTTTTTGAAAGCATTGTAAAAGTCATGAATGTTAGTACCTTTTTTCATTTTATTTTTT 780
ATTTCAAATCCTTATAATTCCAAGTAGCAGGAAGGAAAGACATCAAATTTAAATTCTCGA 840
CAAGGGTTATGTTTAACTATTAGTCTGTACCGTTCGAATGTTCTTGTAATAGTCTGCAGC 900
GTGACATCTTTATTGCTGTGAACCAAAATAGACTCTGTGAGTTCTATGTGTTATTCAATA 960
CATTCTAAATATAACCCAATTTCGTCTTTTTGTTTTATATTATTTGTGGATAGCAAAAAA 1020
TTTTAATATTTCATACATAATGTGAGAAGCAGTGAAGTAAAATGATTGTGTAAAAAAAAA 1080
AAATCAAGATGAAGCTTCACTTTCTTCTTGGTATTCTATTGTCGACGGTAATCATGTTGG 1140
CGGAAGCGAAAAATGGGCTGAGCAGTCGGACAGGCAGTAAAGAACACATGGACATCTAAA 1200
CCAGTGTCCTGGAAATCCTCTTTTCCAGGCAGGATCAACATTTCTTTTTAGTCTCAATAA 1260
AAGAAAATTGTGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA                1305
```

Figure 3.12 The full length cDNA sequences of *SUMO* (1305 bp in length with an ORF of 282 bp corresponding to a polypeptide of 93 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An ubiquitin domain (17th–88th of the deduced protein was highlighted.

Small nuclear ribonucleoprotein polypeptide G (Fig. 3.15) contained a sm domain (positions 7th - 72th, E-value = 1.74e-20). Small nuclear ribonucleoprotein particles (snRNPs or snRNP Sm proteins) involved in pre-mRNA splicing. The expected MW and p*I* of this gene product were 8.65 kDa and 9.98, respectively.

Ubiquitin conjugating enzyme 2 (Fig. 3.17) contained an UBCc domain (positions 81th - 224th, E-value = 1.85e-72). The expected MW and p*I* of this gene product were 23.89 kDa and 7.64, respectively. Ubiquitin proteosome pathway is a major means in eukaryotic cells for targeted protein proteolysis. The system generally includes three classes of ubiquitin enzymes: ubiquitin-activating enzymes (E1s), ubiquitin-conjugating enzymes (E2s or UBC) and ubiquitin protein ligases (E3s).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
GGCACGAGGCAGACCTACGCCAACTTAGCCACCATGGGCAACCCCAAAGTCTTTTTCGAC 60
                                    M  G  N  P  K  V  F  F  D    9
ATTACCGCTGACAACCAGCCCGTTGGCAGGATCGTCATGGAGCTCCGCGCCGACGTGGTC 120
I  T  A  D  N  Q  P  V  G  R  I  V  M  E  L  R  A  D  V  V    29
CCCAAGACCGCCGAGAACTTCCGGTCGCTGTGCACGGGCGAGAAGGGCTTCGGCTACAAG 180
P  K  T  A  E  N  F  R  S  L  C  T  G  E  K  G  F  G  Y  K    49
GGCTCCTGCTTCCACCGCGTGATCCCCAACTTCATGTGTCAGGGAGGCGACTTCACCGCC 240
G  S  C  F  H  R  V  I  P  N  F  M  C  Q  G  G  D  F  T  A    69
GGCAACGGCACGGGCGGCAAGTCCATCTACGGCAACAAATTCGAGGACGAGAACTTCGCA 300
G  N  G  T  G  G  K  S  I  Y  G  N  K  F  E  D  E  N  F  A    89
CTGAAGCACACCGGCCCCGGCACCCTGTCCATGGCCAACGCCGGCCCCAACACCAACGGG 360
L  K  H  T  G  P  G  T  L  S  M  A  N  A  G  P  N  T  N  G    109
TCGCAATTCTTCATCTGCACCGTCAAAACCCCCTGGCTGGACAACAAGCACGTGGTCTTC 420
S  Q  F  F  I  C  T  V  K  T  P  W  L  D  N  K  H  V  V  F    129
GGCTCCGTGGTGGAGGGCATGGACATCGTGCGCCAGGTCGAGGGCTTCGGCACGCCCAAC 480
G  S  V  V  E  G  M  D  I  V  R  Q  V  E  G  F  G  T  P  N    149
GGCTCTTGCAAGCGGAAAGTGATGATCGCCAACTGCGGCCAGCTGTAAAGTCTCAGAACA 540
G  S  C  K  R  K  V  M  I  A  N  C  G  Q  L  *               164
TTCCGCCTTAGCCGCCCACACCTTTTTTTTCTGATGTAATTGAGGATCCAGGATATAATC 600
TTTGCTGTATTGGCACTTCAGTGTTAAATTTCGGCTTGAAAAAAAGTTAAATGCTATATA 660
ACGTAAAGGTGGTGAAACAAGATAGGTGTTCTTCCATTTTTTTTTGTTTTATTAGTTTCA 720
TAAGTGGTCATGTTCTGGAAATGTTGACGCATTATGCTGATATTCAGCATTTCGTCTTCT 780
CACTTCATCAATAAATCACCAACAACCTTTCTTCTTTCTCAACCTTTAAATCTTTGAGAC 840
AAGCTGTAATAAAGACCAGAAACAGTAACAGGTTCAGTAAAGGTCAAATTCAGTAAATCC 900
AAGCCAAAAAAAAAAAAAAAAAAAAAAAA                                929
```

Figure 3.13 The full length cDNA sequences of *cyclophilin A* (929 bp in length with an ORF of 495 bp corresponding to a polypeptide of 164 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A pro-isomerase domain (positions 5th - 164th of the deduced protein is highlighted.

```
CCACGGTGAACGAGACCAACCTTCTCCACCAGTCGAGAATTTAGCGACGATCATCTAGCC 60
GCCATGAAGGTCTTCAAGGATATGCTGACCGGTGATGAGATGTTCACTGACACCTATAAG 120
   M  K  V  F  K  D  M  L  T  G  D  E  M  F  T  D  T  Y  K    19
TATGAGGAGGTGGATGATGCCTTCTACATGGTAATTGGAAAAAATATTACTGTTACTGAA 180
Y  E  E  V  D  D  A  F  Y  M  V  I  G  K  N  I  T  V  T  E    39
GATAACATTGAGCTGGAGGGAGCCAATCCATCAGCTGAAGAGGCAGATGAAGGCACTGAC 240
D  N  I  E  L  E  G  A  N  P  S  A  E  E  A  D  E  G  T  D    59
ACTACTAGTCAGTCTGGTGTTGATGTAGTTATATATATGCGTCTGCAGGAAACCGGCTTC 300
T  T  S  Q  S  G  V  D  V  V  I  Y  M  R  L  Q  E  T  G  F    79
CAAGTCAAGAAGGATTATCTTGCATACATGAAAGAATACCTAAAGAATGTAAAGGCAAAG 360
Q  V  K  K  D  Y  L  A  Y  M  K  E  Y  L  K  N  V  K  A  K    99
TTGGAAGGCACGCCTGAAGCTTCAAAGTTAACATCTATCCAGAAGCCTCTGACAGACCTT 420
L  E  G  T  P  E  A  S  K  L  T  S  I  Q  K  P  L  T  D  L    119
TTGAAGAAGTTCAAGGACTTGCAATTCTTCACTGGAGAATCAATGGACCCTGATGGCATG 480
L  K  K  F  K  D  L  Q  F  F  T  G  E  S  M  D  P  D  G  M    139
GTTGTTCTCATGGATTACAAAGACATTGATGGAGAAGAGCGGCCAGTCTTGTACTTCCCA 540
V  V  L  M  D  Y  K  D  I  D  G  E  E  R  P  V  L  Y  F  P    159
AAATACGGTCTAACAGAGGAGAAGCTATAAACGTTATTTATTTCTGAGTTATAATGCAGC 600
K  Y  G  L  T  E  E  K  L  *                                 168
CCTCGTCATCTGGACTCCCAGGGTCATGAAGCAGATTGTTCATGGTCCTTTTTAATTTAA 660
TTTTAATATACTCAGATAAATTTAGTCATCCTCTCCAAAAAAAAAAAAAAAAAAAAAAAA 720
AAAAAAAAAA                                                   730
```

Figure 3.14 The full length cDNA sequences of *translationally controlled tumor protein* (730 bp in length with an ORF of 507 bp corresponding to a polypeptide of 168 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A TCTP domain (positions 1st - 165th of the deduced protein) is highlighted.

Dynactin subunit 5 has the expected MW and p*I* of 21.54 kDa and 8.84 (Fig. 3.16) whereas those of *BUB3 budding uninhibited by benzimidazoles 3 homolog* (Fig. 3.18) were 37.01 kDa and 7.48, respectively. Functionally important domains were not found in both transcripts.

```
AATCACAATTTCTTTTATTGAGACTAAAAAGAAATGTTGATCCTGCCTGGAAAAGAGGAT 60
TTCCAGGACACTGGTTTAGATGTCCATGTGTTCTTCATTGCCTCTCCGATTGGTCAGCCC 120
ATTTTTCGCTTCCGCCAACATGATTACCGTCGACAATAGAATACCGAGAAGAAAGTGAAG 180
CTTCATCTTGATGTTTTGACTTCGAGAAACAACTGTTTTCGCCCCAAATAAGCTTTCAGG 240
ATGAGCAAGGCACATCCACCAGAGTTGAAGAAGTACATGGACAAGCGCGTCATGACCAAG 300
M  S  K  A  H  P  P  E  L  K  K  Y  M  D  K  R  V  M  T  K
CTGAATGGTGGACGCGTGGTCGAGGGAACACTAAGAGGCTTTGACCCCTTCATGAACCTT 360
L  N  G  G  R  V  V  E  G  T  L  R  G  F  D  P  F  M  N  L
GTGGTGGATGATGGGGTGGAAGTGCGCAGGAGTGGAGATCGTGTCAGGGTTGGCTTTGTG 420
V  V  D  D  G  V  E  V  R  R  S  G  D  R  V  R  V  G  F  V
GTCATCCGAGGCAGTAGCATCATCATGCTTGAAGCCCTGGATCGGATATCGTAGTCTTGT 480
V  I  R  G  S  S  I  I  M  L  E  A  L  D  R  I  S  *
ACCCAAATAATATTAAACTTTAAGTTAGGATCTTAGAAAGATGTAATTCAGCTTAAGTTT 540
TGATTAGAAGCAGCCATTCTTATATGTGAAGAATTATTTAGTTTTTCCGAAGAATTTCCA 600
GTTTTTTCCTGGAATTATTATGTGAATGTGTTTTCAGTAAAATAATGATTTTGTAAAAAA 660
AAAAAAAAAAAAAAAA                                             676
```

Figure 3.15 The full length cDNA sequences of *small nuclear ribonucleoprotein polypeptide G* (676 bp in length with an ORF of 234 bp corresponding to a polypeptide of 77 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A sm domain (positions 7th-72nd of the deduced protein) was highlighted.

```
AATCACAATTTCTTTTATTGAGACTAAAAAGAAATGTTGATCCTGCCTGGAAAAGAGGAT 60
TTCCAGGACACTGGTTTAGATGTCCATGTGTTCTTTACTGCCTGTCCGACTGCTCAGCCC 120
AAGGTCCTGTGACCAAATCTACAGGTACACAAGAAATGACGGACAGGCTGACGAGTCGGG 160
ATTTTTCGCTTCCGCCAACATGATTACCGTCGACAATAGAATACCGAGAAGAAAGTGAAG 180
CTTCATCTTGATTCAGCTGCTGATCCCCGAACTCTGCCGTGCGGACCCGAGGGAGAATTC 240
CCCACGTAAAGGCTTCGTCATGGAATTGCAGGATTCCTTTTATCCCAGAGCTCAGTATAT 300
                   M  E  L  Q  D  S  F  Y  P  R  A  Q  Y  I  14
TGAGACTGCCACAGGAAACCGTGTCAGCCGTGCAAGTGTGTTATGTGGATCTCAGAACAT 360
 E  T  A  T  G  N  R  V  S  R  A  S  V  L  C  G  S  Q  N  I  34
TGTTTTAAGTGGTAAAGTTATCGTGCTTAGTGGTGTGATTATCAGAGGAGACCTTGCCAA 420
 V  L  S  G  K  V  I  V  L  S  G  V  I  I  R  G  D  L  A  N  54
TGTCAGAGTAGGACGCCACTGTGTGATATCATCCAAAGCGGTTATCAGACCTCCATTTAA 480
 V  R  V  G  R  H  C  V  I  S  S  K  A  V  I  R  P  P  F  K  74
AAAATTCAGTAAGGGAGTAGCATTTTTCCCACTTCACATTGGGGATCATGTTTATATTGG 540
 K  F  S  K  G  V  A  F  F  P  L  H  I  G  D  H  V  Y  I  G  94
TGAGGGTTCTGTTGTAAATGCTGCAGTCNCACTGGTTCATATGTCTACATTGGACAAGAA 600
 E  G  S  V  V  N  A  A  V  X  L  V  H  M  S  T  L  D  K  N  114
CTGTGTTATTGGACGCCGTTGCGTGCTCAAAGACTGCTGCATGATTGCGGACAACACAGT 660
 C  V  I  G  R  R  C  V  L  K  D  C  C  M  I  A  D  N  T  V  134
CTTGCCTCCCGAGACTGTTGTTCCACCATTTGCAGTCTACAATGGTTCACCTGCCAAGCA 720
 L  P  P  E  T  V  V  P  P  F  A  V  Y  N  G  S  P  A  K  H  154
CACAGGAGACTTGCCAGAGGCTACCCAAGATCTTATGACTGATTACACTAAATCATGTTA 780
 T  G  D  L  P  E  A  T  Q  D  L  M  T  D  Y  T  K  S  C  Y  174
TCACCATTTCATCAGAGTCAAAGATATGCCACCACAGGCAAAGGAAGGAACTACAAAGCT 840
 H  H  F  I  R  V  K  D  M  P  P  Q  A  K  E  G  T  T  K  L  194
```

```
TATTGAAATTTAGGTACAGAAAGCTGTTTTTTTGCTTCCCTGTTTCCTCAAACCAAAGGT 900
 I  E  I  *                                                  197
ATATTATCCAGCCTGAATCATGGATCTTCAGTTGAACTCAATGAAGAAGAGAAAATATCT 960
GCAGTATACTATTTATTTCAAATCAGTTACTTTAGCATTTAGCAAAGGTATGAATTAAAT 1020
TATGGTATGTTTACATTTAGCAATGGACTATGATGTCATTTGAAATTTGCCTAAATATAT 1080
TACTAGGAGATTTCTGTGTCTACTACTAATAGGAAATGCAGTGTATGTCATTCCAGTATG 1140
ATACTGACAACATTGAGGATATTATTATTTAAGAAGCTTTAAATATTCGTTTTAGTTTTT 1200
TTTATATATAGAACAATAAAAAAGAAAAAGATAAAATTTTCATTATCACCAGGTTATCTT 1260
CTGCAGCCGAATATTATGGCTAGGTAATAGTAATTGATTTAAAGGCAGTTATAAAATGAA 1320
TGAATATATAAGACTTTCTTAAGAATGACTTAATTTTCAAGCTGAATAATGAAGTGATAT 1380
ATATCATTATGTGCTTCTATATGTTAAAATAAAGTAAATAATGGCAAAAAAAAAAA     1436
```

Figure 3.16 The full length cDNA sequences of *dynactin subunit 5* (1436 bp in length with an ORF of 594 bp corresponding to a polypeptide of 197 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined.

```
AGAAACCTCGCATGGAGAAAGAGGAGCTCGAGAGGCAACTCAAAGACAACTTTAACCTTC 60
TGCTTTGTAGCTTTATAACAAAGAAAATTACTTTGGGGTGGGGGGAATAATCGAAGGAAC 120
TTCACCCTTCATTTCTATCTCACAAACGGAACACATTCAGAGGTAGATCATGGTGAGTTA 180
TCAGGAATATGGTCATGTGCATTGATTGATGTCTGCCTCGCCTGCTGGTGCCACCGGGTC 240
                            M  S  A  S  P  A  G  A  T  G  S  11
TGACTCTGTCCCCGTGTCTCCCTCAAACCCCCCCACCACCACTAGTCCCCCCTCGTCAGC 300
 D  S  V  P  V  S  P  S  N  P  P  T  T  T  S  P  P  S  S  A  31
GACCTTGACCGCAGCGTCTGCTGGGTCAGCCCTCTCTCCCTCAAGCACTGCCACCAGTTC 360
 T  L  T  A  A  S  A  G  S  A  L  S  P  S  S  T  A  T  S  S  51
CCAGCAGTCGGCCCCGGTAGAGCCCCCCGTTGTGCGTGAGGTCAGGCCACACAACCCCAA 420
 Q  Q  S  A  P  V  E  P  P  V  V  R  E  V  R  P  H  N  P  K  71
AATGTCAAAGGCACTCAGCACCAGTGCTAAGAGGATACAAAAGGAACTCGCAGAAATAAC 480
 M  S  K  A  L  S  T  S  A  K  R  I  Q  K  E  L  A  E  I  T  91
ACTAGACCCCCCACCCAACTGCAGCGCTGGGCCTAAGGGAGACAATCTGTATGAATGGGT 540
 L  D  P  P  P  N  C  S  A  G  P  K  G  D  N  L  Y  E  W  V  111
GTCGACTATCCTGGGTCCACCTGGGTCCGTGTATGAGGGTGGGGTCTTCTTTCTAGATAT 600
 S  T  I  L  G  P  P  G  S  V  Y  E  G  G  V  F  F  L  D  I  131
CCATTTCTCGGCAGAATACCCATTCAAACCACCTAAGGTCACGTTTCGTACACGAATTTA 660
 H  F  S  A  E  Y  P  F  K  P  P  K  V  T  F  R  T  R  I  Y  151
TCACTGCAACATCAACTCCCAAGGAGTGATATGTTTAGATATTCTTAAGGATAACTGGTC 720
 H  C  N  I  N  S  Q  G  V  I  C  L  D  I  L  K  D  N  W  S  171
TCCAGCCCTCACTATCTCCAAGGTTCTGCTGTCCATCTGCTCTCTTCTTACAGATTGCAA 780
 P  A  L  T  I  S  K  V  L  L  S  I  C  S  L  L  T  D  C  N  191
TCCAGCTGATCCCCTCGTAGGAAGCATTGCCACACAGTACCTCCAAAACAGGGAAGAACA 840
 P  A  D  P  L  V  G  S  I  A  T  Q  Y  L  Q  N  R  E  E  H  211
TGATAGAATTGCACGGCTCTGGACCAAGCGTTATGCTACGTGATCACTTCCCGGTAGCTG 900
 D  R  I  A  R  L  W  T  K  R  Y  A  T  *                    224
TGAACAGGGTATTAGGCATCTCTTTCACCATGCCAATTTGGGTTTAATCATTTTATATAA 960
GTAATTTCCACTCGGCCCCTTGATTTCTAGAGTATCTCAGCTTCTGAAAGAAAAAAAAAG 1020
ATATAAAAATAATAAAACCAAACAAATTAAAAAATCTAACAAAAAAAAATTAAATATAAA 1080
AAAAAAAAAAAAAAACCTCCAAGAAAAGGAATGCACCATGGATAAGCGAATATTAATCCT 1140
GGTTCTTTTAGTGTCTTCGCTGTGCATAATGGCTGATGCTCATGAGAAGAGGAGTCCGCT 1200
AGCCAGAAGTGGCCCAAGCCGTCCGCCAGGGAGTAAAGGAAAGCGCTCAGTTTCCTACGA 1260
AGGCCTGGAGGAGGAATGCCTTTGCCAAAGGTGATCACGGCTCGGCGTGAGACTCGTCGG 1320
CTCAGCGCTCGGTCGTTCGGTTTCGGCGAGTAATTGGCGATGTCACGATTTGGGAATTAA 1380
AGGTTGACATCAATAAAAAAAAAAAAAAAAAAAAAA                         1416
```

Figure 3.17 The fulll length cDNA sequences of *ubiquitin conjugating enzyme 2* (1416 bp in length with an ORF of 675 bp corresponding to a polypeptide of 224 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An UBCc domain (positions 81th - 224thof the deduced protein) is highlighted.

TTTGTTTTCTTCCTTTTCCTTCGTGTGTTTTCGAGTCTGAAGTCGGCAAA**ATG**AGTGAAT 60

M S E S 4

CTCGCATGGAATTTCGGTTGAAGAATACTCCGAGTGACTGTATTCAAAGTGTCAAGTTTG 120

R M E F R L K N T P S D C I Q S V K F G 24

GGCCTTCATCTTCACAGTTCCTTCTAGTAGCATCGTGGGACAAAAGTGTTCGCCTTTATG 180

P S S S Q F L L V A S W D K S V R L Y D 44

ATGTTGTCAATAATAACATGCGGTTACAGTATCAGCATACAGGCCCGGTTTTGGATTGCT 240

V V N N N M R L Q Y Q H T G P V L D C C 64

GCTTCCAGGATGCTGTCCATGCATACAGTGGAGGCTTAGATGGTCAGCTCAAGACCTTTG 300

F Q D A V H A Y S G G L D G Q L K T F D 84

ATCTCAACACAAACACAGAATCTGTGGTTGGCTCTCATGATGCTCCAATCAGGTGTGTGG 360

L N T N T E S V V G S H D A P I R C V E 104

AATTTTGCCCAGAAGTAAATGTTGTGATCACAGGAGCTTGGGATTCCAACATCAAACTCT 420

F C P E V N V V I T G A W D S N I K L W 124

GGGATCCTCGTGGACCACGGGAAGCTGGTACTTTCCAACAGCCAAATAAGGTGTACACCA 480

D P R G P R E A G T F Q Q P N K V Y T M 144

TGGGCCTTGGTGGAGAAAAGTTGGTAGTGGGGACATCCAATAGAAAAGTGATGGTTTGGG 540

G L G G E K L V V G T S N R K V M V W D 164

ATCTGAGGAACATGGGCTTTGCTCAACAGCGCCGAGAATCTTCTCTCAAATACCAGACTC 600

L R N M G F A Q Q R R E S S L K Y Q T R 184

GCTGCATTCAGTGCTTCCCCAACAAACAGGGTTATGTTGTGTCCAGTATTGAGGGTCGTG 660

C I Q C F P N K Q G Y V V S S I E G R V 204

TGGCTGTTGAGTACCTTGACCCGAGCCCGGAAGTCCAGAAGAAGAAGTATGCCTTCAAGT 720

A V E Y L D P S P E V Q K K K Y A F K C 224

GCCACAGACTTAAAGAGGATGGGATTGAGAAAATTTTCCCTGTTAATGCCATAAGTTTCC 780

H R L K E D G I E K I F P V N A I S F H 244

ACAATGGTTACAATACCTTTGCAACAGGAGGTTCTGATGGGTATGTCAATATATGGGACG 840

N G Y N T F A T G G S D G Y V N I W D G 264

GCTTCAACAAGAAGCGCCTGTGCCAGTTCCATCGTTATCCAACTTCCATATCCTCCCTAT 900

F N K K R L C Q F H R Y P T S I S S L C 284

GCTTCAGCAATGATGGTAACACACTAGCAATTGCCTGCTCCTATATGTATGAACAAGAGG 960

F S N D G N T L A I A C S Y M Y E Q E E 304

AAATTGACCCCATGCCAGAGGATTGCATCTTCATCCGTCGTGTGACAGACCAAGAGACGA 1020

I D P M P E D C I F I R R V T D Q E T K 324

AGCCAAAA**TAA**GGAGCCAGAGCAGACAAAGAGCTTGCATCTTTTACAGTAATCCCTCTGT 1080

P K * 326

CACGATGTAGCAAAGAGGGCTGTCCATTACACGGGTCCAATGTGTGTCGGTGGGCTCTGG 1140

GACGTTAACACTCATAGTACCTCAATTTTGTATATGGTGGTATGTATGTATGCCATTCTA 1200

TATGAAGTGTTAGTGAATTTTAAAGTTATGATTAGACCAAGTATTTGACCAGCACTTATT 1260

TAGTGTAGTCAGATAAAAGGTGCTTCTTTCATATCTTTTAATAGTTTTCTATGTATTTTA 1320

TCAGTAGCTTGGCGATTCAGGCCTTAATTACTTTTAAGCAAATGGGTAAATATGCTTAAT 1380

AACCTCCAAATGTCTTTACCATTATTTTCAAAGTTCATTTTCAATGTGGGAACTTGTGAA 1440

TCCTGTATTCATATTCATGCAGTTTCCAGGAAGATCTGAAGAAGAATACCAGTTAATGTT 1500

TAGAAGATACTGAATGAAGTCGGCAGAATTTTTGTGAATGCCTTTGTTGGTTTTATTTAT 1560

CTCCATGTGGATAGTTCTTTAGTGTTGAGAGGTACAGCCCAGTTTACAAGACTTTGGCAG 1620

TGTAGATAAATGATTATTGAAATCTTTTTATTAATGTATTTTTCAATTGCAACAGTACAA 1680

AAGTGTCAGGGGTTATTTCACATATTTACCACTGCTGCATATGCCAAAGTATGCCTAGGG 1740

AGAATTACAATGTCAGCAATAACTACATAGAATGGTTCCATCCTCCTTAGAACTTGATAG 1800

CATTTGGAATAAGGCCCACCGACCTCTGCCTTCTCTTTCATGCCTTTTTTTTTTTCTTCT 1860

TCTTCTTCTTCTTTTGTCTGATTCTTTGTTTCTTCATCATTATGATATGAACAGTGTAAG 1920

TGACAATTCTTTAACTATCAAAGAAAAGAAAAGTGTATTCAGAATTTCTTTAGCCAGCTT 1980

GTGGTATAAAACATTTTTATCTTTTGTTTTAGTTTTTATTGTCCATTTCATTCATTTATC 2040

TTTTGATAAATTCTTTGAGGTTGCCTCCTTCACAAAGCATAAAAAAATTACAATAATAAAT 2100

GCAATTGAAGCTTTTGTTTCTTCAGTTATCATTATCAATGTCATTGTTGTTGTTATTATT 2160

ATTATCATTATTATTATTATTATCATTATTATTATTATTATTATTATTATTATTATCACT 2220

ATTATTATTATCATTATTATTATTGTTATTATTATTATTATCATTATCATCTTCATTATC 2280

ATTATTATTATTACTACTATTTTTTTGTCCCCTTGCCATTTCTCTGCTTGGCTCTTTCTA 2340

CTGCTTGAAACCTTCATAAGAAAAGCAAACAAAAAATATTGCCTATGAAGCTCAAACAGT 2400

ATACCAACAAGTGTGAACCCCCAAGCTTGCATGTAGGTAAAGGTCCAGTACTTTGAGGTA 2460

CAAGGTATACTAAAAATGTCCATTAAATTATTTTGACATGTATCTGCAAAATGTGTTAAG 2520

CATGGGCTTGTGAGATATTTATCTTTTGGAGGAATTTTTAGATAATATTTCCTTAGTGAT 2580

AGAACTGTATTAAAGCTTTCCTTTTACTTAACCTAGTAATTGGAAGAGTTATGCACATAA 2640

TTTTATATTTTGATGAAAAACTAGGGTAGAAATGCAAATTCAGTGGAGATGGTTGAAACT 2700

GTCACTGCCAGAATAAACAACAATAAATACATTCAAAGAAGGATTTACTTATTTTTTTCT 2760

TTTCTTTGAAAGAGATTATCTTTTCCTCAACAATGCTCGTAATAAAAAAAAAACACCAGA 2820

GACTTGAGACTTGGTTTATAGTATTGATGCTCAAGGATACCATTCTTGTCATTCAGATGG 2880

CAATTTTGTTAGTTGCTACTCTTTCATATCAACATTAAAAGGGAAGTAAATGATAGTTTT 2940

ACAGGGTTATGCTTATGGAGGAAATTGAAGGAAATTACTCTTTTCATCTCTCTTATATTT 3000

```
TATCCTTCTATTTTCTAAGTCATACTTAAGCAATGGCTTCTGCACTCACTTTGTCTCAAA 3060
GATATGAAAATAATGACGACATCTGACAATCTCAGCTACAGTCACCATGGAGATTGTAAA 3120
TGACTACTAAAGTTGCTGCTGCATCTTCATCTTTGATTCTTCCTCAAAGGTTTTTCTGAC 3180
CTCCTCCTATCATTAGTCTGTTGAATCCTCCTCCTTTCCTTTTCATTCTGATTGTCCTTC 3240
TCAGCCTCATTAATTTTCATTCTGTCCTATTTTTGGTAGCTCTATTTCTATTGCTCTGCT 3300
GGCTAAGAGCTTTCAGTGCCATGATAGTAAGAATTTTAACTTGTATCTAGTGAGATAACA 3360
GCAAGTTATCTCAAGTGGTATTTTAATTTTACAGCACATAATGAGCATAAAGTTATATCA 3420
AGAGACTGAAAAAAAAGTTGTATTTCCAGTAAATGCTAAAAGTTGCTCTTCAAAAGAAGG 3480
AAAGGCACTCACTGGTGGTATTGAATGAAGACCCTGTATAAATAAATTCTAGAAAGCAAA 3540
AAAAAAAAAAAAAA                                               3554
```

Figure 3.18 The full length cDNA sequences of *BUB3 budding uninhibited by benzimidazoles 3 homolog* (3554 bp in length with an ORF of 981 bp corresponding to a polypeptide of 326 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined.

Cdc2 (Fig. 3.19) contained a S_TKc domain (serine/threonine protein kinase catalytic domain; positions 4th - 287th, E-value = 6.40e-104). The expected MW and p*I* of this deduced protein were 34.69 kDa and 8.46, respectively. Protein kinases are a group of enzymes that possess a catalytic subunit which transfers the gamma phosphate from nucleotide triphosphates (often ATP) amino acid residues in a protein substrate side chain, resulting in a conformational change affecting protein functions. Proteins containing serine/threonine protein kinase catalytic domain can be divided into two broad classes with respect to substrate specificity; serine/threonine or tyrosine.

Thioredoxin 1 (Fig. 3.20) contained a thioredoxin domain (positions 2th - 104th, E-value = 1.30e-43). The expected MW and p*I* of this gene product were 12.0 kDa and 4.57, respectively. Thioredoxins are small enzymes that participate in the redox reactions, via the reversible oxidation of an active centre disulfide bond.

Multiprotein bridging factor 1 (Fig. 3.21) contained a HTH_XRE domain (positions 80th - 135th, E-value = 1.41e-09). The expected MW and p*I* of this gene product were 16.64 kDa and 10.11, respectively. Helix-turn-helix XRE family-like proteins are large family of DNA binding helix-turn helix proteins.

MLC1 (Fig. 3.22) contained two EFh domains (calcium binding motif; positions 12th - 40th and 86th - 114th, E-value = 1.62e+00 and 8.99e+00). The expected MW and p*I* of this gene product were 17.71 kDa and 4.74, respectively. EF-hands are

calcium-binding motifs that occur at least in pairs. Each motif consists of a 12 residue loop flanked on either side by a 12 residue alpha-helix. EF-hands undergo a conformational change upon binding calcium ions.

```
GGCACGAGGCAGCTGAATCAAGATGAAGCTTCACTTTCTTCTCGGTATTCTATTGTCGAC 60
GGTAATCATGTTGGCGGAAGCGAAAAATGGGCTGAGCAGTCGGACAGGCAGTAAAGAACA 120
CATGGACATCTAAACCAGTGTCCGGGAAATCCTCTTTTCCAGGCAGGATCAACATTTCTT 180
TTTAGTCTCAATAAAAGAAATTGTGATTAAAAAAAAAAAACCAAAAAAAAAAAAAAAAAC 240
CAAAGGCGACGGACGTGGGAAGCAAAATGGAGGATTACTTACGTATAGAAAAGCTTGGAG 300
                          M  E  D  Y  L  R  I  E  K  L  G  E 12
AGGGAACATATGGCGTGGTATACAAAGCCAAGAACCGCAAAAGTGGGAAGTTTGTGGCCA 360
  G  T  Y  G  V  V  Y  K  A  K  N  R  K  S  G  K  F  V  A  M 32
TGAAAAAGATCAGACTGGAGAATGAGGAAGAAGGTGTCCCTTCCACAGCTATCAGAGAAA 420
  K  K  I  R  L  E  N  E  E  E  G  V  P  S  T  A  I  R  E  I 52
TCTCTCTCCTGAAAGAACTTCAGCATCCCAACATTGTCCTACTAGAAGATGTATTGATGC 480
  S  L  L  K  E  L  Q  H  P  N  I  V  L  L  E  D  V  L  M  Q 72
AGGAGAGCAAACTTTTCCTTGTGTTCGAGTTCCTCAACATGGATTTGAAGAAATATCTTG 540
  E  S  K  L  F  L  V  F  E  F  L  N  M  D  L  K  K  Y  L  D 92
ACTCTTTGGAATCTGGCAAATATGTAGATAAGAAACTTGTGAAATCTTACTGCTACCAGC 600
  S  L  E  S  G  K  Y  V  D  K  K  L  V  K  S  Y  C  Y  Q  L 112
TTTTCCAAGGAATTCTCTATTGCCATCAGCGAAGGGTGCTCCACAGAGATCTCAAACCAC 660
  F  Q  G  I  L  Y  C  H  Q  R  R  V  L  H  R  D  L  K  P  Q 132
AGAATCTCCTCATCAATGAGCAGGGCGTCATAAAGATTGCTGATTTTGGCCTTGCTCGCG 720
  N  L  L  I  N  E  Q  G  V  I  K  I  A  D  F  G  L  A  R  A 152
CATTTGGAATCCCAGTGAGAGTGTATACACATGAGGTTGTGACTCTGTGGTATCGAGCTC 780
  F  G  I  P  V  R  V  Y  T  H  E  V  V  T  L  W  Y  R  A  P 172
CAGAGGTCCTTCTTGGTTCCTCTCGATACTCCTGTCCTGTTGATGTTTGGTCTCTTGGCT 840
  E  V  L  L  G  S  S  R  Y  S  C  P  V  D  V  W  S  L  G  C 192
GTATATTTGCCGAGATGGTTACTAAACGGCCACTGTTCCATGGTGACTCAGAGATTGACC 900
  I  F  A  E  M  V  T  K  R  P  L  F  H  G  D  S  E  I  D  Q 212
AGCTCTTCAGGATATTCAGAACCTTAACAACCCCCACAGAAGACAACTGGCCTGGTGTAA 960
  L  F  R  I  F  R  T  L  T  T  P  T  E  D  N  W  P  G  V  T 232
CACAACTGCAGGACTACAAGGCCAATTTCCCCAAGTGGACTGATTACAATCTTGGAAATT 1020
  Q  L  Q  D  Y  K  A  N  F  P  K  W  T  D  Y  N  L  G  N  S 252
CCGTCAAACAGATGGACAGCGATGGCTTGGACCTTTTATCGAAAACACTGATCTACGATC 1080
  V  K  Q  M  D  S  D  G  L  D  L  L  S  K  T  L  I  Y  D  P 272
CGACTCGAAGGATTTCTGCCAAGGAGGCCCTGAAGCACCCCTACTTTGATGATCTCGACA 1140
  T  R  R  I  S  A  K  E  A  L  K  H  P  Y  F  D  D  L  D  K 292
AGTCCACTCGTCCAGCCAAGAATTAAACCTTAATTTAGAGGACATTTATCTTTTCTCCAC 1200
  S  T  R  P  A  K  N  *                                     299
TTTCCTTTTTTTAATACTATCATCATCATTGTCATATTTTTATATTCTATTGGATGTGTA 1260
ACTTCTGAACAAAAGTTTCTAAGGGGCAGGACTTTTTGACCATTCGGGCCATATTTACAA 1320
CTATGTTCATGTCTGTCAGATGTGTACTACCAAATTATGGTAAAGTCACCATCACTACTT 1380
GTGAATGTCTGATTGAACAGATATATATTTGCATGGTTTATGACCACTGTTATTGCTCCT 1440
TTTACTTTGTGTTAGCATTCCACCAAGGTGGAAGAAACATATGCAGGACTGTGTGCTTTA 1500
AGTGCACGCTTTCAGTCTGATAAACAATTCTTCTGTTTTCTCTATTCAAGCTCTTCTTTG 1560
TAGCTCTGTCCTTTTGTATTATAGTTTTAGGTCTGCACACTATATATAATTATACAGAAC 1620
AATACATAACAGTATAAAAGGTGAAGTGTCTTTCAATTATAATTGTTGGGCAATAAAAGT 1680
TATTCAAATAAAAAAA                                             1696
```

Figure 3.19 The full length cDNA sequences of *cell division control protein 2* (*cdc2*, 1696 bp in length with an ORF of 900 bp corresponding to a polypeptide of 299 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A S_TKc domain (serine/threonine protein kinase catalytic domain; positions 4th - 287th of the deduced protein) is highlighted.

```
GCAAATTAAATTCTCCGTCCTTCGTGTCTCTTCCTTCAATCGCCAAGATGGTGTACCAAG 60
                                               M  V  Y  Q  V 5
TGAAAGATCAGGAAGATTTTACTAAGCAGCTAAACGAAGCTGGAAACAAGCTGGTCGTCA 120
  K  D  Q  E  D  F  T  K  Q  L  N  E  A  G  N  K  L  V  V  I 25
TCGACTTCTACGCCACCTGGTGTGGGCCATGCAAAATGATTGCACCTAAGCTGGAGGAGC 180
  D  F  Y  A  T  W  C  G  P  C  K  M  I  A  P  K  L  E  E  L 45
TGAGTCAGTCGATGAGCGATGTGGTTTTCCTGAAGGTGGATGTGGACGAATGCGAAGACA 240
  S  Q  S  M  S  D  V  V  F  L  K  V  D  V  D  E  C  E  D  I 65
TTGCCCAAGATAACCAGATTGCATGCATGCCTACTTTTCTGTACATGAAGAATGGCCAGA 300
  A  Q  D  N  Q  I  A  C  M  P  T  F  L  Y  M  K  N  G  Q  K 85
AGCTTGACAGCTTGTCTGGTGCCAACTACGAAAAGCTTGTCGAACTCATCGAGAAGCACA 360
  L  D  S  L  S  G  A  N  Y  E  K  L  V  E  L  I  E  K  H  K 105
AGTAATCCATTCCACTGCTCTCTCTGGCACCAAGAGCATGAAAGATGGACCATCTTTGCA 420
  *
AATTAGATCTGCTTAAAGTATTTTGTTTTAGGTTAATAGTGTTGTGATTAAAGACAAAAA 480
GCAAACTGTTTTGTTGTTTACGGTCATAAATATTCACTTTTTTTCCAAATTTCTCAAAAA 540
AAAAAAAAAAAAAGATTTAACAAGCACCTGTTTGTGGAAATAATTTCTTTTTGTTTTGAT 600
GGAAAATAGTTTAGGCGGCTAATTATTTCAGGTGAATATCCAAGATCTTTTCATTATTGC 660
ATTTAACTACCAGTGACTGTTTGTGCCTTTTTATGTAATGTAGAGCTTGGATTTATAAGA 720
TTTCAATAAAG                                                  731
```

Figure 3.20 The full length cDNA sequences of *thioredoxin 1* (731 bp in length with an ORF of 318 bp corresponding to a polypeptide of 105 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A thioredoxin domain (positions 2th - 104th of the deduced protein) is highlighted.

```
CGCCTCAATTTTCCGTGAATAATGGCTGATAGCGACTGGGACACCGTGACTGTGCTCAGG 60
                     M  A  D  S  D  W  D  T  V  T  V  L  R   13
AAGAAGCCCCAGAAGGCCTCTCAGCTTAAAAGTGAGCAGGCTGTGAACAAGGCCAGACGA 120
K  K  P  Q  K  A  S  Q  L  K  S  E  Q  A  V  N  K  A  R  R   33
ATTGGAGCTCAAGTGGAAACCAGCACGAAATATGGAGCTGCCAGCAACAAGCAACATGCC 180
I  G  A  Q  V  E  T  S  T  K  Y  G  A  A  S  N  K  Q  H  A   53
ACCACCTTCAACACAGCCAAGCTCGATCGTGAGACTGAGGAACTGAAGCACAACAAAGTG 240
T  T  F  N  T  A  K  L  D  R  E  T  E  E  L  K  H  N  K  V   73
GCTCCAGATGTAGGCCGTCTTATCCAGCAAGGACGCCAGAACAAGGGCTGGACGCAGAAG 300
A  P  D  V  G  R  L  I  Q  Q  G  R  Q  N  K  G  W  T  Q  K   93
GACCTGGCCACTCGTGTGAATGAAAAGCCACAGGTGATCCAGGAATATGAGCAAGGCAAG 360
D  L  A  T  R  V  N  E  K  P  Q  V  I  Q  E  Y  E  Q  G  K   113
GCAGTGCCCAACCAAAACATCATAACCAAGATTGAGAAGGCTATTGGGATGCGCCTGCGA 420
A  V  P  N  Q  N  I  I  T  K  I  E  K  A  I  G  M  R  L  R   133
GGCAAGGACAAAGGACAGCCCCTACCAGCTCCAGGGAGCAAGAAGAAATAAGCAGAAGCC 480
G  K  D  K  G  Q  P  L  P  A  P  G  S  K  K  K  *            149
AAGTCCCTCCGCCTTTTGCTCTTTCACCCACAAAGTGTGTACCATGCTACAGGCAGTACC 540
CAGAAAGTTCACCTTCAGTTTACTACTCGGCAATTAGGGCTATTCGGAATATTGTGTTTA 600
GTGTTTCATATGTCAAGTGCTTTTGTTCTCTCACCTAGCTTGGAGGGCATTTTTCTTTTT 660
CTCTTGAGGAAAACAATTTTTCTTGCAGAGTTTGGATAATAGATATATCACTAAA      715
```

Figure 3.21 The full length cDNA sequences of *multiprotein bridging factor 1* (715 bp in length with an ORF of 450 bp corresponding to a polypeptide of 149 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A HTH_XRE domain (positions 80th - 135th of the deduced protein) is highlighted.

```
CTTTTGCCCCGTTCACGTGGTGTGCTAGAGGCTAAAACAAAATGGCCGCGGATCTCAGTG 60
                                        M  A  A  D  L  S  A  7
CTCGTGATGTTGAGAGGGTGAAATTCGCCTTCTCCATCTATGATTTCGAGGGTAATGGCA 120
 R  D  V  E  R  V  K  F  A  F  S  I  Y  D  F  E  G  N  G  T  27
CCATGGATGCCTACTACATTGGCGACTGCCTGCGTGCCCTCAACCTGAACCCGACCCTGT 180
 M  D  A  Y  Y  I  G  D  C  L  R  A  L  N  L  N  P  T  L  S  47
CCGTGATCGAGAAGGTCGGTGGCAAGGAGAAGAGGAAAGAGAAGATGATTAAGCTCGACG 240
 V  I  E  K  V  G  G  K  E  K  R  K  E  K  M  I  K  L  D  E  67
AATTCATGCCCATCTTCGCCCAGGTCAAGAAGGACAAGGATGCCGGCTCCTTCGAAGATT 300
 F  M  P  I  F  A  Q  V  K  K  D  K  D  A  G  S  F  E  D  F  87
TCATGGAAGTCCTGAAGCTTTACGACAAAGCTGAGAACGGCACCATGATGTATGCTGAGC 360
 M  E  V  L  K  L  Y  D  K  A  E  N  G  T  M  M  Y  A  E  L  107
TTGAGCACATCCTTCTGTCCCTTGGTGAGCGTCTTGAGAAAGCTGAGTTGGAGCCCGTCC 420
 E  H  I  L  L  S  L  G  E  R  L  E  K  A  E  L  E  P  V  L  127
TTAAGGAGTGCTGCCCCGAGGAAGACGAAGAAGGCTTCATTCCCTACGAACCGTTCCTTA 480
 K  E  C  C  P  E  E  D  E  E  G  F  I  P  Y  E  P  F  L  K  147
AGAGACTTTTGGCCTTCAAAATCTAGAGGAAGCGTTTACCTTCCTCAAGAAGATCACCCA 540
 R  L  L  A  F  K  I  *                                       154
ACTGCTCTAAACGTCCACCATCTTCTTTTGTGACGTCATCACCAGGGAGTGCACATGCGT 600
ACAACACCACTACCATTCGGAGTGTCCGCGCGTGCAGTGTACAATCATTCTGGGGGAGTG 660
ACCAAACCACAGACAGTACATTAGTTCCAAAAAAAATATGTTAATATATATTGCCAAAAT 720
ATACGAAAATCATAGAAAGTGATTCCTACGAAAAAATGGGAAGCACAAGCTTAAACAAAA 780
AAAAAAATGTGAGGCACCCATTACATGCTCATTATGTCCTTCTTGTCCGTATGTGTGTGT 840
GTTCGTGCAGCGTCGGGTCCCATGGATCTGTTGAGCGTTTTTCTCTCATCCACACGGGGT 900
GTCAAAGGGCCGAGGCATCCGTGACCGCCTCATCATGTAGACCTGAAGTGACCACTGAAT 960
GATCCGCAGTTAGCCACCTGCTCGCTTTTACCACCTGTCTGAAGCGGGGACTAACAGCCA 1020
TTGACGACAAGAGCGGCCGTCTGCCCCACCCGGTTACCATCGTGAAGACCACCTCAATCT 1080
CGGCGGGTGCGGGCGGCGGCGCGAGGGAAAGGAGGGGGCGTTGGAACCCACCATCACTCT 1140
TCTTTCTCCCCCCCTCCTTCTCCCGACCCGGTCACGTGGTGCCTGATCCGGCCCCTTGTG 1200
CCCAATAGTGCCTCTGCGCCTCGCCCAACTCGTAGCGGCCAGGGGGAGGGCCACAGTCAG 1260
CCAGTGCGTCCATGGAGTGAACATACAAGGCCACACGTCAGTCAGTCAGTCAGTCCGAGA 1320
AGTCCGCTTAGCATGGCATCGTTACAAACTGCTGTCTGTTGGCACATTGGTTAAGTAATC 1380
TATGGTCTCAATAAAACAAAACATCTAAAAAAAA                           1414
```

Figure 3.22 The full length cDNA sequences of *myosin 1 light chain* (1414 bp in length with an ORF of 465 bp corresponding to a polypeptide of 154 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. Two EFh domains (positions 12th - 40th and 86th - 114th of the deduced protein) are highlighted.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.8 Summary of the full length cDNA of gene homologues found in the testes and heart cDNA libraries

Clone No.	Transcripts	Closest species	E-value	Full length
TT-N-S01-0144-W	*Small ubiquitin-like modifier, SUMO, small ubiquitin-like modifier SMO-1 (10.2 kD)*	*Apis mellifera*	1e-37	1305 bp (ORF 93 amino acids, 282 bp)
TT-N-S01-0020-W	*Peptidyl-prolyl cis-trans isomerase (cyclophilin A)*	*Maconellicoccus hirsutus*	4e-66	929 bp (ORF 164 amino acids, 495 bp)
TT-N-S01-0067-W	*Dynactin subunit 5 (dynactin subunit p25)*	*Strongylocentrotus purpuratus*	3e-68	1436 bp (ORF 197 amino acids, 594 bp)
TT-N-S01-0153-W	*BUB3 budding uninhibited by benzimidazoles 3 homolog (mitotic checkpoint)*	*Xenopus tropicalis*	1e-143	3554 bp (ORF 326 amino acids, 981 bp)
TT-N-S01-0190-W	*Cell division control protein 2 homolog (Cyclin-dependent kinase 1) (CDK1) (Cell division cycle 2)*	*Carassius auratus*	9e-132	1696 bp (ORF 299 amino acids, 900 bp)
TT-N-S01-0449-W	*Translationally controlled tumor protein (TCTP)*	*Penaeus monodon*	4e-91	730 bp (ORF 168 amino acids, 507 bp)
TT-N-S01-0872-W	*Ubiquitin conjugating enzyme 2 CG6720-PA, isoform A isoform 1*	*Apis mellifera*	2e-86	1416 bp (ORF 224 amino acids, 675 bp)
TT-N-S01-0873-W	*Small nuclear ribonucleoprotein polypeptide G*	*Aedes aegypti*	1e-23	676 bp (ORF 77 amino acids, 234 bp)
TT-N-S01-0897-W	*Thioredoxin 1*	*Litopenaeus vannamei*	1e-51	ORF 105 amino acids, 318 bp*
TT-N-S01-1026-W	*Multiprotein bridging factor 1*	*Bombyx mori*	4e-50	ORF 149 amino acids, 450 bp*
HT-N-S01-0243-W	*Myosin 1 light chain*	*Apis mellifera*	7e-49	1414bp (ORF 154 amino acids, 465 bp)

*Full length cDNA without the poly A tail

In addition, the full length cDNAs of 17 functionally important gene homologues were also characterized. Both 5′ (550 bp) and 3′ RACE-PCR (2 kb) of a *low molecular weight neurofilament protein XNF-L* homologue was successfully carried out (Fig. 3.23). The newly unidentified transcripts of 2194 and 2406 bp with an identical ORF encoding a polypeptide of 112 amino acids but polymorphic 3′ UTR were obtained (Figs. 3.24-3.25). This transcript (called *P. monodon testis-specific transcript, PMTST1*) did not match any gene in the GenBank (E-value > 1e-04). The prediction of translational start site did not reveal the possible site according to the best known Kozak rule (CTCATGG rather than A/GXXATGG, Kozak, 1983). Accordingly, *PMTT1* need to be further characterized. Nucleotide sequences of two *PMTST1* isoforms were aligned and polymorphic 3′ UTRs were observed (Fig.3.26).

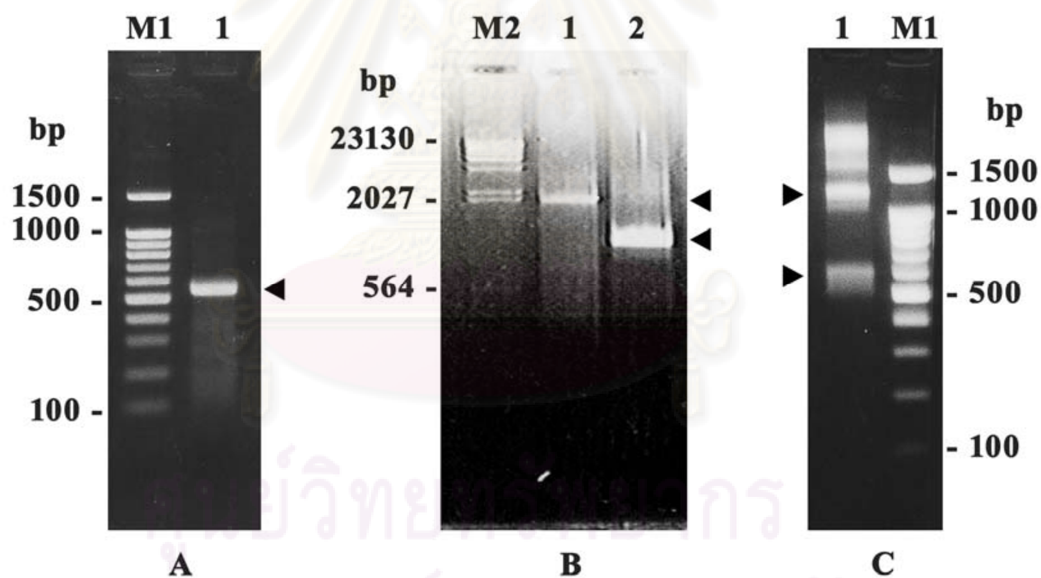


Figure 3.23 5′RACE-PCR (A, lane 1) and 3′RACE-PCR (B, lane 1) products of *low molecular weight neurofilament protein* (called *PMTST1*) and 3′RACE-PCR products of *prohibitin-2* (B, lane 2), and *growth factor receptor-bound protein* (C, lane 1) of *P. monodon*. A 100 bp DNA ladder (lanes M1) and λ *Hind* III (lane M2) were used as the markers.

The 3′ RACE-PCR of *prohibitin-2 (a repressor of estrogen receptor activity)* homologue was successfully carried out. The RACE product of approximately 1000 bp in size was obtained (Fig. 3.23). The full length cDNA of *prohibitin-2* was composed of 1382 bp with the predicted ORF of 891 bp in length corresponding to a polypeptide of 296 amino acids (Fig. 3.27). This sequence significantly matched that of *Tribolium castaneum* (E-value = 2e-128). The expected MW and p*I* of this deduced protein were 32.56 kDa and 9.73, respectively. A deduced prohibitin-2 protein contained a prohibitin domain (PHB, positions 39th - 200th, E-value = 1.25e-42).

The 3′ RACE-PCR products of *growth factor receptor-bound protein* (*GFRBP*) were 600 bp and 1200 bp in length (Fig. 3.23). Nucleotide sequences of EST and 3′ RACE products were assembled and two different forms of the full length cDNA (1188 bp and 1883 bp with an identical ORF of 636 bp corresponding to a polypeptide of 211 amino acids) were found (Fig. 3.28-3.29). Nucleotide sequences of two *GFRBP* isoforms were aligned and polymorphic 3′ UTRs were observed (Fig.3.30). This sequence significantly matched that of *Aedes aegypti* (E-value = 5e-91). A deduced GFRBP protein contained a Src homology 2 domain (SH2, positions 58th - 140th, E-value = 6.37e-35) and 2 domains of Src homology 3 motifs (SH3, positions 1-57 and 155-210, E-value = 3.34e-18 and 4.76e-23, respectively). The expected MW and p*I* of this gene product were 24.44 kDa and 5.51, respectively.

```
GACTCTCGTTCCCGTCTTCTGACAGTATATACAGCTCAATCTCAGTCCTGTCCAAGTCTT  60
GAGGCCGTCCTCATGGCTAACCCACGCCACGAAGCCGTCAAGAACGACTTGCTGGAGGCA  120
            M  A  N  P  R  H  E  A  V  K  N  D  L  L  E  A   16
GTCAAGAGCCGCGAGGCCGGCGCTTACCTGGAAGAGATCGACGCCTTCCTGAAGCAAAAG  180
V  K  S  R  E  A  G  A  Y  L  E  E  I  D  A  F  L  K  Q  K   36
AAGAAGTACAACGCCGACGACGTCAACCTGGCCCACCAGATCATTGACGTGTGTCTCGTC  240
K  K  Y  N  A  D  D  V  N  L  A  H  Q  I  I  D  V  C  L  V   56
AGCGACATCCTCGAGAAGGAGCTGGAGGATGTGAACGGGCGCCTCAAGGAGAAGTACACG  300
S  D  I  L  E  K  E  L  E  D  V  N  G  R  L  K  E  K  Y  T   76
CTTAGCAACGAGAACGCCGAGCTCAAAGTCAAGATGAAGAAGAAGACGACGGTGAAGAAG  360
L  S  N  E  N  A  E  L  K  V  K  M  K  K  K  T  T  V  K  K   96
GGGAAAGCGAGGCGGTACAAGAAGGTTCCGGCCGACAAGGAGGAGTCATAAGGAGGACCA  420
G  K  A  R  R  Y  K  K  V  P  A  D  K  E  E  S  *            112
TGCATCGGATATCCATAATTGGTGTTCAGTTGAAGGAGCAGGCCTTTGGTATTCAAGGAC  480
GACCTCCCACGGACGAGGAGAGCCTCGAGACGACAAAACCTTCGGCAGGATAAGGACCTC  540
CTCGCACCATGGGTCGTCGGAGGGACTTCGGAGGGCCAGGGTGCTTCGTCCAAGAACTGA  600
ACATGGATCTAGTGATATCGGTGCTGTGATCAGTGGCATTGGAGAATCATTGGGGTGTTC  660
TTCCTACCTTTCCTCGTGCTTTTATGCGTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT  720
CTCTCTCGCTCGCTCTCTCTCTGAACTTATATTTGTCGGTCATCATCTACTTTTTTCTCC  780
GTCATTTTGTTATTCAACTTTACCACATCAGTCATTATTTCTTTTCAAAACAACTGTTTT  840
GACCACCATGGCTATCTCGTTTTCGCCTTTCATATTTTACCTTTATTTCTCGAGTTTTCT  900
TTTTTCTTCGCATTTTTTTCACCATCGTACTCTCCTCGTTTTCCTTTTTTTTTTGACTATCT  960
CTTCTCTTCGCAGAAATAGTTAGCCTTTCTCTTTAGTTTACCTTAACATTCTGTCTTTCT  1020
ATGTCTCCTATAATATGTTATATTTTTTCTTGCATTCTCCCGCTTCTGTTCTATCTGTCA  1080
```

```
TATCATTTTTTTTCTCGCTCACTTTCTTCTCATTTTATAATCTTATAATCTTCTTAAATG 1140
CAAGTTGTTTTAAGGAAAGTGAAGTGAATATATGCATTTAACAATTTATTTATTCGGCGA 1200
ATGCAAAGAGTGGAAATGTCTACACAAGAAAATTACAGGATATAAGAGGAAAATGGGATA 1260
ATTGAAAGCGACAGAAAAAGTAGACCATTAGTCTGTGTAAGAATTATCGATGATTGGTTT 1320
TACTGCAGTGAAAAATGATTGAAAGTTTGCAAAGTTAGAAAGATTTTTCATTTAGGTCAG 1380
TGGCTATTCGTAGTTTTCGTTCAGTGATAATTCAACTGCAAGGAATAACTGAAAAAAGTT 1440
AATTAGCAATTTATAGCCACGTTCCCTTGTGTAAATATGTTCCAATATACCCTCTGCATT 1500
TACTTACGAGATATCATGTACTTGCATACATAAGTCTACATTTGTTTTCATGCAAGGAAG 1560
GTGGTAAAACACTCAACCGTGCAGTATACGCTGAATATAGAAATTTTTATATATCTTCAT 1620
TCAAGAAATCTATTTTGTGTGCAGTGGGTTTTTCTGCCAAGTTAGCTGTTTTAGGTATTT 1680
TAGAAATACTTCGGAATTCTCATTTTTTACTTAATCGATTTAATCATTATATTGTTATTA 1740
TTATATATATCTATTTTTTTGTGTGTGTGAACTGCGTAACAAATGTGGTATTTGTGTGTG 1800
CTGATATATATTCATTACTGTGCTGAGTTCAAATTGGTGCTGTAAACCCTCGTTCCTAAA 1860
GAGAAATGCAATCTCATTTAATGTAGACTCGGGACACTCTATAAGATTAGCTTTTGATCT 1920
TTGTCACTGCTGTTATGCTAAATTTCATTTCCTTCTTTTTTTAATTGTATTGGTTGCTCA 1980
TATTCTAGTTTTTGACTCTTGAATATTCCATCTCTTTCATCTTTGAAAAGGAAACTAAAT 2040
AAGAATTTTCAACAGAATGTATGGGTTTAGATACTGATTATTGGTATTGATAACCTTATA 2100
AATTTTATACACTAATTTCTCATATAATGATTTGTTCATTTGTTATAGCTCCCCTCTCCA 2160
TACTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAA                           2194
```

Figure 3.24 The full length cDNA and deduced protein sequences (2194 bp in length with an ORF of 339 bp corresponding to a polypeptide of 112 aa) of a short form of *PMTST1* of *P. monodon* (*PMTST1-s*). Start and stop codons were illustrated in boldface and underlined.

```
GACTCTCGTTCCCGTCTTCTGACAGTATATACAGCTCAATCTCAGTCCTGTCCAAGTCTT 60
GAGGCCGTCCTCATGGCTAACCCACGCCACGAAGCCGTCAAGAACGACTTGCTGGAGGCA 120
            M  A  N  P  R  H  E  A  V  K  N  D  L  L  E  A   16
GTCAAGAGCCGCGAGGCCGGCGCTTACCTGGAAGAGATCGACGCCTTCCTGAAGCAAAAG 180
V  K  S  R  E  A  G  A  Y  L  E  E  I  D  A  F  L  K  Q  K   36
AAGAAGTACAACGCCGACGACGTCAACCTGGCCCACCAGATCATTGACGTGTGTCTCGTC 240
K  K  Y  N  A  D  D  V  N  L  A  H  Q  I  I  D  V  C  L  V   56
AGCGACATCCTCGAGAAGGAGCTGGAGGATGTGAACGGGCGCCTCAAGGAGAAGTACACG 300
S  D  I  L  E  K  E  L  E  D  V  N  G  R  L  K  E  K  Y  T   76
CTTAGCAACGAGAACGCCGAGCTCAAAGTCAAGATGAAGAAGAAGACGACGGTGAAGAAG 360
L  S  N  E  N  A  E  L  K  V  K  M  K  K  K  T  T  V  K  K   96
GGGAAAGCGAGGCGGTACAAGAAGGTTCCGGCCGACAAGGAGGAGTCATAAGGAGGACCA 420
G  K  A  R  R  Y  K  K  V  P  A  D  K  E  E  S  *            112
TGCATCGGATATCCATAATTGGTGTTCAGTTGAAGGAGCAGGCCTTTGGTATTCAAGGAC 480
GACCTCCCACGGACGAGGAGAGCCTCGAGACGACAAAACCTTCGGCAGGATAAGGACCTC 540
CTCGCACCATGGGTCGTCGGAGGAACTTCGGAGGGCCAGGGTGCTTCGTCCAAGAACTGA 600
ACATGGATCTAGTGATATCGGTGCTGTGATCAGTGGCATTGGAGAATCATTGGGGTGTTC 660
TTCCTACCTTTCCTCATGCTTTTATGCGTAATCTCTCTCTCTCTCTCGCTCTCTCTCTGA 720
ACTTATATTTGTCGGTCATCATCTACTTTTTTCTCCTTCATTTTGTTATTAAATTTTAGT 780
CACTCCCAGCTTTTATGGAATCTCTCTCTCTCTCTCGCTCTCTCTCTGAACTTATATTCC 840
CCTTATAGGATCTACTTTTTTCTCCTTCATTTTGTTATTCCCTTTAACCACATCAGTCAT 900
TATTTCTTTTCAAAACAGCTGTTATGACCACCATGGCTATCTCGTTTCCTCCTTTCATAT 960
TTTAGCTTTATTTCTCGAGTTTTCTTTTCTCTTCGCATTTTTTCACCATCTTACTCTCCT 1020
CGTTTTCTTTTTCTTTTCTTTTTTTCTTTTTTCCTTTGACTATCTCTTCTCTTCGCAGAA 1080
ATAGTTAGCCTTTCTCTTAACTTTACCTTAACACTCTGTCTTTCTATGTCTCCTGTAATA 1140
TGATATTTTTTTTTCTTGCATTCTCCCGCTTCTGTTCTATCTGTCATATCATTTTTTTTC 1200
TCGCTCACTTTCTTCTCATTTTATAATCTTATAATCTTCTTAAATGCAAGTTGTTTTAAG 1260
GAAAGTGAAGTGAATATATGCATTTAACAATTTATTTATTCGGCAAATGCAAAGAGTGGA 1320
AATGTCTACACAAGAAAATTACAGGATAAAAGTGAAAATGGGATAATTGAAAGCGACAGA 1380
AAAAGTAGACCATTAGTCTGTGTAAGAATTATCGATGATCGGTTTTACTGCAGTGAAAAA 1440
TGATTGAAAGTTTGCAAAGTTAGAAAGATTTTTCATTTAGGTCAGTGGCTATTCGTAGTT 1500
TTCGTTCAGTGATAATTCAACTGCAAGGAATAACTGAAAAAAGTTAATTAGCAATTTATA 1560
GCCACGTTCCCTTGTGTAAATATGCTCCAATATACCCTCTGCATTTACTTACGAGATACC 1620
ATGTACTTGCATACATAAGTCTACATTTGTTTTCATGCAAGGAAGGTGGTAAAACACTCA 1680
ACCGTGCAGTATACGCTGAATATAGAAATTTTTATATATCTTCATTCAAGAAATCTATTT 1740
TGTGTGCAGTGGGTTTTTCTGCCAAGTTAGCTGTTTTAGGTATTTTAGAAATACTTCGGA 1800
ATTCTCATTTTTTACTTAATCGATTTAATCATTATACTGTTATTATTATATATATCTATT 1860
```

```
TTTTGTGTGTGAACTGCGTAACAAATGTGGTATTTGTGTGTGCTGATATATATTCATTAC 1920
TGTGCTGAGTTCAAATTGGTGCTGTAAACCCTGGTTCCTAAAGAGAAATGGAATTTCATT 1980
TAATGTAGACTCGGGACACTCTATAAGATTAGCTTTTGATTTTTGTCACTGCTGTTATGC 2040
TAAATTTCATTTCCTTTTTTTTTTTAATTGTATTGATTGCTCATATTCTAGTTTTTGACTC 2100
TTGAATATTCCATTTCTTTCATCTTTGTAAAGGAAACTAAATAAGAATTTTCAACAGAAT 2160
GTATGGGTTTAGATACTGATTATTGGTATTGATAACCTTATAAATTTTATACACTAATTT 2220
CTCATATATTGATTTGTTTATTTGTTATAGCTCCCCTCTCCATACTGCCATATATTTTCA 2280
CTTTGTAAAGGAAACGTTGTATAATGATCATATTATTTGAATATAATGATCATGATATTC 2340
AAACGTTATTATATATTCCAATACATTTTATAAGTTGAAAAAAAAAAAAAAAAAAAAAAA 2400
AAAAAA                                                       2406
```

Figure 3.25 The full length cDNA and deduced protein sequences (2406 bp in length with an ORF of 339 bp corresponding to a polypeptide of 112 aa) of a long form of *PMTST1* of *P. monodon* (*PMTST1-l*). Start and stop codons were illustrated in boldface and underlined.

```
PMTST1-s      GACTCTCGTTCCCGTCTTCTGACAGTATATACAGCTCAATCTCAGTCCTGTCCAAGTCTT 60
PMTST1-l      GACTCTCGTTCCCGTCTTCTGACAGTATATACAGCTCAATCTCAGTCCTGTCCAAGTCTT 60
              ************************************************************
PMTST1-s      GAGGCCGTCCTCATGGCTAACCCACGCCACGAAGCCGTCAAGAACGACTTGCTGGAGGCA 120
PMTST1-l      GAGGCCGTCCTCATGGCTAACCCACGCCACGAAGCCGTCAAGAACGACTTGCTGGAGGCA 120
              ************************************************************
PMTST1-s      GTCAAGAGCCGCGAGGCCGGCGCTTACCTGGAAGAGATCGACGCCTTCCTGAAGCAAAAG 180
PMTST1-l      GTCAAGAGCCGCGAGGCCGGCGCTTACCTGGAAGAGATCGACGCCTTCCTGAAGCAAAAG 180
              ************************************************************
PMTST1-s      AAGAAGTACAACGCCGACGACGTCAACCTGGCCCACCAGATCATTGACGTGTGTCTCGTC 240
PMTST1-l      AAGAAGTACAACGCCGACGACGTCAACCTGGCCCACCAGATCATTGACGTGTGTCTCGTC 240
              ************************************************************
PMTST1-s      AGCGACATCCTCGAGAAGGAGCTGGAGGATGTGAACGGGCGCCTCAAGGAGAAGTACACG 300
PMTST1-l      AGCGACATCCTCGAGAAGGAGCTGGAGGATGTGAACGGGCGCCTCAAGGAGAAGTACACG 300
              ************************************************************
PMTST1-s      CTTAGCAACGAGAACGCCGAGCTCAAAGTCAAGATGAAGAAGAAGACGACGGTGAAGAAG 360
PMTST1-l      CTTAGCAACGAGAACGCCGAGCTCAAAGTCAAGATGAAGAAGAAGACGACGGTGAAGAAG 360
              ************************************************************
PMTST1-s      GGGAAAGCGAGGCGGTACAAGAAGGTTCCGGCCGACAAGGAGGAGTCATAAGGAGGACCA 420
PMTST1-l      GGGAAAGCGAGGCGGTACAAGAAGGTTCCGGCCGACAAGGAGGAGTCATAAGGAGGACCA 420
              ************************************************************
PMTST1-s      TGCATCGGATATCCATAATTGGTGTTCAGTTGAAGGAGCAGGCCTTTGGTATTCAAGGAC 480
PMTST1-l      TGCATCGGATATCCATAATTGGTGTTCAGTTGAAGGAGCAGGCCTTTGGTATTCAAGGAC 480
              ************************************************************
PMTST1-s      GACCTCCCACGGACGAGGAGAGCCTCGAGACGACAAAACCTTCGGCAGGATAAGGACCTC 540
PMTST1-l      GACCTCCCACGGACGAGGAGAGCCTCGAGACGACAAAACCTTCGGCAGGATAAGGACCTC 540
              ************************************************************
PMTST1-s      CTCGCACCATGGGTCGTCGGAGGGACTTCGGAGGGCCAGGGTGCTTCGTCCAAGAACTGA 600
PMTST1-l      CTCGCACCATGGGTCGTCGGAGGAACTTCGGAGGGCCAGGGTGCTTCGTCCAAGAACTGA 600
              *********************** ************************************
PMTST1-s      ACATGGATCTAGTGATATCGGTGCTGTGATCAGTGGCATTGGAGAATCATTGGGGTGTTC 660
PMTST1-l      ACATGGATCTAGTGATATCGGTGCTGTGATCAGTGGCATTGGAGAATCATTGGGGTGTTC 660
              ************************************************************
PMTST1-s      TTCCTACCTTTCCTCGTGCTTTTATGCGTAATCTCTCTCTCTCTCTCTCTCTCTCTCT-- 718
PMTST1-l      TTCCTACCTTTCCTCATGCTTTTATGCGTAATCTCTCTCTCTCTCGCTCTCTCTCTCTGA 720
              *************** ******************************* **********
PMTST1-s      ------------------------------------------------------------
PMTST1-l      ACTTATATTTGTCGGTCATCATCTACTTTTTTCTCCTTCATTTTGTTATTAAATTTTAGT 780

PMTST1-s      -----------------------CTCTCTCTCGCTCGCTCTCTCTCTGAACTTATATTTG 755
PMTST1-l      CACTCCCAGCTTTTATGGAATCTCTCTCTCTCTCGCTCTCTCTCTGAACTTATATTCC 840
                                     ********* *************************
PMTST1-s      TCGGTCATCATCTACTTTTTTCTCCGTCATTTTGTTATTCAACTTTACCACATCAGTCAT 815
PMTST1-l      CCTTATAGGATCTACTTTTTTCTCCTTCATTTTGTTATTCCCTTTAACCACATCAGTCAT 900
               *    *  **************** **************   ** **************
PMTST1-s      TATTTCTTTTCAAAACAACTGTTTTGACCACCATGGCTATCTCGTTTTCGCCTTTCATAT 875
PMTST1-l      TATTTCTTTTCAAAACAGCTGTTATGACCACCATGGCTATCTCGTTTCCTCCTTTCATAT 960
              ***************** ***** *********************** * **********
PMTST1-s      TTTACCTTTATTTCTCGAGTTTTCTTTTTTCTTCGCATTTTTTCACCATCGTACTCTCCT 935
PMTST1-l      TTTAGCTTTATTTCTCGAGTTTTCTTTTCTCTTCGCATTTTTTCACCATCTTACTCTCCT 1020
              **** *********************** ********************* *********
PMTST1-s      CGTTTTCCTTTTTTTTT--------------------GACTATCTCTTCTCTTCGCAGAA 975
PMTST1-l      CGTTTTCTTTTTCTTTTCTTTTTTTCTTTTTTCCTTTGACTATCTCTTCTCTTCGCAGAA 1080
```

```
                ******* **** ****                    ***********************
PMTST1-s        ATAGTTAGCCTTTCTCTTTAGTTTACCTTAACATTCTGTCTTTCTATGTCTCCTATAATA 1035
PMTST1-l        ATAGTTAGCCTTTCTCTTAACTTTACCTTAACACTCTGTCTTTCTATGTCTCCTGTAATA 1140
                ****************** * ************ ******************** *****
PMTST1-s        TGTTATATTTTTT-CTTGCATTCTCCCGCTTCTGTTCTATCTGTCATATCATTTTTTTTC 1094
PMTST1-l        TGATATTTTTTTTTCTTGCATTCTCCCGCTTCTGTTCTATCTGTCATATCATTTTTTTTC 1200
                ** *** ****** **********************************************
PMTST1-s        TCGCTCACTTTCTTCTCATTTTATAATCTTATAATCTTCTTAAATGCAAGTTGTTTTAAG 1154
PMTST1-l        TCGCTCACTTTCTTCTCATTTTATAATCTTATAATCTTCTTAAATGCAAGTTGTTTTAAG 1260
                ************************************************************
PMTST1-s        GAAAGTGAAGTGAATATATGCATTTAACAATTTATTTATTCGGCGAATGCAAAGAGTGGA 1214
PMTST1-l        GAAAGTGAAGTGAATATATGCATTTAACAATTTATTTATTCGGCAAATGCAAAGAGTGGA 1320
                ******************************************** ***************
PMTST1-s        AATGTCTACACAAGAAAATTACAGGATATAAGAGGAAAATGGGATAATTGAAAGCGACAG 1274
PMTST1-l        AATGTCTACACAAGAAAATTACAGGATAAAAGTG-AAAATGGGATAATTGAAAGCGACAG 1379
                **************************** *** * *************************
PMTST1-s        AAAAAGTAGACCATTAGTCTGTGTAAGAATTATCGATGATTGGTTTTACTGCAGTGAAAA 1334
PMTST1-l        AAAAAGTAGACCATTAGTCTGTGTAAGAATTATCGATGATCGGTTTTACTGCAGTGAAAA 1439
                **************************************** *******************
PMTST1-s        ATGATTGAAAGTTTGCAAAGTTAGAAAGATTTTTCATTTAGGTCAGTGGCTATTCGTAGT 1394
PMTST1-l        ATGATTGAAAGTTTGCAAAGTTAGAAAGATTTTTCATTTAGGTCAGTGGCTATTCGTAGT 1499
                ************************************************************
PMTST1-s        TTTCGTTCAGTGATAATTCAACTGCAAGGAATAACTGAAAAAAGTTAATTAGCAATTTAT 1454
PMTST1-l        TTTCGTTCAGTGATAATTCAACTGCAAGGAATAACTGAAAAAAGTTAATTAGCAATTTAT 1559
                ************************************************************
PMTST1-s        AGCCACGTTCCCTTGTGTAAATATGTTCCAATATACCCTCTGCATTTACTTACGAGATAT 1514
PMTST1-l        AGCCACGTTCCCTTGTGTAAATATGCTCCAATATACCCTCTGCATTTACTTACGAGATAC 1619
                ************************* ********************************** 
PMTST1-s        CATGTACTTGCATACATAAGTCTACATTTGTTTTCATGCAAGGAAGGTGGTAAAACACTC 1574
PMTST1-l        CATGTACTTGCATACATAAGTCTACATTTGTTTTCATGCAAGGAAGGTGGTAAAACACTC 1679
                ************************************************************
PMTST1-s        AACCGTGCAGTATACGCTGAATATAGAAATTTTTATATATCTTCATTCAAGAAATCTATT 1634
PMTST1-l        AACCGTGCAGTATACGCTGAATATAGAAATTTTTATATATCTTCATTCAAGAAATCTATT 1739
                ************************************************************
PMTST1-s        TTGTGTGCAGTGGGTTTTTCTGCCAAGTTAGCTGTTTTAGGTATTTTAGAAATACTTCGG 1694
PMTST1-l        TTGTGTGCAGTGGGTTTTTCTGCCAAGTTAGCTGTTTTAGGTATTTTAGAAATACTTCGG 1799
                ************************************************************
PMTST1-s        AATTCTCATTTTTTACTTAATCGATTTAATCATTATATTGTTATTATTATATATATCTAT 1754
PMTST1-l        AATTCTCATTTTTTACTTAATCGATTTAATCATTATACTGTTATTATTATATATATCTAT 1859
                ************************************* **********************
PMTST1-s        TTTTTTGTGTGTGTGAACTGCGTAACAAATGTGGTATTTGTGTGTGCTGATATATATTCA 1814
PMTST1-l        TTTTT---GTGTGTGAACTGCGTAACAAATGTGGTATTTGTGTGTGCTGATATATATTCA 1916
                *****   ****************************************************
PMTST1-s        TTACTGTGCTGAGTTCAAATTGGTGCTGTAAACCCTCGTTCCTAAAGAGAAATGCAATCT 1874
PMTST1-l        TTACTGTGCTGAGTTCAAATTGGTGCTGTAAACCCTGGTTCCTAAAGAGAAATGGAATTT 1976
                ************************************ ***************** *** *
PMTST1-s        CATTTAATGTAGACTCGGGACACTCTATAAGATTAGCTTTTGATCTTTGTCACTGCTGTT 1934
PMTST1-l        CATTTAATGTAGACTCGGGACACTCTATAAGATTAGCTTTTGATTTTTGTCACTGCTGTT 2036
                ******************************************** ***************
PMTST1-s        ATGCTAAATTTCATTTCCTTCTTTTTTTAATTGTATTGGTTGCTCATATTCTAGTTTTTG 1994
PMTST1-l        ATGCTAAATTTCATTTCCTTTTTTTTTTAATTGTATTGATTGCTCATATTCTAGTTTTTG 2096
                ******************** ***************** *********************
PMTST1-s        ACTCTTGAATATTCCATCTCTTTCATCTTTGAAAAGGAAACTAAATAAGAATTTTCAACA 2054
PMTST1-l        ACTCTTGAATATTCCATTTCTTTCATCTTTGTAAAGGAAACTAAATAAGAATTTTCAACA 2156
                ***************** ************* ****************************
PMTST1-s        GAATGTATGGGTTTAGATACTGATTATTGGTATTGATAACCTTATAAATTTTATACACTA 2114
PMTST1-l        GAATGTATGGGTTTAGATACTGATTATTGGTATTGATAACCTTATAAATTTTATACACTA 2216
                ************************************************************
PMTST1-s        ATTTCTCATATAATGATTTGTTCATTTGTTATAGCTCCCCTCTCCATACTGCCAAAAAAA 2174
PMTST1-l        ATTTCTCATATATTGATTTGTTTATTTGTTATAGCTCCCCTCTCCATACTGCCATATATT 2276
                ************ ********* ******************************* * *
PMTST1-s        AAAAAAAAAAAAAAAAAAAA---------------------------------------- 2194
PMTST1-l        TTCACTTTGTAAAGGAAACGTTGTATAATGATCATATTATTTGAATATAATGATCATGAT 2336
                   *       ***  ***
PMTST1-s        ------------------------------------------------------------
PMTST1-l        ATTCAAACGTTATTATATATTCCAATACATTTTATAAGTTGAAAAAAAAAAAAAAAAAAA 2396

PMTST1-s        ----------
PMTST1-l        AAAAAAAAAA 2406
```

Figure 3.26 Pairwise alignments of different isoforms of *PMTST1* cDNAs of *P. monodon.*

```
GCACGAGGGTGTTTACATGTGAGTGAAGTTACATCAGCAAGAATGGGCGACAAACTGAAC 60
                                          M  G  D  K  L  N   6
GACCTCGCTGGACGCTTCGGCAAAGGACCTCGTGGTTTGGGTCTGGGTCTCAAGCTCTTG 120
D  L  A  G  R  F  G  K  G  P  R  G  L  G  L  G  L  K  L  L   26
GCAACGGCGGGAGCAGCGGCGTATGGCATCTCGCAGTCCATGTACACCGTTGAGGGTGGT 180
A  T  A  G  A  A  A  Y  G  I  S  Q  S  M  Y  T  V  E  G  G   46
CACAGAGCCATCATCTTCAACCGTATTGGAGGAGTGCAGCCAGATATTTACACTGAAGGG 240
H  R  A  I  I  F  N  R  I  G  G  V  Q  P  D  I  Y  T  E  G   66
TTGCACTTCAGGATTCCATGGTTCCAGTACCCAGTAGTCTATGATATCAGGGCTCGGCCT 300
L  H  F  R  I  P  W  F  Q  Y  P  V  V  Y  D  I  R  A  R  P   86
AGAAAGATCAGCTCACCCACAGGTAGCAAAGACTTGCAGATGGTGAACATTTCCCTTAGG 360
R  K  I  S  S  P  T  G  S  K  D  L  Q  M  V  N  I  S  L  R   106
GTCTTGTCACGCCCTGTAGGTACAGCCATCCCTAACATCCACCAGACCTTAGGGCCAGAC 420
V  L  S  R  P  V  G  T  A  I  P  N  I  H  Q  T  L  G  P  D   126
TTCGATGAGAAGGTGCTTCCATCTATTTGCAATGAAGTCCTCAAATCAGTTGTAGCAAAA 480
F  D  E  K  V  L  P  S  I  C  N  E  V  L  K  S  V  V  A  K   146
TTTAATGCTGCTCAATTAATCACAATGAGACAGCAAGTCTCTTTGATGATCCGCCGTGAT 540
F  N  A  A  Q  L  I  T  M  R  Q  Q  V  S  L  M  I  R  R  D   166
TTGACTCAGAGAGCAGAAGACTTCAACATAATCCTTGATGACGTTTCCATTACTGAGCTC 600
L  T  Q  R  A  E  D  F  N  I  I  L  D  D  V  S  I  T  E  L   186
AGCTTTGGCAGAGAATACACCAGTGCTGTTGAAGCCAAACAGGTGGCCCAGCAAGAGGCC 660
S  F  G  R  E  Y  T  S  A  V  E  A  K  Q  V  A  Q  Q  E  A   206
CAGCGAGCCTCTTTCATTGTCGAGAGAGCGAGACAGGAGAGGCAGCAGAAGATTGTGCAA 720
Q  R  A  S  F  I  V  E  R  A  R  Q  E  R  Q  Q  K  I  V  Q   226
GCTGAGGGTGAAGCTGAAGCTGCCAAATTGATTGGTAATGCCATTGGGTTGAATCCAGGT 780
A  E  G  E  A  E  A  A  K  L  I  G  N  A  I  G  L  N  P  G   246
TATCTGAAGCTCCGAAAGATCAAGGCTGCTGCTAGCATTGGCAAGACAATTTCTCAGGCA 840
Y  L  K  L  R  K  I  K  A  A  A  S  I  G  K  T  I  S  Q  A   266
CAGAACAGAGTGTATCTTGGTGCTGACACACTGATGCTCAACCTGAATGACAAGGATTTC 900
Q  N  R  V  Y  L  G  A  D  T  L  M  L  N  L  N  D  K  D  F   286
GATGCCAGCGCAACCCGTGTGACCAAGAAGTAAAACAGATGTTTAGCCAAGAAATATGAA 960
D  A  S  A  T  R  V  T  K  K  *                             296
AGTTAATGGGCAGTAGAGCAACTACAGTGTATTGAAGACTTATTGTCAGCCGCACTATAA 1020
TCCCAATTAATTTCTCCTTTTATGATGTTTTCTGTAGGAAAGCTCCTCATTATGTAACCC 1080
AGTGCTTTGATTTGCAGATACATGTTTTCCCTCTTCTCCCATCCAATTGTTAGAAGGTTT 1140
GACAACCATACTGGTTTACTTGCCTTCAGTGGATATGTATGGAAGGATTTAATTTTTGTT 1200
AGTTTATAGTAATAAACCTTGTTTGTACATGAAAAGTTGAAAGATTTGTGGATATTTTCA 1260
GACAAATGTACATATATTATGAATTCATTTGTCAAAGATGTATATAGATAGAAACAAATA 1320
TGGGCATCCACCTTTTTGCAAATAAATGTATAGGAAAAAAAAAAAAAAAAAAAAAAAAAA 1380
```

Figure 3.27 The full length cDNA and deduced protein sequences of *prohibitin-2* (1380 bp; ORF 891 bp, 296 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A prohibitin domain (PHB, positions 39th - 200th of the deduced protein) is highlighted.

```
GCGCCTTCCACCCTCACGCACTCTTAGGGCGAATCTCGGGCGTTTTCACCCTTTGCCACG 60
GCGAGAAAGGGCGGCTCTCGCGGCGTGCGCTCCGGAGGGACAGATTTCACCATGGAGGCG 120
                                                   M  E  A   3
ATAGCAAAACATGACTTTAACGCCACAGCTGAGGACGAGCTCAGTTTTAGGAAAGGGCAG 180
I  A  K  H  D  F  N  A  T  A  E  D  E  L  S  F  R  K  G  Q   23
ATTCTTAAGGTACTAAATATGGAAGATGATATGAACTGGTTCAGAGCAGAGCTGGACGGC 240
I  L  K  V  L  N  M  E  D  D  M  N  W  F  R  A  E  L  D  G   43
AGAGAAGGACTCATCCCTAGCAACTACATCGAGATGAAGAGTCATGAATGGTATTATGGA 300
R  E  G  L  I  P  S  N  Y  I  E  M  K  S  H  E  W  Y  Y  G   63
AGGATAACTCGCGCAGATGCGGAAAAACTCTTGCTTAATAAACACGAAGGAGCGTTCCTC 360
R  I  T  R  A  D  A  E  K  L  L  L  N  K  H  E  G  A  F  L   83
ATCCGAGTTAGTGAGAGTTCTCCGGGAGATTTTTCATTATCCGTCAAATGTGGAGATGGT 420
I  R  V  S  E  S  S  P  G  D  F  S  L  S  V  K  C  G  D  G   103
GTTCAGCACTTTAAGGTCTTGAGGGACACACAGGGCAAGTTTTTCCTCTGGGTCGTCAAG 480
V  Q  H  F  K  V  L  R  D  T  Q  G  K  F  F  L  W  V  V  K   123
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
TTCAACTCCCTAAATGAATTGGTGGAGTACCATCGGTCAGCGTCTGTGTCCCGGTCCCAT 540
F  N  S  L  N  E  L  V  E  Y  H  R  S  A  S  V  S  R  S  H   143
GACATTAAGCTCAAAGACATGACTCCAGAAGAATTCTTAGTGCAAGCCCTATACGACTTC 600
D  I  K  L  K  D  M  T  P  E  E  F  L  V  Q  A  L  Y  D  F   163
ACCCCTCAGGAGCAGGGCGAGTTGGAATTCAAGCGAGGTGATGTCATCACTGTCACAGAC 660
T  P  Q  E  Q  G  E  L  E  F  K  R  G  D  V  I  T  V  T  D   183
CGGTCAGACCCCCACTGGTGGAGCGGCGAAATGGGCAATCGCAGGGGGCTCTTTCCTGCC 720
R  S  D  P  H  W  W  S  G  E  M  G  N  R  R  G  L  F  P  A   203
ACCTACGTGGCTCCCTACCACACCTAGATGCCCAGTGCAGGAGCTCCACCTCGAGTACCA 780
T  Y  V  A  P  Y  H  T  *                                    211
CGTCATAACCGGAGTCAGCAGCCATTCGTACCAGGAGGCTGCTCAAATAGTATCTTAACA 840
GAAACAATGAAAGAGACCTTGCTGAAAACAATGAATGGAAACTTGGCCAGGCTTAAGGGT 900
GCTTTGGCCTACACACAGTGACAGACTGAGGGAGGCCTTGCAGGAGATGAATAGTAGTTG 960
GCTGGCACCCTATACAGTTTTTGGTTTGTGTTTCTGTGGCTTTCACCCCAGTCAGTAATT 1020
GTGCCTACATTCTACTTTGATTTTGCCCCCATCCATTTATATTGGATTAGCCAGTAAATA 1080
TGTTTTATTATTGGTGCCATCTCATACCTTTTCATCCTACATGGTTTTGTAAAGAATGAC 1140
TTTTTAAAAAAAAGGAATTACAAAAAAAAAAAAAAAAAAAAAAAAAAA             1188
```

Figure 3.28 The full length cDNA and deduced protein sequences of the short form of *growth factor receptor-bound protein* (*GFRBP-s*, 1188 bp; ORF 636 bp, 211 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A Src homology 2 domain (SH2, positions 58th - 140th) and 2 domains of Src homology 3 domain (SH3, positions 1th - 57th and 155th - 210th of the deduced protein) are highlighted.

```
GCGCCTTCCACCCTCACGCACTCTTAGGGCGAATCTCGGGCGTTTTCACCCTTTGCCACG 60
GCGAGAAAGGGCGGCTCTCGCGGCGTGCGCTCCGGAGGGACAGATTTCACCATGGAGGCG 120
                                                   M  E  A   3
ATAGCAAAACATGACTTTAACGCCACAGCTGAGGACGAGCTCAGTTTTAGGAAAGGGCAG 180
I  A  K  H  D  F  N  A  T  A  E  D  E  L  S  F  R  K  G  Q   23
ATTCTTAAGGTACTAAATATGGAAGATGATATGAACTGGTTCAGAGCAGAGCTGGACGGC 240
I  L  K  V  L  N  M  E  D  D  M  N  W  F  R  A  E  L  D  G   43
AGAGAAGGACTCATCCCTAGCAACTACATCGAGATGAAGAGTCATGAATGGTATTATGGA 300
R  E  G  L  I  P  S  N  Y  I  E  M  K  S  H  E  W  Y  Y  G   63
AGGATAACTCGCGCAGATGCGGAAAAACTCTTGCTTAATAAACACGAAGGAGCGTTCCTC 360
R  I  T  R  A  D  A  E  K  L  L  L  N  K  H  E  G  A  F  L   83
ATCCGAGTTAGTGAGAGTTCTCCGGGAGATTTTTCATTATCCGTCAAATGTGGAGATGGT 420
I  R  V  S  E  S  S  P  G  D  F  S  L  S  V  K  C  G  D  G   103
GTTCAGCACTTTAAGGTCTTGAGGGACACACAGGGCAAGTTTTTCCTCTGGGTCGTCAAG 480
V  Q  H  F  K  V  L  R  D  T  Q  G  K  F  F  L  W  V  V  K   123
TTCAACTCCCTAAATGAATTGGTGGAGTACCATCGGTCAGCGTCTGTGTCCCGGTCCCAT 540
F  N  S  L  N  E  L  V  E  Y  H  R  S  A  S  V  S  R  S  H   143
GACATTAAGCTCAAAGACATGACTCCAGAAGAATTCTTAGTGCAAGCCCTATACGACTTC 600
D  I  K  L  K  D  M  T  P  E  E  F  L  V  Q  A  L  Y  D  F   163
ACCCCTCAGGAGCAGGGCGAGTTGGAATTCAAGCGAGGTGATGTCATCACTGTCACAGAC 660
T  P  Q  E  Q  G  E  L  E  F  K  R  G  D  V  I  T  V  T  D   183
CGGTCAGACCCCCACTGGTGGAGCGGCGAAATGGGCAATCGCAGGGGGCTCTTTCCTGCC 720
R  S  D  P  H  W  W  S  G  E  M  G  N  R  R  G  L  F  P  A   203
ACCTACGTGGCTCCCTACCACACCTAGATGCCCAGTGCAGGAGCTCCACCTCGAGTACCA 780
T  Y  V  A  P  Y  H  T  *                                    211
CGTCATAACCGGAGTCAGCAGCCATTCGTACCAGGAGGCTGCTCAAATAGTATCTTAACA 840
GAAACAATGAAAGAGACCTTGCTGAAAACAATGAATGGAAACTTGGCCAGGCTTAAGGGT 900
GCTTTGGCCTACACACAGTGACAGACTGAGGGAGGCCTTGCAGGAGATGAATAGTAGTTG 960
GCTGGCACCCTATACAGTTTTTGGTTTGTGTTTCTGTGGCTTTCACCCCAGTCAGTAATT 1020
GTGCCTACATTCTACTTTGATTTTGCCCCCATCCATTTATATTGGATTAGCCAGTAAATA 1080
TGTTTTATTATTGGTGCCATCTCATACCTTTTCATCCTACATGGTTTTGTAAAGAATGAC 1140
TTTTTAAAAAAAAGGAATTGCATGTGTATTATTCTTGCCACAAGCTTTGATTGCAGAATT 1200
TTTAGAAATATATCAATGGAAAATACATTAGTATCCATAACATAAAATTGAATATTCAAT 1260
AGCTTGAAGAATAAACCAATTAGCCTTGCATGGCTTTTGCAGTGTTGCAGTGATTCATCT 1320
```

```
GGTACATAACAATTTGTGCACAAAATTATTGGTAATGTACTTTACAAAATGTTTATGTCC 1380
CACCTATGCACCTGGAGTGATTCTTTTTTTTCTTTCTTTCTTTTTTTCATTTTTTCTATT 1440
TTATTTTTTATTTTTTTTTCTCTCCTTTTGATAGGAAGCTGTTTCCACCTTTACCTGAAT 1500
TTAAGAAATCTTTACAGACAATTTGCCCAAAAGTCCATGAAATGTCTTTTTAGAGATAAG 1560
GTACATTTAGAGACACTATATTTTTTTGTTTTCTCGTTTCTCTTCGGACCTTAAATCCAG 1620
GTAAAGTTGGTCCAGATTTTAGAATTTATGATTCCTCGAATAGATCAAGCCATGTTTGCA 1680
AGAGGAAAGCACAAGTGATTGTTCGAGGTCTCTCTTTTACATTTTGTTTTATTCAGGCCA 1740
GGCAAGCACAAAAAACTTTTTTTTTTTTTTTAATACAGTTAGGCTTTTTTTTTTTTTTTT 1800
TTTTTTCTATCTTTCTTTCTTTTTGATTTTTGATTTTTTATATCTATATAGTGTGTAAAA 1860
AAAAAAAAAAAAAAAAAAAAAAAAA                                    1883
```

Figure 3.29 The full length cDNA and deduced protein sequences of the long form of *growth factor receptor-bound protein* (*GFRBP-l*, 1885 bp; ORF 636 bp, 211 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A Src homology 2 domain (SH2, positions 58th - 140th) and 2 domains of Src homology 3 domain (SH3, positions 1th - 57th and 155th - 210th of the deduced protein) were highlighted.

```
GFRBP-s      GCGCCTTCCACCCTCACGCACTCTTAGGGCGAATCTCGGGCGTTTTCACCCTTTGCCACG 60
GFRBP-l      GCGCCTTCCACCCTCACGCACTCTTAGGGCGAATCTCGGGCGTTTTCACCCTTTGCCACG 60
             ************************************************************
GFRBP-s      GCGAGAAAGGGCGGCTCTCGCGGCGTGCGCTCCGGAGGGACAGATTTCACCATGGAGGCG 120
GFRBP-l      GCGAGAAAGGGCGGCTCTCGCGGCGTGCGCTCCGGAGGGACAGATTTCACCATGGAGGCG 120
             ************************************************************
GFRBP-s      ATAGCAAAACATGACTTTAACGCCACAGCTGAGGACGAGCTCAGTTTTAGGAAAGGGCAG 180
GFRBP-l      ATAGCAAAACATGACTTTAACGCCACAGCTGAGGACGAGCTCAGTTTTAGGAAAGGGCAG 180
             ************************************************************
GFRBP-s      ATTCTTAAGGTACTAAATATGGAAGATGATATGAACTGGTTCAGAGCAGAGCTGGACGGC 240
GFRBP-l      ATTCTTAAGGTACTAAATATGGAAGATGATATGAACTGGTTCAGAGCAGAGCTGGACGGC 240
             ************************************************************
GFRBP-s      AGAGAAGGACTCATCCCTAGCAACTACATCGAGATGAAGAGTCATGAATGGTATTATGGA 300
GFRBP-l      AGAGAAGGACTCATCCCTAGCAACTACATCGAGATGAAGAGTCATGAATGGTATTATGGA 300
             ************************************************************
GFRBP-s      AGGATAACTCGCGCAGATGCGGAAAAACTCTTGCTTAATAAACACGAAGGAGCGTTCCTC 360
GFRBP-l      AGGATAACTCGCGCAGATGCGGAAAAACTCTTGCTTAATAAACACGAAGGAGCGTTCCTC 360
             ************************************************************
GFRBP-s      ATCCGAGTTAGTGAGAGTTCTCCGGGAGATTTTTCATTATCCGTCAAATGTGGAGATGGT 420
GFRBP-l      ATCCGAGTTAGTGAGAGTTCTCCGGGAGATTTTTCATTATCCGTCAAATGTGGAGATGGT 420
             ************************************************************
GFRBP-s      GTTCAGCACTTTAAGGTCTTGAGGGACACACAGGGCAAGTTTTTCCTCTGGGTCGTCAAG 480
GFRBP-l      GTTCAGCACTTTAAGGTCTTGAGGGACACACAGGGCAAGTTTTTCCTCTGGGTCGTCAAG 480
             ************************************************************
GFRBP-s      TTCAACTCCCTAAATGAATTGGTGGAGTACCATCGGTCAGCGTCTGTGTCCCGGTCCCAT 540
GFRBP-l      TTCAACTCCCTAAATGAATTGGTGGAGTACCATCGGTCAGCGTCTGTGTCCCGGTCCCAT 540
             ************************************************************
GFRBP-s      GACATTAAGCTCAAAGACATGACTCCAGAAGAATTCTTAGTGCAAGCCCTATACGACTTC 600
GFRBP-l      GACATTAAGCTCAAAGACATGACTCCAGAAGAATTCTTAGTGCAAGCCCTATACGACTTC 600
             ************************************************************
GFRBP-s      ACCCCTCAGGAGCAGGGCGAGTTGGAATTCAAGCGAGGTGATGTCATCACTGTCACAGAC 660
GFRBP-l      ACCCCTCAGGAGCAGGGCGAGTTGGAATTCAAGCGAGGTGATGTCATCACTGTCACAGAC 660
             ************************************************************
GFRBP-s      CGGTCAGACCCCCACTGGTGGAGCGGCGAAATGGGCAATCGCAGGGGGCTCTTTCCTGCC 720
GFRBP-l      CGGTCAGACCCCCACTGGTGGAGCGGCGAAATGGGCAATCGCAGGGGGCTCTTTCCTGCC 720
             ************************************************************
GFRBP-s      ACCTACGTGGCTCCCTACCACACCTAGATGCCCAGTGCAGGAGCTCCACCTCGAGTACCA 780
GFRBP-l      ACCTACGTGGCTCCCTACCACACCTAGATGCCCAGTGCAGGAGCTCCACCTCGAGTACCA 780
             ************************************************************
GFRBP-s      CGTCATAACCGGAGTCAGCAGCCATTCGTACCAGGAGGCTGCTCAAATAGTATCTTAACA 840
GFRBP-l      CGTCATAACCGGAGTCAGCAGCCATTCGTACCAGGAGGCTGCTCAAATAGTATCTTAACA 840
             ************************************************************
GFRBP-s      GAAACAATGAAAGAGACCTTGCTGAAAACAATGAATGGAAACTTGGCCAGGCTTAAGGGT 900
GFRBP-l      GAAACAATGAAAGAGACCTTGCTGAAAACAATGAATGGAAACTTGGCCAGGCTTAAGGGT 900
             ************************************************************
GFRBP-s      GCTTTGGCCTACACACAGTGACAGACTGAGGGAGGCCTTGCAGGAGATGAATAGTAGTTG 960
GFRBP-l      GCTTTGGCCTACACACAGTGACAGACTGAGGGAGGCCTTGCAGGAGATGAATAGTAGTTG 960
```

```
                ************************************************************
GFRBP-s         GCTGGCACCCTATACAGTTTTTGGTTTGTGTTTCTGTGGCTTTCACCCCAGTCAGTAATT 1020
GFRBP-l         GCTGGCACCCTATACAGTTTTTGGTTTGTGTTTCTGTGGCTTTCACCCCAGTCAGTAATT 1020
                ************************************************************
GFRBP-s         GTGCCTACATTCTACTTTGATTTTGCCCCCATCCATTTATATTGGATTAGCCAGTAAATA 1080
GFRBP-l         GTGCCTACATTCTACTTTGATTTTGCCCCCATCCATTTATATTGGATTAGCCAGTAAATA 1080
                ************************************************************
GFRBP-s         TGTTTTATTATTGGTGCCATCTCATACCTTTTCATCCTACATGGTTTTGTAAAGAATGAC 1140
GFRBP-l         TGTTTTATTATTGGTGCCATCTCATACCTTTTCATCCTACATGGTTTTGTAAAGAATGAC 1140
                ************************************************************
GFRBP-s         TTTTTAAAAAAAAGGAATTACAAAAAAAAAAAAAAAAAAAAAAAAAAA------------ 1188
GFRBP-l         TTTTTAAAAAAAAGGAATTGCATGTGTATTATTCTTGCCACAAGCTTTGATTGCAGAATT 1200
                ******************* **     *  *        * **
GFRBP-s         ------------------------------------------------------------
GFRBP-l         TTTAGAAATATATCAATGGAAAATACATTAGTATCCATAACATAAAATTGAATATTCAAT 1260

GFRBP-s         ------------------------------------------------------------
GFRBP-l         AGCTTGAAGAATAAACCAATTAGCCTTGCATGGCTTTTGCAGTGTTGCAGTGATTCATCT 1320

GFRBP-s         ------------------------------------------------------------
GFRBP-l         GGTACATAACAATTTGTGCACAAAATTATTGGTAATGTACTTTACAAAATGTTTATGTCC 1380

GFRBP-s         ------------------------------------------------------------
GFRBP-l         CACCTATGCACCTGGAGTGATTCTTTTTTTTTTTCTTTCTTTCTTTTTTTAATTTTTTCT 1440

GFRBP-s         ------------------------------------------------------------
GFRBP-l         TTTCTTTTATTATTTTTTTTTTCTCTCCTTTTGATAGGAAGCTGTTTCCACCTTTACCTG 1500

GFRBP-s         ------------------------------------------------------------
GFRBP-l         AATTTAAGAAATCTTTACAGACAATTTGCCCAAAAGTCCATGAAATGTCTTTTTAGAGAT 1560

GFRBP-s         ------------------------------------------------------------
GFRBP-l         AAGGTACATTTAGAGACACTATATTTTTTTGTTTTCTCGTTTCTCTTCGGACCTTAAATC 1620

GFRBP-s         ------------------------------------------------------------
GFRBP-l         CAGGTAAAGTTGGTCCAGATCTTAGAATTTATGATTCCTCGAATAGATCAAGCCATGTTT 1680

GFRBP-s         ------------------------------------------------------------
GFRBP-l         GCAAGAGGAAAGCACAAGTGATTGTTCGAGGTCTCTCTTTTACATTTTGTTTTATTCAGG 1740

GFRBP-s         ------------------------------------------------------------
GFRBP-l         CCAGGCAAGCACAAAAAACTTTTTTTTTTTTTTTTTTAATACTGTTAGGCTTTTTTTTTT 1800

GFRBP-s         ------------------------------------------------------------
GFRBP-l         TTTTTCTATCTTTCTTTCTTTTTGATTTTTGATTTTTTATATCTATATAGTGTGTAAAAA 1860

GFRBP-s         -------------------------
GFRBP-l         AAAAAAAAAAAAAAAAAAAAAAAAA 1885
```

Figure 3.30 Pairwise alignments of different isoforms of *growth factor receptor-bound protein* (*GFRBP*) of *P. monodon*.

Two different isoforms of *multiple inositol polyphosphate phosphatase* were found and further characterized. 5′ RACE-PCR of *MIPP* (clone no. TT-N-S01-0004-W) was carried out and the amplification product of 1,100 was obtained (Fig. 3.31). Whereas 3′ RACE-PCR of this gene do not identified. The full length cDNA of *MIPP* of *P. monodon* (1644 bp in length with an ORF of 1374 bp corresponding to a polypeptide of 457 amino acids, Fig. 3.32) was identified. This transcript significantly matched *multiple inositol polyphosphate phosphatase* of *Nasonia vitripennis* (E-value = 9e-35). A deduced MIPP protein contained an acid phosphatase A domain

(positions 52th - 398th, E-value = 3.20e-06). The expected MW and p*I* of a deduced *P. monodon* MIPP were 52.59 kDa and 8.83, respectively.

In addition, 5′ and 3′RACE-PCR of *MIPP2* (clone no. TT-N-S01-0678-W) generated fragments of 650 and 200 bp, respectively. After sequence assembly, the full length cDNA of *MIPP2* of *P. monodon* (1746 bp in length with an ORF of 1194 bp corresponding to 397 amino acids, Fig. 3.33) was successfully characterized. This transcript significantly matched with *MIPP2* (*CG4317-PA*) of *Apis mellifera* (E-value = 4e-31). A deduced MIPP2 protein contained an acid phosphatase A domain located at positions 27th - 338th of the deduced protein (E-value = 9.70e-07). The expected MW and p*I* of the deduced MIPP2 were 45.58 kDa and 8.91, respectively.

Deduced amino acid and nucleotide sequences of different isoforms of *MIPP* were aligned (Fig. 3.34-3.35) and results indicated large differences between these genes. This suggested that *MIPP* isoforms should be encoded from different loci.

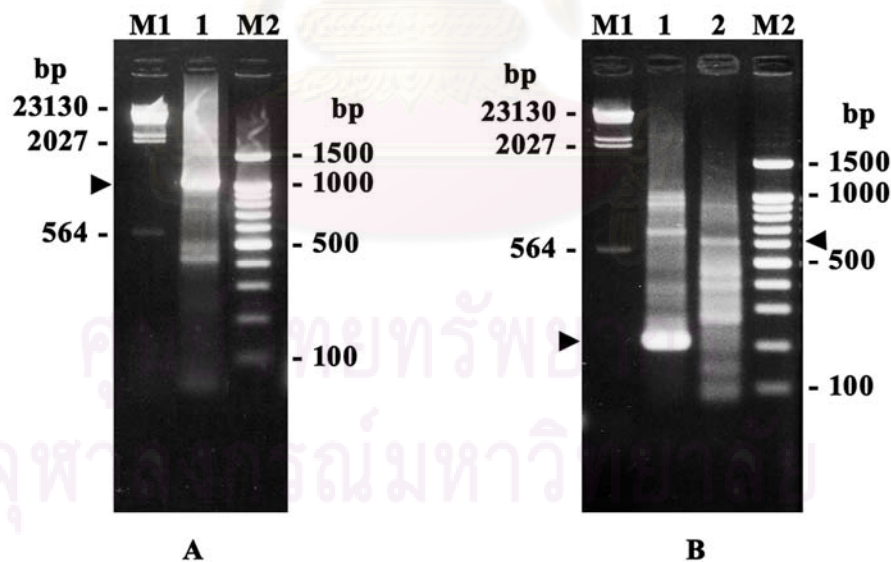


Figure 3.31 5′RACE-PCR products of *MIPP* (A, lane 1) and 3′ (B, lane 1) and 5′RACE-PCR of *MIPP2* (B, lanes 2) of *P. monodon*. λ-*Hind* III (lanes M1) and 100 bp markers (lanes M2) were used as the marker.

```
ACTCGGAGGCGGCAGCTTACGGAGACTCAGGGCCTTTAGACTTAACACAAAAGCTTATGG 60
                                                        M  D 2
ACTCTGCTAAGAGGAAAGCAGGGCAACGCAGGACCACGGGCAGTATGTATTGTCTCTCCG 120
  S  A  K  R  K  A  G  Q  R  R  T  T  G  S  M  Y  C  L  S  E 22
AAGACGCTAATCCTTACACGGGCTTCGCTACCATGACTCCATACAGGATCGCCTCCACGC 180
  D  A  N  P  Y  T  G  F  A  T  M  T  P  Y  R  I  A  S  T  P 42
CCTTGAAGGCGGACGATGTTATTCCGAGAGAATGTAAGCCAGTGCAGATATGGCACCTGA 240
  L  K  A  D  D  V  I  P  R  E  C  K  P  V  Q  I  W  H  L  I 62
TTCGCCACGGAAGCCACGGTGCTCACAGGAACGACTACATAAAGTTTGAAGATGAGTTGC 300
  R  H  G  S  H  G  A  H  R  N  D  Y  I  K  F  E  D  E  L  P 82
CATTTCTTAGGAGAAAGATCTTCCGAGCTCGCGCATTGGAAGGGGAATCTCTGTGACA 360
  F  L  R  R  K  I  F  R  A  R  A  H  W  K  G  N  L  C  D  K 102
AAGACCTGAAATTAATTAGGCTCTGGAAAGTGTCTAAGATGATGAGCAAGACAGGTACTC 420
  D  L  K  L  I  R  L  W  K  V  S  K  M  M  S  K  T  G  T  L 122
TTTCTGTGGAAGGTATGGAAGAAATTGCCGGTTTAGCTGACCGGTTCAAGTCCGTTTTCC 480
  S  V  E  G  M  E  E  I  A  G  L  A  D  R  F  K  S  V  F  P 142
CAGGTCTTCTTGAAAAAAAAATTCTCAGCTAAGCTACATCCCGGAGCTGGTAATAAGATTG 540
  G  L  L  E  K  K  F  S  A  K  L  H  P  G  A  G  N  K  I  A 162
CCTTCGGTACCGGCCGCCAGAACCAGCAGAGCGCCGTCGCCTATGTTTCTAGCATGTACG 600
  F  G  T  G  R  Q  N  Q  Q  S  A  V  A  Y  V  S  S  M  Y  G 182
GCCCCTTCGCGCGGTTCGTCCCCGTAGGTTCAATACCATCGAGGGATCTGCAGTTCTACG 660
  P  F  A  R  F  V  P  V  G  S  I  P  S  R  D  L  Q  F  Y  D 202
ACTACTGTAGAAACTACATTGAAAGTGTATTGAACATGCACAAGAAGTTGAAGCCGTACC 720
  Y  C  R  N  Y  I  E  S  V  L  N  M  H  K  K  L  K  P  Y  H 222
ACAACTTCATGTCAGGGAGCATCATGAATTCTGTTCTTCGAAGAGTGTCCGAACGTCTTG 780
  N  F  M  S  G  S  I  M  N  S  V  L  R  R  V  S  E  R  L  G 242
GATTTCCGGTAACCGTGGCCAATGTTCGTGTGATGTACAACGCATGTCGGTACTACTATG 840
  F  P  V  T  V  A  N  V  R  V  M  Y  N  A  C  R  Y  Y  Y  A 262
CGTGGTACAAAAATATTGTGTCGCCCTGGTGCACTGTCTTCACCCCGATGGATCTGAAGG 900
  W  Y  K  N  I  V  S  P  W  C  T  V  F  T  P  M  D  L  K  V 282
TGCTGGAATACTGGGAAGACCTAAGGGTATATCACGATCAAGGCCACCGCTTCGAGATCA 960
  L  E  Y  W  E  D  L  R  V  Y  H  D  Q  G  H  R  F  E  I  S 302
GCTCAAAGCAAGCCTGCGTTCTTGGCAAAGACGTGATGGATCAGTTCCGAAATCGAGTGG 1020
  S  K  Q  A  C  V  L  G  K  D  V  M  D  Q  F  R  N  R  V  E 322
AAAACGGCAGTACGGAACTCTACGCCGCCTCGTATTTCGTCAATCCGGAGGCTTTGGTAA 1080
  N  G  S  T  E  L  Y  A  A  S  Y  F  V  N  P  E  A  L  V  T 342
CGTTCATTACCTTACTGGGGCTGTTCAATGATGAGGAGCCCATAACTGAGTTTTACATCC 1140
  F  I  T  L  L  G  L  F  N  D  E  E  P  I  T  E  F  Y  I  P 362
CCTCGTCCCGCTTGTGGAAAACATCCCAGTTTGCTGGTTTCGGTAGCAACTTGGCCATCT 1200
  S  S  R  L  W  K  T  S  Q  F  A  G  F  G  S  N  L  A  I  L 382
TGCTCTCTTTGTGTGCTGATGACAGCTTTTGGGTGAGCGCTCTCCTCAACGAGAAACCAG 1260
  L  S  L  C  A  D  D  S  F  W  V  S  A  L  L  N  E  K  P  V 402
TTCAACTACCGGGATGTGACAGCAGCCTGGGTTGTCCTTGGAATAATTTCAGTCAATACT 1320
  Q  L  P  G  C  D  S  S  L  G  C  P  W  N  N  F  S  Q  Y  Y 422
ACGACTACCTAAGTGACTGCAACTTCGATGAGCTTTGTGGAAGCTTTTCGAGTCTGCTGA 1380
  D  Y  L  S  D  C  N  F  D  E  L  C  G  S  F  S  S  L  L  T 442
CGCAGAGTCGCCACTGGTATGCTCTCTACATGAATGAACAATGGATGTAGAAAACTCAGC 1440
  Q  S  R  H  W  Y  A  L  Y  M  N  E  Q  W  M  *             457
AGATGAATGCAGCTTTTATTTTCATCGTCCTTTCAGACCAGAAATGCATCACTTATGTGT 1500
TTAGTCCTACATTTTCATTTATTTCAAGAGAAAGGATTTAAGAAGGAAGAATTTTGGAGG 1560
AAATTTGTATATTGTTTGTGAACTTACCGGATGTATGGAAAACATGTATGATTTCAATTC 1620
AACATTAGAGTGAACCCTCGTGCC                                     1644
```

Figure 3.32 The full length cDNA and deduced amino acid sequences of *multiple inositol polyphosphate phosphatase* (*MIPP*, 1644 bp; ORF 1374 bp, 457 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An acid phosphatase A domain (positions 52th - 398th, E-value = 3.20e-06) is highlighted.

```
CACGGCAGCCGCAGTATGCACAGGATGGACTTCATGAAGATGGAGACCCAGCTGCCCATC 60
                        M  D  F  M  K  M  E  T  Q  L  P  I   12
CTCAAGAGGAAAATTCTAGCGGCTCACTCTTTAGGCAGCGGCGAACTGTGCACCCAGGAC 120
L  K  R  K  I  L  A  A  H  S  L  G  S  G  E  L  C  T  Q  D   32
CTGGCCCTGATCCGCGGCTGGAAACTCAAAGACATGGACAAGGGCAAGGCAGGCACGCTC 180
L  A  L  I  R  G  W  K  L  K  D  M  D  K  G  K  A  G  T  L   52
ACCCCGGAAGGGAGAGCCGAGGTGGAGAGCATCGCGTCCAGGTTCAAGGCTGCGTTCCCC 240
T  P  E  G  R  A  E  V  E  S  I  A  S  R  F  K  A  A  F  P   72
GACCTGGTCTACAAGCGGTACAGCATCTCAAAAACTGCCGCGGAGAACGTTCGACAACAA 300
D  L  V  Y  K  R  Y  S  I  S  K  T  A  A  E  N  V  R  Q  Q   92
CCTGGCGAAGGAACCAAGGTTGCCTTTGCGCCTTCCCGCCAGACCTACGAAAGCGCCGTG 360
P  G  E  G  T  K  V  A  F  A  P  S  R  Q  T  Y  E  S  A  V   112
AGCTACCTCGAGGCTCTGTACGGCCGCCAGTGGGGTCACGTCGGCCTCCCCGTGAGAGGC 420
S  Y  L  E  A  L  Y  G  R  Q  W  G  H  V  G  L  P  V  R  G   132
TCTCAAAGCTTGCAGTACTACGACTACTGCAAGAATTACATCAACAAAGTGGTGGCTCTG 480
S  Q  S  L  Q  Y  Y  D  Y  C  K  N  Y  I  N  K  V  V  A  L   152
AAAAAGAAGAGCAAACCCTTTCATATCTTCACCAAAGGGAAGTCTATGGAAGCCGTCATG 540
K  K  K  S  K  P  F  H  I  F  T  K  G  K  S  M  E  A  V  M   172
GACAGGGTCTCCAGACGAACTGGTGTTACCGTCAACTTAACCAAACTCCGCACCATGTAC 600
D  R  V  S  R  R  T  G  V  T  V  N  L  T  K  L  R  T  M  Y   192
AACGCCTGTCGGTACCAGAAGGCATGGGCTCCGCAGGATCCGTCCCCGTGGTGCGTCGTC 660
N  A  C  R  Y  Q  K  A  W  A  P  Q  D  P  S  P  W  C  V  V   212
TTCACCCCAAGCGATCTTCAGGTTCTGGAATACTGGGAAGACCTTCGGATACACCACGAC 720
F  T  P  S  D  L  Q  V  L  E  Y  W  E  D  L  R  I  H  H  D   232
CACGGTTACGCTCACTCAATTAACTACAAGCAGGCCTGTATTTTTGGCAACGATATCTTC 780
H  G  Y  A  H  S  I  N  Y  K  Q  A  C  I  F  G  N  D  I  F   252
CTCCATTTCCGAAACCGGATTGAATCTGGGGTCACTAATACCGACTCTACAACGTACTTC 840
L  H  F  R  N  R  I  E  S  G  V  T  N  T  D  S  T  T  Y  F   272
GTGAATATGGAAGCATTTGTGCCGTTCATGGCACTGTTGGGTTTATTCAAGGACCCCGAG 900
V  N  M  E  A  F  V  P  F  M  A  L  L  G  L  F  K  D  P  E   292
CCACTCACGAGTGAAGTTGCTAACCCGGATCGTGTGTGGAAAACGTCCAAGTTCGCTGGG 960
P  L  T  S  E  V  A  N  P  D  R  V  W  K  T  S  K  F  A  G   312
TATGGAAGTAACCTAGGGTTCCTCCTCTCGACTTGTGGAAACGACAGTGCCAGTTGGTGG 1020
Y  G  S  N  L  G  F  L  L  S  T  C  G  N  D  S  A  S  W  W   332
ATAACCGCCATTCTCAACGAAGAGAAGATCAAGTTGCCAGGCTGCGAAACTAGCTTGGGT 1080
I  T  A  I  L  N  E  E  K  I  K  L  P  G  C  E  T  S  L  G   352
TGCCCCTGGGAGAGGTTTGTAAGCGAGTACGATTTCCTTGAAGACTGTGACTTCACTCGC 1140
C  P  W  E  R  F  V  S  E  Y  D  F  L  E  D  C  D  F  T  R   372
CTCTGTGGACGTTATTCAGATAAACTGTGGCGGTCGCAGCACTGGCGTTTATCCTACATC 1200
L  C  G  R  Y  S  D  K  L  W  R  S  Q  H  W  R  L  S  Y  I   392
ATGCACAACTGGATGTAAGGGGACCGGTTAGGCCGTGCGAAGAAAGCTGTTTTCTTTTTC 1260
M  H  N  W  M  *                                             397
AACATTTCATTGAATAATGGATATTGGAGGATTTGTTTTAGTTCATGCAGTTATTACTAA 1320
CTAGAGGGAGGCTCTGATTTCTTTCACTTTTTGTGAAGCTCAGTGTTAGAGGTTCATTGT 1380
TAATTCCAAGATTTATTTTCTTTTATGTGTAGTTGAATGACAAAGGTGAACCGTTAATTT 1440
CAAGATGTAATGTGAAAAAGGTATAAGGAATACATATGTGCCAATATTCAAAAAAAAAAA 1500
AAAACTGAATCAAGATGAAGCTTCACTTTCTTCTCGGTATTCTATTGTCGACGGTAATCA 1560
TGTTGGCGGAAGCGAAAAATGGGCTGAGCAGTCGGAGAGGCAATAAAGAACACATGGACA 1620
TCTAAACCAGTGTCCTGGAAATCCTCTTTTCCAGGCAGGATCAACATTTCTTTTTAGTCT 1680
CAATAAAAGAAATTGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1740
AAAAAA                                                       1746
```

Figure 3.33 The full length cDNA and deduced amino acid sequences of *multiple inositol polyphosphate phosphatase 2* (*MIPP2*, 1746 bp; ORF 1194 bp, 397 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An acid phosphatase A domain (positions 27th - 338th, E-value = 9.70e-07) is highlighted.

```
MIPP            MDSAKRKAGQRRTTGSMYCLSEDANPYTGFATMTPYRIASTPLKADDVIPRECKPVQIWH 60
MIPP2           ------------------------------------------------------------

MIPP            LIRHGSHGAHRNDYIKFEDELPFLRRKIFRARAHWKGNLCDKDLKLIRLWKVSKMMS-KT 119
MIPP2           -----------MDFMKMETQLPILKRKILAAHSLGSGELCTQDLALIRGWKLKDMDKGKA 49
                            *::*:* :**:*:***: *::  .*:** :** *** **:..* . *:
MIPP            GTLSVEGMEEIAGLADRFKSVFPGLLEKKFS--------AKLHPGAGNKIAFGTGRQNQQ 171
MIPP2           GTLTPEGRAEVESIASRFKAAFPDLVYKRYSISKTAAENVRQQPGEGTKVAFAPSRQTYE 109
                ***: **  *: .:*.***:.**.*: *::*        .: :** *.*:**...**. :
MIPP            SAVAYVSSMYGPFARFVPVGSIPSRDLQFYDYCRNYIESVLNMHKKLKPYHNFMSGSIMN 231
MIPP2           SAVSYLEALYGRQWGHVGLPVRGSQSLQYYDYCKNYINKVVALKKKSKPFHIFTKGKSME 169
                ***:*:.::**    .* :    *:.**:****:***:.*: ::** **:* * .*. *:
MIPP            SVLRRVSERLGFPVTVANVRVMYNACRYYYAWYKNIVSPWCTVFTPMDLKVLEYWEDLRV 291
MIPP2           AVMDRVSRRTGVTVNLTKLRTMYNACRYQKAWAPQDPSPWCVVFTPSDLQVLEYWEDLRI 229
                :*: ***.* *..*.::::*.*******  **  :  ****.**** **:*********:
MIPP            YHDQGHRFEISSKQACVLGKDVMDQFRNRVENGSTELYAASYFVNPEALVTFITLLGLFN 351
MIPP2           HHDHGYAHSINYKQACIFGNDIFLHFRNRIESGVTNTDSTTYFVNMEAFVPFMALLGLFK 289
                :**:*:  ..*. ****::*::: :****:*.* *:  :::**** **:*.*::*****:
MIPP            DEEPITEFYIPSSRLWKTSQFAGFGSNLAILLSLCADD--SFWVSALLNEKPVQLPGCDS 409
MIPP2           DPEPLTSEVANPDRVWKTSKFAGYGSNLGFLLSTCGNDSASWWITAILNEEKIKLPGCET 349
                * **:*.    ..*:****:***:****.:*** *.:*  *:*::*:***: ::****::
MIPP            SLGCPWNNFSQYYDYLSDCNFDELCGSFSSLLTQSRHWYALYMNEQWM 457
MIPP2           SLGCPWERFVSEYDFLEDCDFTRLCGRYSDKLWRSQHWRLSYIMHNWM 397
                ******:.* . **:*.**:* .*** :*. * :*:**   *: .:**
```

Figure 3.34 Amino acid alignments of two isoforms of *multiple inositol polyphosphate phosphatase* (*MIPP*) of *P. monodon*.

```
MIPP            CTTATGGACTCTGCTAAGAGGAAAGCAGGGCAACGCAGGACCACGGGCAGTATGTATTGT 60
MIPP2           ------------------------------------------------------------

MIPP            CTCTCCGAAGACGCTAATCCTTACACGGGCTTCGCTACCATGACTCCATACAGGATCGCC 120
MIPP2           ------------------------------------------------------------

MIPP            TCCACGCCCTTGAAGGCGGACGATGTTATTCCGAGAGAATGTAAGCCAGTGCAGATATGG 180
MIPP2           ------------------------------------------------------------

MIPP            CACCTGATTCGCCACGGAAGCCACGGTGCTCACAGGAACGACTACATAAAGTTTGAAGAT 240
MIPP2           ------------------------------------ATGGACTTCATGAAGATGGAGACC 24
                                                    *  **** *** *** * **
MIPP            GAGTTGCCATTTCTTAGGAGAAAGATCTTCCGAGCTCGCGCGCATTGGAAGGGGAATCTC 300
MIPP2           CAGCTGCCCATCCTCAAGAGGAAAATTCTAGCGGCTCACTCTTTAGGCAGCGGCGAACTG 84
                 ** ****  * ** * *** ** **  *    **** * *     * *  **  * **
MIPP            TGTGACAAAGACCTGAAATTAATTAGGCTCTGGAAAGTGTCTAAGATG---ATGAGCAAG 357
MIPP2           TGCACCCAGGACCTGGCCCTGATCCGCGGCTGGAAACTCAAAGACATGGACAAGGGCAAG 144
                **   * * ******    * **  *   ******* *     * ***   * * *****
MIPP            ACAGGTACTCTTTCTGTGGAAGGTATGGAAGAAATTGCCGGTTTAGCTGACCGGTTCAAG 417
MIPP2           GCAGGCACGCTCACCCCGGAAGGGAGAGCCGAGGTGGAGAGCATCGCGTCCAGGTTCAAG 204
                 **** ** **  *   ****** *  *  **  * *   *  * **   * ********
MIPP            TCCGTTTTCCCAGGTCT--TCTTGAAA--------------AAAAATTCTCA-------- 453
MIPP2           GCTGCGTTCCCCGACCTGGTCTACAAGCGGTACAGCATCTCAAAAACTGCCGCGGAGAAC 264
                 * *  ***** *  **  ***  **               ***** *  *
MIPP            GCTAAGCTACATCCCGGAGCTGGTAATAAGATTGCCTTCGGTACCGGCCGCCAGAACCAG 513
MIPP2           GTTCGACAACAACCTGGCGAAGGAACCAAGGTTGCCTTTGCGCCTTCCCGCCAGACCTAC 324
                * *   * *** ** ** *  ** *  *** ******* *   *   ******** * *
MIPP            CAGAGCGCCGTCGCCTATGTTTCTAGCATGTACGGCCCCTTCGCGCGGTTCGTCCCCGTA 573
MIPP2           GAAAGCGCCGTGAGCTACCTCGAGGCTCTGTACGGCCGCCAGTGGGGTCACGTCGGCCTC 384
                 * ********   ***  *        ********* *     * *   ****  * *
MIPP            GGTTCAATACCATCGAGGGATCTGCAGTTCTACGACTACTGTAGAAACTACATTGAAAGT 633
MIPP2           CCCGTGAGAGGCTCTCAAAGCTTGCAGTACTACGACTACTGCAAGAATTACATCAACAAA 444
                      * *   **        ****** ************ *  ** *****  * *
MIPP            GTATTGAACATGCACAAGAAGTTGAAGCCGTACCACAACTTCATGTCAGGGAGCATCATG 693
MIPP2           GTGGTGGCTCTGAAAAAGAAGAGCAAACCCTTTCATATCTTCACCAAAGGGAAGTCTATG 504
                **  **    ** * ******   ** ** *  ** * *****    *****     ***
MIPP            AATTCTGTTCTTCGAAGAGTGTCCGAACGTCTTGGATTTCCGGTAACCGTGGCCAATGTT 753
MIPP2           GAAGCCGTCATGGACAGGGTCTCCAGACGAACTGGTGTTACCGTCAACTTAACCAAACTC 564
                 *  * **  *    ** ** ***  ***   ***  ** * ** * * *  ****  *
MIPP            CGTGTGATGTACAACGCATGTCGGTACTACTATGCGTGGTACAAAAATATTGTGTCGCCC 813
MIPP2           CGCACCATGTACAACGCCTGTCGGTACCAGAAGGCATGGGCTCCGCAGGATCCGTCCCCG 624
                **    *********** ********* *  * ** ***       *   *  *** **
```

```
MIPP             TGGTGCACTGTCTTCACCCCGATGGATCTGAAGGTGCTGGAATACTGGGAAGACCTAAGG 873
MIPP2            TGGTGCGTCGTCTTCACCCCAAGCGATCTTCAGGTTCTGGAATACTGGGAAGACCTTCGG 684
                 ******   *********** *  *****  **** ********************  **

MIPP             GTATATCACGATCAAGGCCACCGCTTCGAGATCAGCTCAAAGCAAGCCTGCGTTCTTGGC 933
MIPP2            ATACACCACGACCACGGTTACGCTCACTCAATTAACTACAAGCAGGCCTGTATTTTTGGC 744
                  ** * ***** ** **  **     *   ** * **  ***** *****  ** *****

MIPP             AAAGACGTGATGGATCAGTTCCGAAATCGAGTGGAAAACGGCAGTACGGAACTCTACGCC 993
MIPP2            AACGATATCTTCCTCCATTTCCGAAACCGGATTGAATCTGGGGTCACTAATACCGACTCT 804
                 ** **  *  *    ** ******** **  * ***   **    **  *   * ** *

MIPP             GCCTCGTATTTCGTCAATCCGGAGGCTTTGGTAACGTTCATTACCTTACTGGGGCTGTTC 1053
MIPP2            ACAACGTACTTCGTGAATATGGAAGCATTTGTGCCGTTCATGGCACTGTTGGGTTTATTC 864
                  *  **** ***** ***  *** ** ** **  *******  *  *  ****  * ***

MIPP             AATGATGAGGAGCC---CATAACTGAGTTTTACATCCCCTCGTCCCGCTTGTGGAAAACA 1110
MIPP2            AAGGACCCCGAGCCACTCACGAGTGAAGTTGCTAACC---CGGATCGTGTGTGGAAAACG 921
                 ** **    *****   **  * ***  **   * **   **   **  **********

MIPP             TCCCAGTTTGCTGGTTTCGGTAGCAACTTGGCCATCTTGCTCTC--TTTGTG----TGCT 1164
MIPP2            TCCAAGTTCGCTGGGTATGGAAGTAACCTAGGGTTCCTCCTCTCGACTTGTGGAAACGAC 981
                 *** **** ***** *  ** ** *** * *   ** * *****   *****     *

MIPP             GATGACAGCTTTTGGGTGAGCGCTCTCCTCAACGAGAAACCAGTTCAACTACCGGGATGT 1224
MIPP2            AGTGCCAGTTGGTGGATAACCGCCATTCTCAACGAAGAGAAGATCAAGTTGCCAGGCTGC 1041
                   ** *** *  *** * * ***  * ********  *     *  *  * ** ** **

MIPP             GACAGCAGCCTGGGTTGTCCTTGGAATAATTTCAGTCAATACTACGACTACCTAAGTGAC 1284
MIPP2            GAAACTAGCTTGGGTTGCCCCTGGGAGAGGTTTGTAAGCGAGTACGATTTCCTTGAAGAC 1101
                 ** *  *** ******* ** *** * *  **        * ***** * ***    ***

MIPP             TGCAACTTCGATGAGCTTTGTGGAAGCTTTTCGAGTCTGCTGACGCAGAGTCGCCACTGG 1344
MIPP2            TGTGACTTCACTCGCCTCTGTGGACGTTATTCAGATAAACTGTGGCGGTCGCAGCACTGG 1161
                 **  *****  *   ** ****** * * ***   *   ***  ** *   *  ******

MIPP             TATGCTCTCTACATGAATGAACAA-TGGATGTAG 1377
MIPP2            CGTTTATCCTACATCA-TGCACAACTGGATGTAA 1194
                   *     ****** * ** **** ********
```

Figure 3.35 Pairwise alignment of different forms of *P. monodon multiple inositol polyphosphate phosphatase* (*MIPP*).

A 450 bp fragment was obtained from 5′ RACE-PCR of *transformer-2* (*Tra-2*, Fig. 3.36). After sequence assembly, two forms of the full length cDNA of *P. monodon Tra*-2 (*Tra-2l*; 2673 bp, 804 bp, 267 aa and *Tra-2s*; 2658 bp, 789 bp, 262 aa, Figs. 3.37 and 3.38) was obtained. These different isoforms sequences significantly matched that of *Bombyx mori* (E-value = 1e-40 and 4e-40, respectively). Amino acid and nucleotide sequences of *Tra-2l* and *Tra-2s* were aligned and these deduced proteins were different due to an indel of a SVHSS pentapeptide (Figs. 3.39-3.40). A deduced nonsecretory Tra-2 protein contained a RNA recognition motif domain (RRM, positions 109th - 182th, E-value = 1.49e-22 for *Tra-2l* and positions 104th - 177th, E-value = 1.49e-22 for *Tra-2s*, respectively). The expected MW and p*I* of these gene products were 30.23 kDa and 11.46 and 29.73 kDa and 11.46, respectively.

The 5′ RACE-PCR product (1100 bp) of *Rac GTPase-activating protein 1* homologue was characterized (Fig. 3.36). Its full length cDNA was composed of 2838 bp with an ORF of 1812 bp in length corresponding to a polypeptide of 603 amino acids (Fig. 3.41). This sequence significantly matched that of *Apis mellifera* (E-value = 3e-99). A deduced Rac GTPase-activating protein 1 protein contained a RhoGAP

domain (positions 364th - 540th, E-value = 6.4e-47) and protein kinase C conserved region 1 domains or cysteine-rich domains (C1, positions 317th - 365th, E-value = 5.32e-06). The expected MW and p*I* of these gene products were 68.62 kDa and 8.58, respectively.

A 550 bp fragment was obtained from 3′ RACE-PCR of *flotillin-2* was carried out (Fig. 3.36). The full length cDNA of *Flotillin-2* of *P. monodon* was composed of 1937 bp with ORF of 1320 bp in length encoding a polypeptide of 439 amino acids (Fig. 3.42). This sequence significantly matched that of *Drosophila melanogaster* (E-value = 0.0). A deduced flotillin-2 contained a prohibitin domain (PHB, positions 87th -268th, E-value = 4.24e-04). The expected MW and p*I* of these gene products were 47.69 kDa and 5.38, respectively.

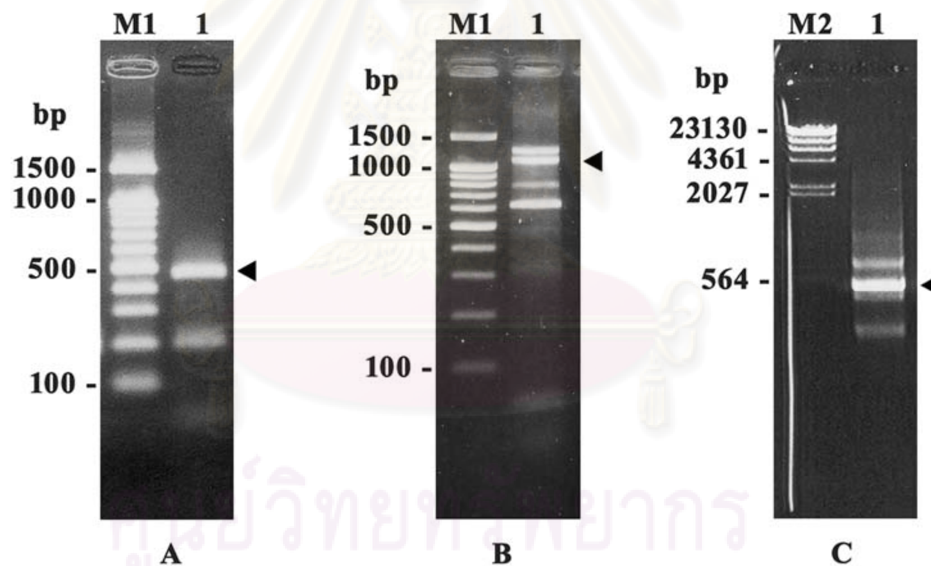


Figure 3.36 5′RACE-PCR products of *transformer-2* (A, lane 1) and *Rac GTPase-activating protein 1* (B, lane 1), and 3′RACE-PCR products of *flotillin-2* (C, lane 1) of *P. monodon*. A 100 bp marker (lanes M1) and λ-*Hind* III (lane M2) were used as the marker.

```
TACATGGTGGGAAAAGGCGGCTTCAACGCGTGAGGTTCCCGAATTTTTGCGATTTTTTGG 60
CTTTTTCTACTTAAAATGGAGAGTCCAGAGGGTGAACGGAACAGTCTCACACCTCGTTCG 120
               M  E  S  P  E  G  E  R  N  S  L  T  P  R  S   15
CGGTCGCGCTCAAGGTCACGTCTAGAGTCACCAGCCGCTTCCCCTGCTCATCGCCGCACA 180
R  S  R  S  R  S  R  L  E  S  P  A  A  S  P  A  H  R  R  T   35
GCCACTTCGCAGTCAAGGTCTCCCCAGCCTCGCAGACGCTCATTTTCCAGGTCGCGATCC 240
A  T  S  Q  S  R  S  P  Q  P  R  R  R  S  F  S  R  S  R  S   55
CGAACTCCACGCAGGCATCGCTCCCGGAGTGGCTCACCTCGAAATGGTCATGATGGAAGC 300
R  T  P  R  R  H  R  S  R  S  G  S  P  R  N  G  H  D  G  S   75
AGTCGCCGCTCCCGACGCAGCCGGTCGGTCCACTCAAGCTCCCCGATGTCGAACCGGAGG 360
S  R  R  S  R  R  S  R  S  V  H  S  S  S  P  M  S  N  R  R   95
CGGCATATTGGCACCAGAGACAATCCCGAGCCATCTAAGTGCCTGGGGGTGTTCGGCCTG 420
R  H  I  G  T  R  D  N  P  E  P  S  K  C  L  G  V  F  G  L   115
AGTGTGCATACCACGGAAAGACAGCTGTACACCATTTTTGACAAGTTTGGCCCTCTGGAG 480
S  V  H  T  T  E  R  Q  L  Y  T  I  F  D  K  F  G  P  L  E   135
AAAGTACAAGTAGTCCTGGATTCGAAAACGGGCAAGTCAAGAGGCTTTGCGTTTGTGTAC 540
K  V  Q  V  V  L  D  S  K  T  G  K  S  R  G  F  A  F  V  Y   155
TTTGAGTCACTTAAGGATGCCTCAGAAGCCAAAAATGAGTGTTCTGGGATGGAGATTGAC 600
F  E  S  L  K  D  A  S  E  A  K  N  E  C  S  G  M  E  I  D   175
GGCCGGAGGATTAGAGTTGATTATTCCATTACCAAGCGGCCACACACTCCAACTCCAGGA 660
G  R  R  I  R  V  D  Y  S  I  T  K  R  P  H  T  P  T  P  G   195
ATATACATGGGTAGACCCACATCTCGCGGTGGCTACGACCGGGGCTACGGCCGAGGAGGC 720
I  Y  M  G  R  P  T  S  R  G  G  Y  D  R  G  Y  G  R  G  G   215
CACCGCGGAGACAGGTACCGTTCCCCGTCACCCCGTTACCGGCCTCGCTCCAGCGGTGGT 780
H  R  G  D  R  Y  R  S  P  S  P  R  Y  R  P  R  S  S  G  G   235
CGCCGCGACTATTACGATCGTGGGTATGTACTCGGCCAGGGCGTTACCGCACAACTCAAG 840
R  R  D  Y  Y  D  R  G  Y  V  L  G  Q  G  V  T  A  Q  L  K   255
CCACCACAAGATATAATAGTTCATGTGGAGGCAAGCTAGTTAAAGCAAGTTTACAGAAGC 900
P  P  Q  D  I  I  V  H  V  E  A  S  *                        267
TTCTTAAGTAGGCCAAGGGGCTGAAGATGATCAACTCCCAAGTTAATGGTGAAGGTGGTT 960
AAGAATTGATGGATTGATGGAAAGGATTCACGGTTTGTTCAGAGCAACCCTCCAAAGGCA 1020
ATTGACAAGTGCAGATGAAAAAACAGTAAAATTGTGAGGAAAAACCTGTGAACTGAAGAT 1080
TTCTAAGAGATTTGAGGTATGACCGCGGGGACAGGAGCTACGACAGAGGCTATGACCGCT 1140
ACGACAGGCCTCAATACCATGACCGCTATGATAGGTATGACCGTGCTTATGACAAGTACG 1200
ATCGCTACGATAGGTCCAGATCTCGCTCCTATTCTCCACGAAGATACAAGTATTGATGAT 1260
TGTACACCAGACAAAACTGCAAGATTACTACAGAAAGGTTGTAGATGCTCCATACTTGAC 1320
GGTGCTGTGCAGACGGCTGGCTCCAGAGAACGTTTCGACCGTTATACAGTTACAAATTCT 1380
CAGTCTCATTTGTCATGCTAGTTATCAAAATATTTTAAGATAAGCTGTAGATAACAAGGT 1440
CTAAGAAAGAAAAACACTTTTTTTTTTTTCTATTTTTTCTTTCTTTTTTCTGTCTTTTTT 1500
CTGTCTTTTCTTATAGGTAATGTACTCTCCTCCACCCATATCCCCTTCCCTTCTTCCCGA 1560
GTCAAGCACGTGTGTTGAAATTTTTATCAATCAGTTGTTTGTATGTGATTCATTTGTAGC 1620
ATGGCTGATGTTGTGTGTCAAAAAAGTGGTGAGGATTCAGGAATTTCTGTGCTAATTGTG 1680
TGCAGATTGACTTCAGAAACAAAACCAAAAACAAAAACAAAACAAGGAAATGTAAAAATA 1740
TATATTAATGTATGTGAAAGACATTGAAGTGTTTGATACATGGCAGCATCTGTGGAGTAT 1800
GGGTGTGTAATGACCTTGTGGTTCAATGCAACTTTGACAGATATCTTTTTTTTTCTCTCT 1860
CTCTCTTTCCACTTTTCTCTATTTCCATCAAAAGAAAAAACATTGGAAGTGTGAGTGTAT 1920
GATTGATAGTCGCACGGTAAACTTGCACAGATACTTTTTTATTATTATTATTATTTCCTA 1980
AATTGTTTTATCACTTCCCTTACAAAGAAAATTAGTGATAAGAAATATGCAGCTGAATCT 2040
CCAACAAGTTGCATTTGTGAGAAGTCTCGGTGTGTAGCTTCAAATTATTGATATTTTACT 2100
TGCTGTCACTCAAACACCTTCTCATCTTTTTTTTTTTTTTTTCTTCTTTTTTTTTCTTTTAT 2160
TGTACTTTAATAAAGAATCCCCATAAGCTTGAAAGGACTTCTTTGTTTTAACTTCATGAA 2220
TATGGTTACTATAATTTTTGAACGTTTTATTATGTTCTCTCTTCATTCAGAAAGTATCTT 2280
TTAAGTTCTCACTTAAGTTTTGTATATTTACCACATGATGATTATGTGATATGTGTATTT 2340
TTTCTTACCTTTTTTGGGATTTTGTGTTCCTTGGAAATAAAAACAAAGATTGATAATTTG 2400
TAGAGAAAAAGAAGAGAGATGTAAAAAGCTTTCTTTACCTGTGTCAGTTACAAAAGAAAA 2460
TTACTGCGATATGTTTTGTGTTTCATTAAACGCACTTGAATATCATCAACTCGTTGGTAA 2520
ACTTTGGCCTCCAATGTGCTTGATGTGAGGTATGACTACAATGTGAAAACCCTTTGTAAC 2580
AAAGCTCAGGTCATCAAGTTGTCTCCTCCTGGTTCTTGCACTTGCTTTGTTAAGTTAGGG 2640
TTTTGTCAAACCAATTATGGACGGTAAAAAAAA                            2673
```

Figure 3.37 The full length cDNA and deduced protein sequences of the long form of *Tra-2* (*Tra-2l*, 2673 bp; ORF 804 bp, 267 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. The RNA recognition motif domain (RRM, positions 109th - 182th of the deduced prtein) is highlighted.

```
TACATGGTGGGAAAAGGCGGCTTCAACGCGTGAGGTTCCCGAATTTTTGCGATTTTTTGG 60
CTTTTTCTACTTAAAATGGAGAGTCCAGAGGGTGAACGGAACAGTCTCACACCTCGTTCG 120
               M  E  S  P  E  G  E  R  N  S  L  T  P  R  S   15
CGGTCGCGCTCAAGGTCACGTCTAGAGTCACCAGCCGCTTCCCCTGCTCATCGCCGCACA 180
R  S  R  S  R  S  R  L  E  S  P  A  A  S  P  A  H  R  R  T   35
GCCACTTCGCAGTCAAGGTCTCCCCAGCCTCGCAGACGCTCATTTTCCAGGTCGCGATCC 240
A  T  S  Q  S  R  S  P  Q  P  R  R  R  S  F  S  R  S  R  S   55
CGAACTCCACGCAGGCATCGCTCCCGGAGTGGCTCACCTCGAAATGGTCATGATGGAAGC 300
R  T  P  R  R  H  R  S  R  S  G  S  P  R  N  G  H  D  G  S   75
AGTCGCCGCTCCCGACGCAGCCGCTCCCCGATGTCGAACCGGAGGCGGCATATTGGCACC 360
S  R  R  S  R  R  S  R  S  P  M  S  N  R  R  R  H  I  G  T   95
AGAGACAATCCCGAGCCATCTAAGTGCCTGGGGGTGTTCGGCCTGAGTGTGCATACCACG 420
R  D  N  P  E  P  S  K  C  L  G  V  F  G  L  S  V  H  T  T   115
GAAAGACAGCTGTACACCATTTTTGACAAGTTTGGCCCTCTGGAGAAAGTACAAGTAGTC 480
E  R  Q  L  Y  T  I  F  D  K  F  G  P  L  E  K  V  Q  V  V   135
CTGGATTCGAAAACGGGCAAGTCAAGAGGCTTTGCGTTTGTGTACTTTGAGTCACTTAAG 540
L  D  S  K  T  G  K  S  R  G  F  A  F  V  Y  F  E  S  L  K   155
GATGCCTCAGAAGCCAAAAATGAGTGTTCTGGGATGGAGATTGACGGCCGGAGGATTAGA 600
D  A  S  E  A  K  N  E  C  S  G  M  E  I  D  G  R  R  I  R   175
GTTGATTATTCCATTACCAAGCGGCCACACACTCCAACTCCAGGAATATACATGGGTAGA 660
V  D  Y  S  I  T  K  R  P  H  T  P  T  P  G  I  Y  M  G  R   195
CCCACATCTCGCGGTGGCTACGACCGGGGCTACGGCCGAGGAGGCCACCGCGGAGACAGG 720
P  T  S  R  G  G  Y  D  R  G  Y  G  R  G  G  H  R  G  D  R   215
TACCGTTCCCCGTCACCCCGTTACCGGCCTCGCTCCAGCGGTGGTCGCCGCGACTATTAC 780
Y  R  S  P  S  P  R  Y  R  P  R  S  S  G  G  R  R  D  Y  Y   235
GATCGTGGGTATGTACTCGGCCAGGGCGTTACCGCACAACTCAAGCCACCACAAGATATA 840
D  R  G  Y  V  L  G  Q  G  V  T  A  Q  L  K  P  P  Q  D  I   255
ATAGTTCATGTGGAGGCAAGCTAGTTAAAGCAAGTTTACAGAAGCTTCTTAAGTAGGCCA 900
I  V  H  V  E  A  S  *                                       262
AGGGGCTGAAGATGATCAACTCCCAAGTTAATGGTGAAGGTGGTTAAGAATTGATGGATT 960
GATGGAAAGGATTCACGGTTTGTTCAGAGCAACCCTCCAAAGGCAATTGACAAGTGCAGA 1020
TGAAAAAACAGTAAAATTGTGAGGAAAAACCTGTGAACTGAAGATTTCTAAGAGATTTGA 1080
GGTATGACCGCGGGGACAGGAGCTACGACAGAGGCTATGACCGCTACGACAGGCCTCAAT 1140
ACCATGACCGCTATGATAGGTATGACCGTGCTTATGACAAGTACGATCGCTACGATAGGT 1200
CCAGATCTCGCTCCTATTCTCCACGAAGATACAAGTATTGATGATTGTACACCAGACAAA 1260
ACTGCAAGATTACTACAGAAAGGTTGTAGATGCTCCATACTTGACGGTGCTGTGCAGACG 1320
GCTGGCTCCAGAGAACGTTTCGACCGTTATACAGTTACAAATTCTCAGTCTCATTTGTCA 1380
TGCTAGTTATCAAAATATTTTAAGATAAGCTGTAGATAACAAGGTCTAAGAAAGAAAAAC 1440
ACTTTTTTTTTTTTCTATTTTTTCTTTCTTTTTTCTGTCTTTTTTCTGTCTTTTCTTATA 1500
GGTAATGTACTCTCCTCCACCCATATCCCCTTCCCTTCTTCCCGAGTCAAGCACGTGTGT 1560
TGAAATTTTTATCAATCAGTTGTTTGTATGTGATTCATTTGTAGCATGGCTGATGTTGTG 1620
TGTCAAAAAAGTGGTGAGGATTCAGGAATTTCTGTGCTAATTGTGTGCAGATTGACTTCA 1680
GAAACAAAACCAAAAACAAAAACAAAACAAGGAAATGTAAAAATATATATTAATGTATGT 1740
GAAAGACATTGAAGTGTTTGATACATGGCAGCATCTGTGGAGTATGGGTGTGTAATGACC 1800
TTGTGGTTCAATGCAACTTTGACAGATATCTTTTTTTTTCTCTCTCTCTCTTTCCACTTT 1860
TCTCTATTTCCATCAAAAGAAAAAACATTGGAAGTGTGAGTGTATGATTGATAGTCGCAC 1920
GGTAAACTTGCACAGATACTTTTTTATTATTATTATTATTTCCTAAATTGTTTTATCACT 1980
TCCCTTACAAAGAAAATTAGTGATAAGAAATATGCAGCTGAATCTCCAACAAGTTGCATT 2040
TGTGAGAAGTCTCGGTGTGTAGCTTCAAATTATTGATATTTTACTTGCTGTCACTCAAAC 2100
ACCTTCTCATCTTTTTTTTTTTTTTTTCTTCTTTTTTTTTCTTTTATTGTACTTTAATAAAG 2160
AATCCCCATAAGCTTGAAAGGACTTCTTTGTTTTAACTTCATGAATATGGTTACTATAAT 2220
TTTTGAACGTTTTATTATGTTCTCTCTTCATTCAGAAAGTATCTTTTAAGTTCTCACTTA 2280
AGTTTTGTATATTTACCACATGATGATTATGTGATATGTGTATTTTTTCTTACCTTTTTT 2340
GGGATTTTGTGTTCCTTGGAAATAAAAACAAAGATTGATAATTTGTAGAGAAAAAGAAGA 2400
GAGATGTAAAAAGCTTTCTTTACCTGTGTCAGTTACAAAAGAAAATTACTGCGATATGTT 2460
TTGTGTTTCATTAAACGCACTTGAATATCATCAACTCGTTGGTAAACTTTGGCCTCCAAT 2520
GTGCTTGATGTGAGGTATGACTACAATGTGAAAACCCTTTGTAACAAAGCTCAGGTCATC 2580
AAGTTGTCTCCTCCTGGTTCTTGCACTTGCTTTGTTAAGTTAGGGTTTTGTCAAACCAAT 2640
TATGGACGGTAAAAAAAA                                           2658
```

Figure 3.38 The full length cDNA and deduced protein sequences of the short form of *Tra-2* (*Tra-2s*, 2658 bp; ORF 789 bp, 262 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined obtained from EST. The RNA recognition motif domain (RRM, positions and 104th-177th of the deduced prtein) are highlighted.

```
Tra-2l      MESPEGERNSLTPRSRSRSRSRLESPAASPAHRRTATSQSRSPQPRRRSFSRSRSRTPRR 60
Tra-2s      MESPEGERNSLTPRSRSRSRSRLESPAASPAHRRTATSQSRSPQPRRRSFSRSRSRTPRR 60
            ************************************************************

Tra-2l      HRSRSGSPRNGHDGSSRRSRRSRSVHSSSPMSNRRRHIGTRDNPEPSKCLGVFGLSVHTT 120
Tra-2s      HRSRSGSPRNGHDGSSRRSRRSR-----SPMSNRRRHIGTRDNPEPSKCLGVFGLSVHTT 115
            ***********************     ********************************

Tra-2l      ERQLYTIFDKFGPLEKVQVVLDSKTGKSRGFAFVYFESLKDASEAKNECSGMEIDGRRIR 180
Tra-2s      ERQLYTIFDKFGPLEKVQVVLDSKTGKSRGFAFVYFESLKDASEAKNECSGMEIDGRRIR 175
            ************************************************************

Tra-2l      VDYSITKRPHTPTPGIYMGRPTSRGGYDRGYGRGGHRGDRYRSPSPRYRPRSSGGRRDYY 240
Tra-2s      VDYSITKRPHTPTPGIYMGRPTSRGGYDRGYGRGGHRGDRYRSPSPRYRPRSSGGRRDYY 235
            ************************************************************

Tra-2l      DRGYVLGQGVTAQLKPPQDIIVHVEAS 267
Tra-2s      DRGYVLGQGVTAQLKPPQDIIVHVEAS 262
            ***************************
```

Figure 3.39 Amino acid alignments of two isoforms of *Tra-2* of *P. monodon*.

```
Tra-2s      TACATGGTGGGAAAAGGCGGCTTCAACGCGTGAGGTTCCCGAATTTTTGCGATTTTTTGG 60
Tra-2l      TACATGGTGGGAAAAGGCGGCTTCAACGCGTGAGGTTCCCGAATTTTTGCGATTTTTTGG 60
            ************************************************************
Tra-2s      CTTTTTCTACTTAAAATGGAGAGTCCAGAGGGTGAACGGAACAGTCTCACACCTCGTTCG 120
Tra-2l      CTTTTTCTACTTAAAATGGAGAGTCCAGAGGGTGAACGGAACAGTCTCACACCTCGTTCG 120
            ************************************************************
Tra-2s      CGGTCGCGCTCAAGGTCACGTCTAGAGTCACCAGCCGCTTCCCCTGCTCATCGCCGCACA 180
Tra-2l      CGGTCGCGCTCAAGGTCACGTCTAGAGTCACCAGCCGCTTCCCCTGCTCATCGCCGCACA 180
            ************************************************************
Tra-2s      GCCACTTCGCAGTCAAGGTCTCCCCAGCCTCGCAGACGCTCATTTTCCAGGTCGCGATCC 240
Tra-2l      GCCACTTCGCAGTCAAGGTCTCCCCAGCCTCGCAGACGCTCATTTTCCAGGTCGCGATCC 240
            ************************************************************
Tra-2s      CGAACTCCACGCAGGCATCGCTCCCGGAGTGGCTCACCTCGAAATGGTCATGATGGAAGC 300
Tra-2l      CGAACTCCACGCAGGCATCGCTCCCGGAGTGGCTCACCTCGAAATGGTCATGATGGAAGC 300
            ************************************************************
Tra-2s      AGTCGCCGCTCCCGACGCAGCCG---------------CTCCCCGATGTCGAACCGGAGG 345
Tra-2l      AGTCGCCGCTCCCGACGCAGCCGGTCGGTCCACTCAAGCTCCCCGATGTCGAACCGGAGG 360
            ***********************               **********************
Tra-2s      CGGCATATTGGCACCAGAGACAATCCCGAGCCATCTAAGTGCCTGGGGGTGTTCGGCCTG 405
Tra-2l      CGGCATATTGGCACCAGAGACAATCCCGAGCCATCTAAGTGCCTGGGGGTGTTCGGCCTG 420
            ************************************************************
Tra-2s      AGTGTGCATACCACGGAAAGACAGCTGTACACCATTTTTGACAAGTTTGGCCCTCTGGAG 465
Tra-2l      AGTGTGCATACCACGGAAAGACAGCTGTACACCATTTTTGACAAGTTTGGCCCTCTGGAG 480
            ************************************************************
Tra-2s      AAAGTACAAGTAGTCCTGGATTCGAAAACGGGCAAGTCAAGAGGCTTTGCGTTTGTGTAC 525
Tra-2l      AAAGTACAAGTAGTCCTGGATTCGAAAACGGGCAAGTCAAGAGGCTTTGCGTTTGTGTAC 540
            ************************************************************
Tra-2s      TTTGAGTCACTTAAGGATGCCTCAGAAGCCAAAAATGAGTGTTCTGGGATGGAGATTGAC 585
Tra-2l      TTTGAGTCACTTAAGGATGCCTCAGAAGCCAAAAATGAGTGTTCTGGGATGGAGATTGAC 600
            ************************************************************
Tra-2s      GGCCGGAGGATTAGAGTTGATTATTCCATTACCAAGCGGCCACACACTCCAACTCCAGGA 645
Tra-2l      GGCCGGAGGATTAGAGTTGATTATTCCATTACCAAGCGGCCACACACTCCAACTCCAGGA 660
            ************************************************************
Tra-2s      ATATACATGGGTAGACCCACATCTCGCGGTGGCTACGACCGGGGCTACGGCCGAGGAGGC 705
Tra-2l      ATATACATGGGTAGACCCACATCTCGCGGTGGCTACGACCGGGGCTACGGCCGAGGAGGC 720
            ************************************************************
Tra-2s      CACCGCGGAGACAGGTACCGTTCCCCGTCACCCCGTTACCGGCCTCGCTCCAGCGGTGGT 765
Tra-2l      CACCGCGGAGACAGGTACCGTTCCCCGTCACCCCGTTACCGGCCTCGCTCCAGCGGTGGT 780
            ************************************************************
Tra-2s      CGCCGCGACTATTACGATCGTGGGTATGTACTCGGCCAGGGCGTTACCGCACAACTCAAG 825
Tra-2l      CGCCGCGACTATTACGATCGTGGGTATGTACTCGGCCAGGGCGTTACCGCACAACTCAAG 840
            ************************************************************
Tra-2s      CCACCACAAGATATAATAGTTCATGTGGAGGCAAGCTAGTTAAAGCAAGTTTACAGAAGC 885
Tra-2l      CCACCACAAGATATAATAGTTCATGTGGAGGCAAGCTAGTTAAAGCAAGTTTACAGAAGC 900
            ************************************************************
Tra-2s      TTCTTAAGTAGGCCAAGGGGCTGAAGATGATCAACTCCCAAGTTAATGGTGAAGGTGGTT 945
Tra-2l      TTCTTAAGTAGGCCAAGGGGCTGAAGATGATCAACTCCCAAGTTAATGGTGAAGGTGGTT 960
            ************************************************************
```

```
Tra-2s        AAGAATTGATGGATTGATGGAAAGGATTCACGGTTTGTTCAGAGCAACCCTCCAAAGGCA 1005
Tra-2l        AAGAATTGATGGATTGATGGAAAGGATTCACGGTTTGTTCAGAGCAACCCTCCAAAGGCA 1020
              ************************************************************

Tra-2s        ATTGACAAGTGCAGATGAAAAAACAGTAAAATTGTGAGGAAAAACCTGTGAACTGAAGAT 1065
Tra-2l        ATTGACAAGTGCAGATGAAAAAACAGTAAAATTGTGAGGAAAAACCTGTGAACTGAAGAT 1080
              ************************************************************

Tra-2s        TTCTAAGAGATTTGAGGTATGACCGCGGGGACAGGAGCTACGACAGAGGCTATGACCGCT 1125
Tra-2l        TTCTAAGAGATTTGAGGTATGACCGCGGGGACAGGAGCTACGACAGAGGCTATGACCGCT 1140
              ************************************************************

Tra-2s        ACGACAGGCCTCAATACCATGACCGCTATGATAGGTATGACCGTGCTTATGACAAGTACG 1185
Tra-2l        ACGACAGGCCTCAATACCATGACCGCTATGATAGGTATGACCGTGCTTATGACAAGTACG 1200
              ************************************************************

Tra-2s        ATCGCTACGATAGGTCCAGATCTCGCTCCTATTCTCCACGAAGATACAAGTATTGATGAT 1245
Tra-2l        ATCGCTACGATAGGTCCAGATCTCGCTCCTATTCTCCACGAAGATACAAGTATTGATGAT 1260
              ************************************************************

Tra-2s        TGTACACCAGACAAAACTGCAAGATTACTACAGAAAGGTTGTAGATGCTCCATACTTGAC 1305
Tra-2l        TGTACACCAGACAAAACTGCAAGATTACTACAGAAAGGTTGTAGATGCTCCATACTTGAC 1320
              ************************************************************

Tra-2s        GGTGCTGTGCAGACGGCTGGCTCCAGAGAACGTTTCGACCGTTATACAGTTACAAATTCT 1365
Tra-2l        GGTGCTGTGCAGACGGCTGGCTCCAGAGAACGTTTCGACCGTTATACAGTTACAAATTCT 1380
              ************************************************************

Tra-2s        CAGTCTCATTTGTCATGCTAGTTATCAAAATATTTTAAGATAAGCTGTAGATAACAAGGT 1425
Tra-2l        CAGTCTCATTTGTCATGCTAGTTATCAAAATATTTTAAGATAAGCTGTAGATAACAAGGT 1440
              ************************************************************

Tra-2s        CTAAGAAAGAAAAACACTTTTTTTTTTTTCTATTTTTTCTTTCTTTTTTCTGTCTTTTTT 1485
Tra-2l        CTAAGAAAGAAAAACACTTTTTTTTTTTTCTATTTTTTCTTTCTTTTTTCTGTCTTTTTT 1500
              ************************************************************

Tra-2s        CTGTCTTTTCTTATAGGTAATGTACTCTCCTCCACCCATATCCCCTTCCCTTCTTCCCGA 1545
Tra-2l        CTGTCTTTTCTTATAGGTAATGTACTCTCCTCCACCCATATCCCCTTCCCTTCTTCCCGA 1560
              ************************************************************

Tra-2s        GTCAAGCACGTGTGTTGAAATTTTTATCAATCAGTTGTTTGTATGTGATTCATTTGTAGC 1605
Tra-2l        GTCAAGCACGTGTGTTGAAATTTTTATCAATCAGTTGTTTGTATGTGATTCATTTGTAGC 1620
              ************************************************************

Tra-2s        ATGGCTGATGTTGTGTGTCAAAAAAGTGGTGAGGATTCAGGAATTTCTGTGCTAATTGTG 1665
Tra-2l        ATGGCTGATGTTGTGTGTCAAAAAAGTGGTGAGGATTCAGGAATTTCTGTGCTAATTGTG 1680
              ************************************************************

Tra-2s        TGCAGATTGACTTCAGAAACAAAACCAAAAACAAAAACAAAACAAGGAAATGTAAAAATA 1725
Tra-2l        TGCAGATTGACTTCAGAAACAAAACCAAAAACAAAAACAAAACAAGGAAATGTAAAAATA 1740
              ************************************************************

Tra-2s        TATATTAATGTATGTGAAAGACATTGAAGTGTTTGATACATGGCAGCATCTGTGGAGTAT 1785
Tra-2l        TATATTAATGTATGTGAAAGACATTGAAGTGTTTGATACATGGCAGCATCTGTGGAGTAT 1800
              ************************************************************

Tra-2s        GGGTGTGTAATGACCTTGTGGTTCAATGCAACTTTGACAGATATCTTTTTTTTTCTCTCT 1845
Tra-2l        GGGTGTGTAATGACCTTGTGGTTCAATGCAACTTTGACAGATATCTTTTTTTTTCTCTCT 1860
              ************************************************************

Tra-2s        CTCTCTTTCCACTTTTCTCTATTTCCATCAAAAGAAAAAACATTGGAAGTGTGAGTGTAT 1905
Tra-2l        CTCTCTTTCCACTTTTCTCTATTTCCATCAAAAGAAAAAACATTGGAAGTGTGAGTGTAT 1920
              ************************************************************

Tra-2s        GATTGATAGTCGCACGGTAAACTTGCACAGATACTTTTTTATTATTATTATTATTTCCTA 1965
Tra-2l        GATTGATAGTCGCACGGTAAACTTGCACAGATACTTTTTTATTATTATTATTATTTCCTA 1980
              ************************************************************

Tra-2s        AATTGTTTTATCACTTCCCTTACAAAGAAAATTAGTGATAAGAAATATGCAGCTGAATCT 2025
Tra-2l        AATTGTTTTATCACTTCCCTTACAAAGAAAATTAGTGATAAGAAATATGCAGCTGAATCT 2040
              ************************************************************

Tra-2s        CCAACAAGTTGCATTTGTGAGAAGTCTCGGTGTGTAGCTTCAAATTATTGATATTTTACT 2085
Tra-2l        CCAACAAGTTGCATTTGTGAGAAGTCTCGGTGTGTAGCTTCAAATTATTGATATTTTACT 2100
              ************************************************************

Tra-2s        TGCTGTCACTCAAACACCTTCTCATCTTTTTTTTTTTTTTTCTTCTTTTTTTTCTTTTAT 2145
Tra-2l        TGCTGTCACTCAAACACCTTCTCATCTTTTTTTTTTTTTTTCTTCTTTTTTTTCTTTTAT 2160
              ************************************************************

Tra-2s        TGTACTTTAATAAAGAATCCCCATAAGCTTGAAAGGACTTCTTTGTTTTAACTTCATGAA 2205
Tra-2l        TGTACTTTAATAAAGAATCCCCATAAGCTTGAAAGGACTTCTTTGTTTTAACTTCATGAA 2220
              ************************************************************

Tra-2s        TATGGTTACTATAATTTTTGAACGTTTTATTATGTTCTCTCTTCATTCAGAAAGTATCTT 2265
Tra-2l        TATGGTTACTATAATTTTTGAACGTTTTATTATGTTCTCTCTTCATTCAGAAAGTATCTT 2280
              ************************************************************

Tra-2s        TTAAGTTCTCACTTAAGTTTTGTATATTTACCACATGATGATTATGTGATATGTGTATTT 2325
Tra-2l        TTAAGTTCTCACTTAAGTTTTGTATATTTACCACATGATGATTATGTGATATGTGTATTT 2340
              ************************************************************

Tra-2s        TTTCTTACCTTTTTTGGGATTTTGTGTTCCTTGGAAATAAAAACAAAGATTGATAATTTG 2385
Tra-2l        TTTCTTACCTTTTTTGGGATTTTGTGTTCCTTGGAAATAAAAACAAAGATTGATAATTTG 2400
              ************************************************************

Tra-2s        TAGAGAAAAAGAAGAGAGATGTAAAAAGCTTTCTTACCTGTGTCAGTTACAAAAGAAAA 2445
Tra-2l        TAGAGAAAAAGAAGAGAGATGTAAAAAGCTTTCTTACCTGTGTCAGTTACAAAAGAAAA 2460
              ************************************************************
```

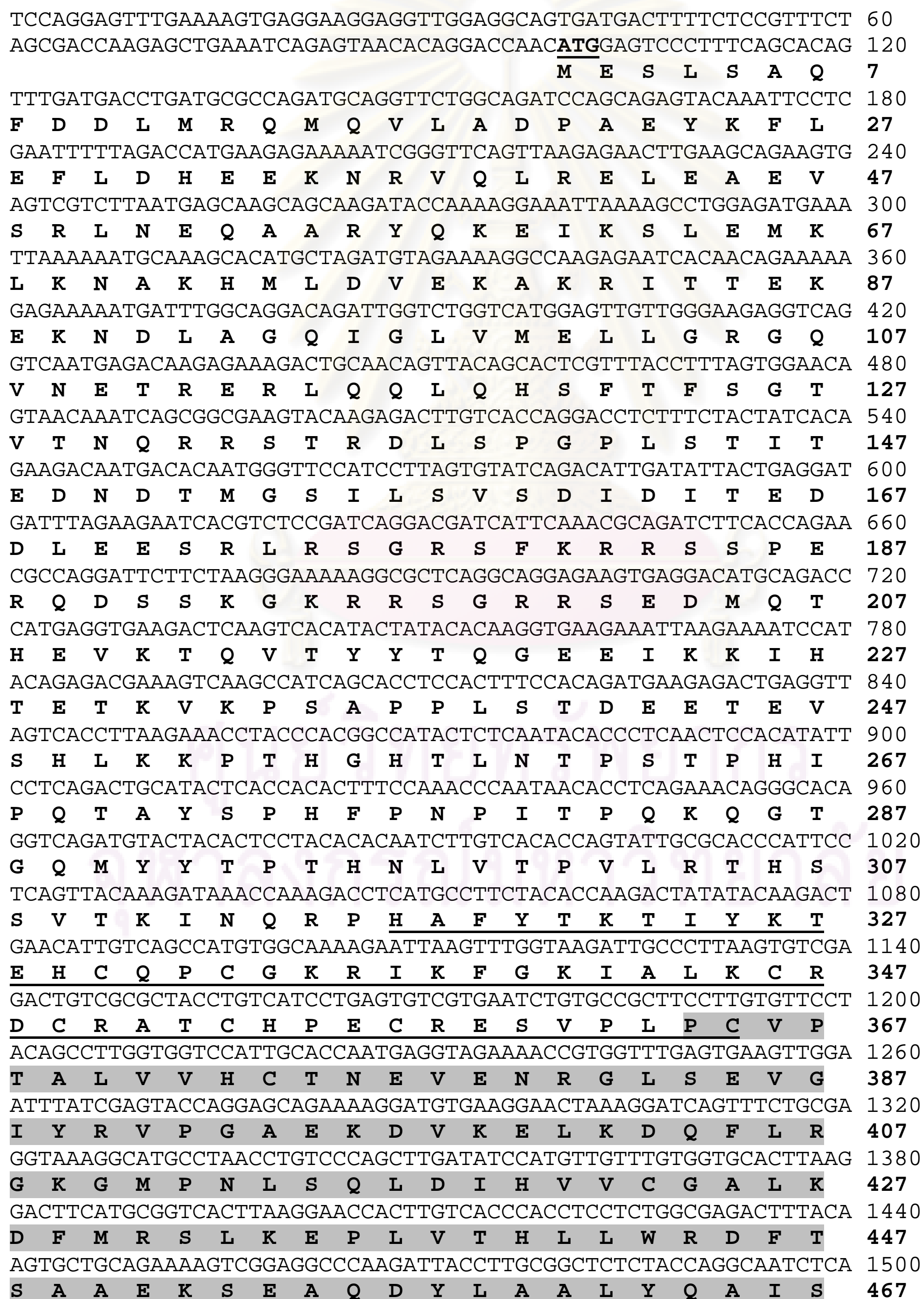
```
Tra-2s          TTACTGCGATATGTTTTGTGTTTCATTAAACGCACTTGAATATCATCAACTCGTTGGTAA 2505
Tra-2l          TTACTGCGATATGTTTTGTGTTTCATTAAACGCACTTGAATATCATCAACTCGTTGGTAA 2520
                ************************************************************

Tra-2s          ACTTTGGCCTCCAATGTGCTTGATGTGAGGTATGACTACAATGTGAAAACCCTTTGTAAC 2565
Tra-2l          ACTTTGGCCTCCAATGTGCTTGATGTGAGGTATGACTACAATGTGAAAACCCTTTGTAAC 2580
                ************************************************************

Tra-2s          AAAGCTCAGGTCATCAAGTTGTCTCCTCCTGGTTCTTGCACTTGCTTTGTTAAGTTAGGG 2625
Tra-2l          AAAGCTCAGGTCATCAAGTTGTCTCCTCCTGGTTCTTGCACTTGCTTTGTTAAGTTAGGG 2640
                ************************************************************

Tra-2s          TTTTGTCAAACCAATTATGGACGGTAAAAAAAA 2658
Tra-2l          TTTTGTCAAACCAATTATGGACGGTAAAAAAAA 2673
                *********************************
```

Figure 3.40 Pairwise alignments of different forms of *Tra-2* of *P. monodon*.



```
GAATTACCACAGCCCAACAGGGATACTTTGGCTTGGATCATGACTCATCTTCAAAGAGTA 1560
E  L  P  Q  P  N  R  D  T  L  A  W  I  M  T  H  L  Q  R  V   487
GCTGAATGTCCTGAATGCAAAATGCCGGCTAGCAACCTAGCCAAGGTGTTTGGGCCAACA 1620
A  E  C  P  E  C  K  M  P  A  S  N  L  A  K  V  F  G  P  T   507
CTTGTAGGATACTCAGTACCAGAACCTGATCCAGCCACTATGCTGACTGAAACTCGACAA 1680
L  V  G  Y  S  V  P  E  P  D  P  A  T  M  L  T  E  T  R  Q   527
CAGCAAATGGTCATGGAAAAGCTGCTTGAAATCTCCACAGACTACTGGAACACTTTCATT 1740
Q  Q  M  V  M  E  K  L  L  E  I  S  T  D  Y  W  N  T  F  I   547
AACGTTACTGATGAGAATGTGCACCAGGGAGTTCAGCAGGTTCCTACTCTAGAAGGTGGC 1800
N  V  T  D  E  N  V  H  Q  G  V  Q  Q  V  P  T  L  E  G  G   567
ACTCTCCTTGGAGGTTTCCCATCCTCCAACACGCGTCGACGCTCTATACTTACTCGCACT 1860
T  L  L  G  G  F  P  S  S  N  T  R  R  R  S  I  L  T  R  T   587
CCACTAACCCCCAGGGAAACTCCAAAGAACCGCTATGTCTTCCGGAAGTGAGGTTGCTGA 1920
P  L  T  P  R  E  T  P  K  N  R  Y  V  F  R  K  *            603
AATCTTTTTTTGTTTGATACAAATCAAGATGGACATAATATTATTTTATAAAATGTATTA 1980
CTAGGGCAATATGTGCCAAAAGAAATTGTATACTTTAACCATGCTTCTAGTTTAGTTAGC 2040
CTTTTGTGTATATATATTTCATTTATTTCATATTTGTATACATGTTTGTGCGTGTGCATG 2100
CGAGTTTACATTGTTTTGTGTGAGCATGTACTTATGTAACTGTACAGAACGAAAGATGAT 2160
TAGATGCTGATATAGTAGTGTGAAGTGGAGATATTGAAATTGTGGTACCAGGACATCAGC 2220
CATTCCTTTGAATCAGAATTGAGCCTATTAAGATGATTATGTTATACAAGTTTATACCAC 2280
TGTAGGAGAAGTTTATCAACTGATATTCTAGTTTTAAGGTTCTTCTAATTAAGTATTTTT 2340
GTAATAACTTTACATTTTGAAATATTGTCACAGTTGTACAGATTTTGATAATAACCTGAC 2400
TTAACCTTGTTACATCTGGTATGAATGCAATATCTAGTTTTTAGTTATTGCTGTGTTAGT 2460
CATGAAACAAAATACATTAAATATATACTACATTAATGCATGTAAATAATGGGTCAGGTC 2520
ATAACCAGACAAATTCTCTTAAACAGAAATGCTTTACTGTGAAGTTGATTAGAATCATTT 2580
GTGTTATTTTATTTTATTTATTTATTATTATTATTATTTTTTTTTTTTTTTATTGACTAT 2640
TTTTCATCTATCCACACCAAAGGAGCATTGTGAGATATGTATGTTTCTTTTATATTATTT 2700
TTAATGTGCATTATTAATACATTTCACATTACTCATAAATTCGCAGATAATCTATTATTT 2760
CTTGTATTTTTTAGGAAATTTCTCTTTTGAAGAAAATTTTTTATATATTGTGGAAAAAAA 2820
AAAAAAAATAAAAAAAAA                                           2838
```

Figure 3.41 The full length cDNA and deduced protein sequences of *Rac GTPase-activating protein 1* (2838 bp; ORF 1812 bp, 603 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A RhoGAP domain (positions 364th - 540th) and Protein kinase C conserved region 1 domains or Cysteine-rich domains (C1, positions 317th - 365th of the deduced amino acid, underlined) are highlighted.

```
CACGGTCTAGGTGTTGCTCGTCCTTCCCACGTTCGAAATACCTTAAAATCTGAGATAACG 60
AGTGCTTCCGATCTGTTACGCGAGAATCGGGGTATCGCGACGGTGTCTTGCGCAAGAGGT 120
CACGAGGGATAGATATGCTGTGAGTGTGAGATTCCAACTTCAGAGAAGCGCTGTTTGTAT 180
GAGGAGTCGCGGATTGGAAGAAGCGCAGGGCGTCGTCTGCATGTTTACATAGCGCAGTTT 240
GGGACGCGTTCCTCACCCTCCGATATTTGTAACCTCTTCGAGCGACAGGAGAACACGGCG 300
AAGACAAGCTAGACAACATGGGCAACATACACACCGTCGGACCAAACGAAGCTCTCGTGG 360
                 M  G  N  I  H  T  V  G  P  N  E  A  L  V  V 15
TGTCAGGTGGCTGCTGCGGCGCCACGTCCAAGAAGACCATCGTCGGCGGATGGGCGTGGG 420
  S  G  G  C  C  G  A  T  S  K  K  T  I  V  G  G  W  A  W  A 35
CGTGGTGGTTCGTGACCGACGTGCAGAGGCTCTCCCTCGAGGTGATCACGCTAAACCCAC 480
  W  W  F  V  T  D  V  Q  R  L  S  L  E  V  I  T  L  N  P  R 55
GATGCGAGAATGTAGAAACCTCCCTGGGTGTACAGGTGACAGTGACGGGCGTAGCACAGG 540
  C  E  N  V  E  T  S  L  G  V  Q  V  T  V  T  G  V  A  Q  V 75
TGAAGGTCATGAAGGAAGAGAAGGTGCTGAAGATTGCCTCAGAGCAGTTCCTGGGCATGT 600
  K  V  M  K  E  E  K  V  L  K  I  A  S  E  Q  F  L  G  M  S 95
CGAGCGATGAGATCAAGGGCACCATCCTCATGACACTCGAGGGCCATCTCAGGGCCATCT 660
  S  D  E  I  K  G  T  I  L  M  T  L  E  G  H  L  R  A  I  L 115
TGGCCACACTGACGGTAGAAGAGGTGTACCGTGACCGAGACCAGTTTGCATCCCTTGTGA 720
  A  T  L  T  V  E  E  V  Y  R  D  R  D  Q  F  A  S  L  V  R 135
GAGAAGTGGCTGGAATGGATGTTGGAAGAATGGGTATTGAGATTCTGTCCTTTACCATCA 780
  E  V  A  G  M  D  V  G  R  M  G  I  E  I  L  S  F  T  I  K 155
```

```
AGGATGTGTATGATCGTGTTGATTACTTGGCATCATTGGGCAAGTCCCAGACAGCTGCAG 840
  D  V  Y  D  R  V  D  Y  L  A  S  L  G  K  S  Q  T  A  A  V 175
TAAAGAGGGATGCTGACATCGGTGTGGCTCAGGCAAACAGAGATGCAGGAATTAGGGAAG 900
  K  R  D  A  D  I  G  V  A  Q  A  N  R  D  A  G  I  R  E  A 195
CTGAAGCAGAGAAAGCTGCTATGGATGTGAAGTATAACACAGACACCAAGATTGAAGACA 960
  E  A  E  K  A  A  M  D  V  K  Y  N  T  D  T  K  I  E  D  N 215
ACGCTCGACTGTACAAGTTGCAGAAGGCTCAGTTTGACCGTGAGATCAATACTGCGAAAG 1020
  A  R  L  Y  K  L  Q  K  A  Q  F  D  R  E  I  N  T  A  K  A 235
CTGAGGCACAGCTGGCATATGAACTCCAGGCAGCAAAGACACAACAGAAAATCCGAAACG 1080
  E  A  Q  L  A  Y  E  L  Q  A  A  K  T  Q  Q  K  I  R  N  E 255
AAGAAATTGCGATTGAAGTTGTGGAGAGGAGAAAGCAGATTGAGGTTGAAGAACAAGAAA 1140
  E  I  A  I  E  V  V  E  R  R  K  Q  I  E  V  E  E  Q  E  I 275
TCAAGCGAAAAGAGAAGGAGCTGATGGCAACCATCCGTCTCCCAGCTGAAGCAGAAAGCT 1200
  K  R  K  E  K  E  L  M  A  T  I  R  L  P  A  E  A  E  S  Y 295
ATCGTGTTGAAACCATCGCTCAGGGTCGTCGTACTCAGACAGTGGAAGCTGCCAGAGCTG 1260
  R  V  E  T  I  A  Q  G  R  R  T  Q  T  V  E  A  A  R  A  E 315
AAGCCGAACGTATTAAGCGCATTGGAGAAGCTGAGGCTTATGCTGTTGAAGCTGTGGGTA 1320
  A  E  R  I  K  R  I  G  E  A  E  A  Y  A  V  E  A  V  G  K 335
AGGCAGACGCAGAGAGTATGAAGTTGAAGGCAGTAGCCCTCAAGCAGTATGGTGATGCAG 1380
  A  D  A  E  S  M  K  L  K  A  V  A  L  K  Q  Y  G  D  A  A 355
CTATTACAGCTATGGTGCTTGAGAGCCTGCCACAGATTGCTGCTGAAGTTGCTGCTCCTC 1440
  I  T  A  M  V  L  E  S  L  P  Q  I  A  A  E  V  A  A  P  L 375
TAGCCAAGACTGAGGAGAGCGTGATGGTTGGTGGAGGAGACACAGTGAGCAATGCCATCA 1500
  A  K  T  E  E  S  V  M  V  G  G  G  D  T  V  S  N  A  I  N 395
ACAAGATGTGCAGTGAATTGCCACCTGCTGTCCATGCTCTGACTGGTGTTGATCTTACAA 1560
  K  M  C  S  E  L  P  P  A  V  H  A  L  T  G  V  D  L  T  K 415
AGGTGCTAAGCAAAGTACCAGGTGCAACTGTCACCCAACCCACAGCAGCTCCACGTGTGG 1620
  V  L  S  K  V  P  G  A  T  V  T  Q  P  T  A  A  P  R  V  A 435
CCCCAACAGCCGTGTAAAGCTCCACTTGACAGGGATATCTATGTATCCAGTGGAAGGACA 1680
  P  T  A  V  *                                              439
CCTAGAGAACTACTATTAGAAATAAGATATAACTTTCTCCATAGATGTAACAAGTTAAAC 1740
AGCACAGTTTTATTGTGAAATGGTTATTAACAGGATACTAGAGTAAAGTCAAATTCATTG 1800
ATACTGAGTAGTTAATTTGCAAAGAAAGAGACAGTATGCATCATATAAATGTTTTTGTTT 1860
GTTCAGTATTTATTGATGGAAGTTAATTATTTTTTTTTCTAATCTTACAAAAAAAAAAA 1920
AAAAAAAAAAAAAAAAA                                            1937
```

Figure 3.42 The full length cDNA and deduced protein sequences of *flotillin 2* (1937 bp; ORF 1320 bp, 439 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A prohibitin domain (PHB, positions 87th - 268th of the deduced protein) is highlighted.

A homologue of *DMC1/LIM15 isoform 1* initially found in the forward SSH library was further characterized by RACE-PCR. The amplification products of 1000 bp and 600 bp were obtained from 5′ and 3′RACE-PCR, respectively (Fig. 3.43). The full length cDNA of *DMC1/LIM15 homolog isoform 1* was composed of 1661 bp with an ORF of 1026 bp in length encoding a polypeptide of 341 amino acids. The predicted p*I* and MW of the deduced DMC1/LIM15 isoform 1 protein were 5.35 and 37.54 kDa, respectively. The deduced DMC1/LIM15 homolog isoform 1 protein contained an AAA domain (positions 118th - 308th, E-value = 9.73e-06, Fig. 3.44). The closest similarity of this transcript was *meiotic recombination protein DMC1/LIM15 homolog isoform 1* of *Canis familiaris* (E-value = 1e-146).

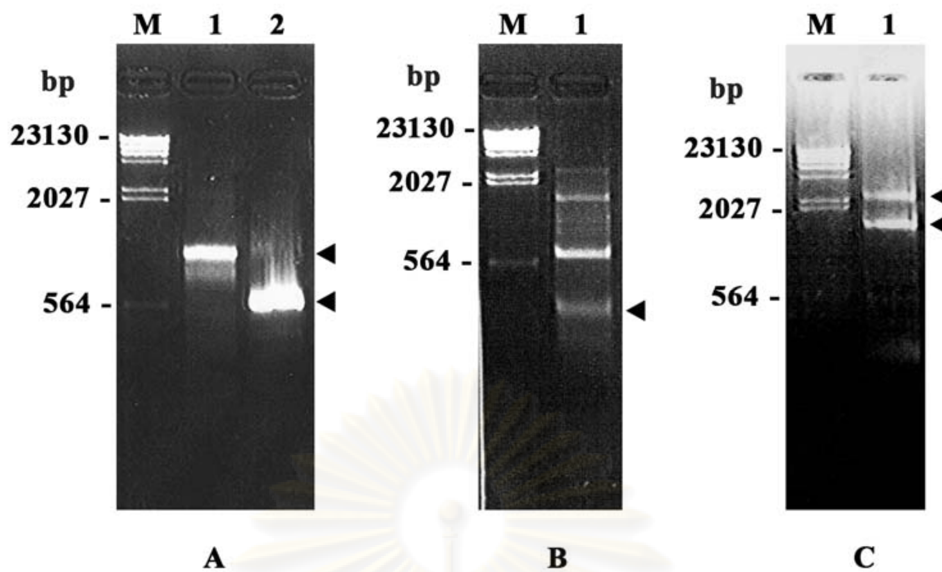


Figure 3.43 The RACE-PCR products of 5′- and 3′- *DMC1/LIM15 homolog isoform 1* (A), 5′- *PGMRC1* (B), and 3′- *PGMRC1* (C) of *P. monodon*. The λ *Hind* III (lane M) was used as the markers.

A homologue of *progestin receptor membrane component 1* (*PGMRC1*) initially found in the reverse SSH library was also further characterized. The full length cDNA of this gene homologue was successfully identified by RACE-PCR. The 5′ and 3′RACE-PCR products of 200 bp and 1500 and 2300 bp were obtained, respectively (Fig. 3.43). Three full length cDNAs of *PGMRC1*: *PGMRC1*-s (1980 bp), *PGMRC1-m* (2848 bp), and *PGMRC1-l* (2971 bp) were found in *P. monodon* (Fig. 3.45-3.47). These transcripts shared an identical ORF of 573 bp deducing to a 190 aa polypeptide but differed in length of the 3′UTR region. Neucleotide sequences of different isoforms of *PGMRC1* were multiple aligned and length polymorphism of the 3′UTR was observed (Fig. 3.48). The predicted p*I* and MW of the protein encoded by this cDNA were 4.60 and 20.98 kDa, respectively. A cytochrome b-5 like heme/steroid binding domain (Cyt-b5 domain; positions 68th- 166th, E-value = 1.3e-19) functionally important for ubiquitous electron transportation (Ozols, 1989) in heme-binding protein and progesterone receptor (Meyer *et al*., 1996) was found in the deduced PGMRC1. The closest similarity of these transcripts was *PGMRC1* of the medaka, *Oryzias latipes* (E-value = 1e-41, 2e-41, and 2e-41, respectively).

```
AAAAAAAAAAAATCCGACGGACAGGCGAAGACGTACATCTTTTCCGCCGCCGATTTTTTTT 60
TACAAAAAATAACCTCGACACACAACCTTCTCCTTCCGTTTAGAAGAACCGAAGAATCAC 120
AATGGAAGATCAGGCTTTAGATGCCGAGGAATGCATCCTGGACGATGAAATAAGCTTCTT 180
 M  E  D  Q  A  L  D  A  E  E  C  I  L  D  D  E  I  S  F  F  20
CACAGATATAGATGAATTACAAGCTCATGGCATCAACGCGGCGGATATTAAGAAGCTAAA 240
 T  D  I  D  E  L  Q  A  H  G  I  N  A  A  D  I  K  K  L  K  40
ATCTGCGGGAATTTGTACAGTCAAGGGAGTACAGATGATCACCCGACGTAGACTTTGCAT 300
 S  A  G  I  C  T  V  K  G  V  Q  M  I  T  R  R  R  L  C  M  60
GATCAAGGGAATCTCTGAGGCAAAAGTAGACAAGATAAAGGAGGTGGCAGCCAAGCTATG 360
 I  K  G  I  S  E  A  K  V  D  K  I  K  E  V  A  A  K  L  C  80
CGGAGGCGATGGGTTCGTGACGGCGCTGGTCATGTGCGAGAAGCGAAGGCTAGTGTTCCG 420
 G  G  D  G  F  V  T  A  L  V  M  C  E  K  R  R  L  V  F  R  100
AGTAAGCACGGGATCTGCTGAACTGGATGCACTGCTCGGCGGCGGGATCGAGAGCATGGC 480
 V  S  T  G  S  A  E  L  D  A  L  L  G  G  G  I  E  S  M  A  120
CATAACGGAGGTGTTCGGAGAGTTTCGTACGGGGAAGACACAAATCTCCCACACTCTCTG 540
 I  T  E  V  F  G  E  F  R  T  G  K  T  Q  I  S  H  T  L  C  140
CGTCACAGCACAGATTCCCAATGAAGCTGGGACGTACTCCGGAGGGAAGGTCATCTTCAT 600
 V  T  A  Q  I  P  N  E  A  G  T  Y  S  G  G  K  V  I  F  I  160
CGATACGGAGAATACCTTTCGGCCCGACCGACTGCGTCCCATTGCTGACCGCTACAACCT 660
 D  T  E  N  T  F  R  P  D  R  L  R  P  I  A  D  R  Y  N  L  180
CGAACAAGACGCAGTGCTCGACAACGTCCTGTACACCAGAGCCTTCACCTCGGAGCACCA 720
 E  Q  D  A  V  L  D  N  V  L  Y  T  R  A  F  T  S  E  H  Q  200
GCTGGAGATCCTGGACCACGTGGCCGCGCAGTTTCACGAAGAGCCTGGCATTTTCAAACT 780
 L  E  I  L  D  H  V  A  A  Q  F  H  E  E  P  G  I  F  K  L  220
GCTCATTGTCGACTCCGTTATGGCTCTCTTTCGGGTCGATTTCAGTGGACGCGGAGAGCT 840
 L  I  V  D  S  V  M  A  L  F  R  V  D  F  S  G  R  G  E  L  240
GGCTGACCGACAGCAGAAACTGGCGCAGTATATGTCACGACTGCAGAAGATCAGCGAGGA 900
 A  D  R  Q  Q  K  L  A  Q  Y  M  S  R  L  Q  K  I  S  E  E  260
ATACAACGTGTCTGTCTTCATCACTAATCAGATGACAGCAGATCCAGGGGCTGCGATGTC 960
 Y  N  V  S  V  F  I  T  N  Q  M  T  A  D  P  G  A  A  M  S  280
TTTCCAGGCGGACCCGAAGAAACCGATTGGCGGCCATATCCTAGCCCATGCCCCGACAAC 1020
 F  Q  A  D  P  K  K  P  I  G  G  H  I  L  A  H  A  P  T  T  300
AAGGGTATGTCTGCGCAAGGGGCGCGGCGAAACTCGTATTGCAAAGATCTACGACAGCCC 1080
 R  V  C  L  R  K  G  R  G  E  T  R  I  A  K  I  Y  D  S  P  320
CGAATTGCCTGAGAACGAGTGCACCTTCGCTATCACAGCAGGAGGCATTGCTGATGCTAA 1140
 E  L  P  E  N  E  C  T  F  A  I  T  A  G  G  I  A  D  A  K  340
GGAGTAACCATCTTCATATGCTGACGTGTTTATTTGTTTCAATTATTGTTTTTATACAGG 1200
 E  *                                                        341
TTTTACAATTTACTGTATTATGATTCAACTGAATTGAGAAAATTACTTCCAGGTAACCAT 1260
GTTTTGCTTGTATTTTTTGGTTAACAGTAATGCATATGTTAAACTATTTATTGCTTACAA 1320
ATACCCATATTTTAGTTTCAAAAGTGGAATTGTCTGTTTGATTACAAAGTAGTCTAGATA 1380
TTGAGATATCTTTATAGTTTTTGCAATAAGTTCATTGCTTGATTCTGGTCAGGCAATCAA 1440
TATTATGGATATCTGTGGTTGTACATCTTCATGGTGTATGAATAATCTTAGTATGCTACT 1500
CATTGGGGCATGATTGCATACATGCCCTTACACATTATGCTAGCATTGTATTAATTTGAT 1560
CACAGCACCCTTCTCTGACGTTTCATAACATCGTGGATATGATTGTATGTGGAAGTAAAG 1620
GGTTACTTCAGATGAAAAAAAAAAAAAAAAAAAAAAAAAAA                    1661
```

Figure 3.44 The full length cDNA and deduced protein sequences of *DMC1/LIM15 homolog isoform 1* (1661 bp; ORF 1026 bp, 341 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An AAA domain (positions 118th - 308th of the deduced protein) is highlighted.

```
TACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGACTCGACTATC 60
ATGGCGGACGAGGGAGCGGACGCCGTCTCCATCGAAGAGTCCTTCCTGGGCTCACTACTC 120
M  A  D  E  G  A  D  A  V  S  I  E  E  S  F  L  G  S  L  L   20
AAAGAAATATTCACCTCCCCACTTAATGTGTTCCTCTTGGGTGTCTGTACCGTCCTCATC 180
K  E  I  F  T  S  P  L  N  V  F  L  L  G  V  C  T  V  L  I   40
TATAAGATATTCCGTTCGTCCGATGGCAGTGGAGGAGCAACAGGTCCAGTGGAACCTCCT 240
Y  K  I  F  R  S  S  D  G  S  G  G  A  T  G  P  V  E  P  P   60
GTGCCCAAGATGAAACGACAGGACATGACCTTGGAGCAGTTGAAGCAGTATGATGGCATG 300
V  P  K  M  K  R  Q  D  M  T  L  E  Q  L  K  Q  Y  D  G  M   80
```

```
GGGGAGCATGGGCGTGTGTGTGCGGCAGTTAATGGCAAGATCTTTGATGTCACCCGAGGC 360
G  E  H  G  R  V  C  A  A  V  N  G  K  I  F  D  V  T  R  G  100
TCCAAGTTCTATGGCCCAGGTGGGCCGTATTCTGCCTTTGCTGGCCGAGATGCAACAAGA 420
S  K  F  Y  G  P  G  G  P  Y  S  A  F  A  G  R  D  A  T  R  120
GCTCTGGCAACCTTCAGTGTAAAGGATGTAAAGGAAGAGTACGATGACCTCAGTGACCTC 480
A  L  A  T  F  S  V  K  D  V  K  E  E  Y  D  D  L  S  D  L  140
TCCTCTATGCAGATGGACTCTGTCAGGGAATGGGAGATGCAGTTCACAGAAAAGTACGAT 540
S  S  M  Q  M  D  S  V  R  E  W  E  M  Q  F  T  E  K  Y  D  160
TATATTGGTAAATTTTTGAAACCAGGAGAACAGCCCACAGAGTACTCAGATGATGAGGAA 600
Y  I  G  K  F  L  K  P  G  E  Q  P  T  E  Y  S  D  D  E  E  180
GCAAAGGACACCAAAGCGAAGACGGATGATTAGATGTAGTTGAGGTGATTGCGCATTGCT 660
A  K  D  T  K  A  K  T  D  D  *                                190
GTATAGGTTAAGGCCTCTCGGTTCCACCAGACTCCAAAGCCCTTGAGCATGGTCTTAAGA 720
TTAGGATGTGGACGTGAAAAAAAGTAAAAAAAAAAAAAAAGAACCCCACTCAATTAGTCA 780
CTAATGATACGGTGTGATGGAAAAAGCCTACATTAGGTTGGGGGTTGGAGGTTTAAACTA 840
TATGTAAAACTACTACTTTATATTTTTTCTCATAAGGGGCTAATTAATCCCAAATATGTT 900
CTCAATAAAGATTGTCTACTTTGAACAATTTATCGATATGTGGTGACTTTGGTTAGTCTG 960
GGTGAGCCATGAAAGTTTGAGTAGGAGGAGCAGGAGGTGACAAGATCAGTCATTATCAGG 1020
CTTATTGGGGTATTTCATAAGGTATAATCTTGCAGTTAAAATGGAAAATAAAGTCTCTTA 1080
CAAAGGAGAGAGAAGGCTGATAGATATGCAGCTTTGTAGACCAATGCAAGCGACAAGTAT 1140
GTGTATACAGATTAATATAATTATAGAAGTGATTATTGAAGGATTGGGTCCCATTGAAAC 1200
ACAGCACCTACCAAACTTATCCTATTGTGTGATATATTTGTATAGATGGTTGAGATGTTG 1260
TTTTGTGTGGAATAAATGAATCATAGTAGTTTTGAAAATTGTTTTATGAGAATGATTGGA 1320
TATAGTTTATGAATGAGCAGCCCAAAGATGATGAGTTGGGAAGAGTGCAAGTGCAAGGAA 1380
TTCATCCTCAAATCAAACTTTCAGCCTTATAGAATACTGCAGAGGACTCATAATTGCTGG 1440
TCTGACTCAGAGTTATTTTGATACCTAACCTCTTGCCAGCATGGCATGATCCCCATCTTT 1500
TTCTAATCTGCCATGATTTATATTGTACTGTGGATACTCAGTGTGTGGATCCTTTATTCA 1560
GTCAATGTTTTAACATGTAAATATAGTGTATTCACCGTTGCCAAGTCCTGAAAAGACGTC 1620
CTCCAAATCTGCCTGCCTATCACGTTTGGGAATGGTAAATGACTTAGATATTGGAATGAG 1680
AGTGCAAGGGGATTATCTATTTTTTCTAGAAGTTTAGAGAGATAATGTTAACATGATTAC 1740
TCTGAACTTACTGGTTGAGTGTTTTCAGTACTCTTTTCTAACGAGTCTTTAGATGACATG 1800
TAGGTTCTGGCACAAACATGTGAAAGATGTATCTCGAGAGACAACAGGATCATAATGCTG 1860
CCTTGTTAACTGTTCTTCATCTTTAAGCAAGTAAGGCCTTTAGGTAGTGTCAGTCATTGT 1920
AAAGAGTTTGTTTTGAGAAAATGAAGGCATAAACACTAGGCTTAGTTGACTGGGGACTGT 1980
TCATCATTCAGAAATTTGTACAAAAAAAAAAAAAAAAAAAAAAAAAAA             2028
```

Figure 3.45 The full length cDNA and deduced protein sequences of a short form of *progestin receptor membrane component 1* (*PGMRC1-s*) (2028 bp; ORF 573 bp, 190 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A cytochrome b-5 like heme/steroid binding domain (Cyt-b5 domain; positions 68th-166th of the deduced protein) is highlighted.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
TACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGACTCGACTATC 60
ATGGCGGACGAGGGAGCGGACGCCGTCTCCATCGAAGAGTCCTTCCTGGGCTCACTACTC 120
M  A  D  E  G  A  D  A  V  S  I  E  E  S  F  L  G  S  L  L  20
AAAGAAATATTCACCTCCCCACTTAATGTGTTCCTCTTGGGTGTCTGTACCGTCCTCATC 180
K  E  I  F  T  S  P  L  N  V  F  L  L  G  V  C  T  V  L  I  40
TATAAGATATTCCGTTCGTCCGATGGCAGTGGAGGAGCAACAGGTCCAGTGGAACCTCCT 240
Y  K  I  F  R  S  S  D  G  S  G  G  A  T  G  P  V  E  P  P  60
GTGCCCAAGATGAAACGACAGGACATGACCTTGGAGCAGTTGAAGCAGTATGATGGCATG 300
V  P  K  M  K  R  Q  D  M  T  L  E  Q  L  K  Q  Y  D  G  M  80
GGGGAGCATGGGCGTGTGTGTGCGGCAGTTAATGGCAAGATCTTTGATGTCACCCGAGGC 360
G  E  H  G  R  V  C  A  A  V  N  G  K  I  F  D  V  T  R  G  100
TCCAAGTTCTATGGCCCAGGTGGGCCGTATTCTGCCTTTGCTGGCCGAGATGCAACAAGA 420
S  K  F  Y  G  P  G  G  P  Y  S  A  F  A  G  R  D  A  T  R  120
GCTCTGGCAACCTTCAGTGTAAAGGATGTAAAGGAAGAGTACGATGACCTCAGTGACCTC 480
A  L  A  T  F  S  V  K  D  V  K  E  E  Y  D  D  L  S  D  L  140
TCCTCTATGCAGATGGACTCTGTCAGGGAATGGGAGATGCAGTTCACAGAAAAGTACGAT 540
S  S  M  Q  M  D  S  V  R  E  W  E  M  Q  F  T  E  K  Y  D  160
TATATTGGTAAATTTTTGAAACCAGGAGAACAGCCCACAGAGTACTCAGATGATGAGGAA 600
```

```
Y   I   G   K   F   L   K   P   G   E   Q   P   T   E   Y   S   D   D   E   E    180
GCAAAGGACACCAAAGCGAAGACGGATGATTAGATGTAGTTGAGGTGATTGCGCATTGCT 660
A   K   D   T   K   A   K   T   D   D   *                                        190
GTATAGGTTAAGGCCTCTCGGTTCCACCAGACTCCAAAGCCCTTGAGCATGGTCTTAAGA 720
TTAGGATGTGGACGTGAAAAAAAGTAAAAAAAAAAAAAAAAGAACCCCACTCAATTAGTCA 780
CTAATGATACGGTGTGATGGAAAAAGCCTACATTAGGTTGGGGGTTGGAGGTTTAAACTA 840
TATGTAAAACTACTACTTTATATTTTTTCTCATAAGGGGCTAATTAATCCCAAATATGTT 900
CTCAATAAAGATTGTCTACTTTGAACAATTTATCGATATGTGGTGACTTTGGTTAGTCTG 960
GGTGAGCCATGAAAGTTTGAGTAGGAGGAGCAGGAGGTGACAAGATCAGTCATTATCAGG 1020
CTTATTGGGGTATTTCATAAGGTATAATCTTGCAGTTAAAATGGAAAATAAAGTCTCTTA 1080
CAAAGGAGAGAGAAGGCTGATAGATATGCAGCTTTGTAGACCAATGCAAGCAACAAGTAT 1140
GTGTATACAGATTAATATAATTATAGAAGTAATTATTGAAGGATTGGATCCCATTGAAAC 1200
ACAGCACCTACCAAACTTATCCTATTGTGTGATATATTTGTATAGATGGTTGAAATGTTG 1260
TTTTGTGTGGAATAAATGAATCATAGTAGTTTTGAAAATTGTTTTATGAGAATGATTGGA 1320
TATAGTTTATGAATGAGCAGCCCAAAGATGATGAGTTGGGAAGAGTGCAAGTGCAAGGAA 1380
TTCATCCTCAAATCAAACTTTCAGCCTTATAGAATACTGCAGAGGACTCATAATTGCTGG 1440
TCTGACTCAGAGTTATTTTGATACCTAACCTCTTGCCAGCATGGCATGATCCCCATCTTT 1500
TTCTAATCTACCATGATTTATATTGTACTGTGGATACTCAGTGTGTGGATCCTTTATTCA 1560
GTCAATGTTTTAACATGTAAATATAGTGTGTTCACCGTTGCCAAGTCCTGAAAAGACGTC 1620
CTCCAAATCTGCCTGCCTATCACGTTTGGGAATGGTAAATGACTTAGATATTGGAATGAG 1680
AGTGCAAGGGGATTATCTATTTTTTCTAGAAGTTTAGAGAGATAATGTTAACATGATTAC 1740
TCTGAACTTACTGGTTGAGTGTTTTCAGTACTCTTTTCTAACGAGTCTTTAGATGACATG 1800
TAGGTTCTGGCACAAACATGTGAAAGATGTATCTCGAGAGACAACAGGATCATAATGCTG 1860
CCTTGTTAACTGTTCTTCATCTTTAAGCAAGTAAGGCCTTTAGGTAGTGTCAGTCATTGT 1920
AAAGAGTTTGTTTTGAGAAAATGAAGGCATAAACACTAGGCTTAGTTGACTGGGGACTGT 1980
TCATCATTCAGAAATTTGTACAAAAAAAAAAAGTTATGACTTGCTCTCTTAAGTAAATTC 2040
CTTGGCAACTAAAAGAAAAGAGGTGTTTTTAATAAGAATAAGATGATTGGGCATATAGAT 2100
TTATATGCTTTGTTACCTCCAGCCAGTAGAGGTAAATAAGATTATGCTAACAGCTTCTAT 2160
GTTCAACAGGATTATATTTTGATGTTGTAGTTGATTCACCCTTATAAACGTATGAAGAAA 2220
TGTTCATTTTAAAGCTTACGAGTTTTCATTTCTTATAAAAACTGATAAACAGAAAGTTGA 2280
ATGAGTGCTTCTCCCATGGCTTGATGGTTGCAACACTAGATGTCATATGATCAAGGCTTT 2340
CCTTCTTTTCACACATCAATGTTTGATAAATGGCAGTTGTGAAAGGAAGATCCAGGAAGC 2400
TTCCACTAATGGCTTAAAAGCCTGATAAGTGAGTGTATTCTTAACAAAGGGAACTCCCGA 2460
GGCAGCTGTTGCAGTGCTGGTGGTGTAAGGTCCCTGGGGGATGTGATCTTCATTGATGAT 2520
CTAGATTTATCATTATAAGGTATACTACTTTGTTATGTTCATTTTTGTTATTTTTCATAT 2580
CCTTTATTTTCTATTTATTTATTTTCTTTTTTTTTCTTCCATATAGGGAGAATTTATATT 2640
TTGGACATCAGAGTTTCGTGAGCTGGAAATTCCTGTAATTGTGTGTGCATACCAGTTTTT 2700
GGCTAGCTATATCCAGCAATCTGTTCATTTGGCATTCCTGAAAGTTGCTCTTCAGGCTTT 2760
GCGTGAAGGAGTTGATTTAACTTTGTGATATTTAGAGAAGCAGAATTTGTATTTATATTT 2820
TTTACACATTGTCAATGTAGTTTGATAATATCCACATGGAACAAATGTTAAAAAAAAAAA 2880
AAAAAAAAAAAAAAAA                                             2896
```

Figure 3.46 The full length cDNA and deduced protein sequences of the medium form of *PGMRC1* (*PGMRC1-m*, 2896 bp; ORF 573 bp, 190 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A cytochrome b-5 like heme/steroid binding domain (Cyt-b5 domain; positions 68th - 166th of the deduced protein) is highlighted.

```
TACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGACTCGACTATC 60
ATGGCGGACGAGGGAGCGGACGCCGTCTCCATCGAAGAGTCCTTCCTGGGCTCACTACTC 120
M   A   D   E   G   A   D   A   V   S   I   E   E   S   F   L   G   S   L   L    20
AAAGAAATATTCACCTCCCCACTTAATGTGTTCCTCTTGGGTGTCTGTACCGTCCTCATC 180
K   E   I   F   T   S   P   L   N   V   F   L   L   G   V   C   T   V   L   I    40
TATAAGATATTCCGTTCGTCCGATGGCAGTGGAGGAGCAACAGGTCCAGTGGAACCTCCT 240
Y   K   I   F   R   S   S   D   G   S   G   G   A   T   G   P   V   E   P   P    60
GTGCCCAAGATGAAACGACAGGACATGACCTTGGAGCAGTTGAAGCAGTATGATGGCATG 300
V   P   K   M   K   R   Q   D   M   T   L   E   Q   L   K   Q   Y   D   G   M    80
GGGGAGCATGGGCGTGTGTGTGCGGCAGTTAATGGCAAGATCTTTGATGTCACCCGAGGC 360
G   E   H   G   R   V   C   A   A   V   N   G   K   I   F   D   V   T   R   G    100
TCCAAGTTCTATGGCCCAGGTGGGCCGTATTCTGCCTTTGCTGGCCGAGATGCAACAAGA 420
```

```
S   K   F   Y   G   P   G   G   P   Y   S   A   F   A   G   R   D   A   T   R    120
GCTCTGGCAACCTTCAGTGTAAAGGATGTAAAGGAAGAGTACGATGACCTCAGTGACCTC 480
A   L   A   T   F   S   V   K   D   V   K   E   E   Y   D   D   L   S   D   L    140
TCCTCTATGCAGATGGACTCTGTCAGGGAATGGGAGATGCAGTTCACAGAAAAGTACGAT 540
S   S   M   Q   M   D   S   V   R   E   W   E   M   Q   F   T   E   K   Y   D    160
TATATTGGTAAATTTTTGAAACCAGGAGAACAGCCCACAGAGTACTCAGATGATGAGGAA 600
Y   I   G   K   F   L   K   P   G   E   Q   P   T   E   Y   S   D   D   E   E    180
GCAAAGGACACCAAAGCGAAGACGGATGATTAGATGTAGTTGAGGTGATTGCGCATTGCT 660
A   K   D   T   K   A   K   T   D   D   *                                        190
GTATAGGTTAAGGCCTCTCGGTTCCACCAGACTCCAAAGCCCTTGAGCATGGTCTTAAGA 720
TTAGGATGTGGACGTGAAAAAAAGTAAAAAAAAAAAAAAAAGAACCCCACTCAATTAGTCA 780
CTAATGATACGGTGTGATGGAATAAGCCTACATTAGGTTGGGGGTTGGAGGTTTAAACTA 840
TATGTAAAACTACTACTTTATATTTTTTCTCATAAGGGGCTAATTAATCCCAAATATGTT 900
CTCAATAAAGATTGTCTACTTTGAACAATTTATCGATATGTGGTGACTTTGGTTAGTCTG 960
GGTGAGCCATGAAAGTTTGAGTAGGAGGAGCAGGAGGTGACAAGATCAGTCATTATCAGG 1020
CTTATTGGGGTATTTCATAAGGTATAATCTTGCAGTTAAAATGGAAAATAAAGTCTCTTA 1080
CAAAGGAGAGAGAAGGCTGATAGATATGCAGCTTTGTAGACCAATGCAAGCAACAAGTAT 1140
GTGTATACAGATTAATATAATTATAGAAGTAATTATTGAAGGATTGGATCCCATTGAAAC 1200
ACAGCACCTACCAAACTTATCCTATTGTGTGATATATTTGTATAGATGGTTGAAATGTTG 1260
TTTTGTGTGGAATAAATGAATCATAGTAGTTTTGAAAATTGTTTTATGAGAATGATTGGA 1320
TATAGTTTATGAATGAGCAGCCCAAAGATGATGAGTTGGGAAGAGTGCAAGTGCAAGGAA 1380
TTCATCCTCAAATCAAACTTTCAGCCTTATAGAATACTGCAGAGGACTCATAATTGCTGG 1440
TCTGACTCAGAGTTATTTTGATACCTAACCTCTTGCCAGCATGGCATGATCCCCATCTTT 1500
TTCTAATCTGCCATGATTTATATTGTACTGTGGATACTCAGTGTGTGGATCCTTTATTCA 1560
GTCAATGTTTTAACATGTAAATATAGTGTATTCACCGTTGCCAAGTCCTGAAAAGACGTC 1620
CTCCAAATCTGCCTGCCTATCACGTTTGGGAATGGTAAATGACTTAGATATTGGAATGAG 1680
AGTGCAAGGGGATTATCTATTTTTTCTAGAAGTTTAGAGAGATAATGTTAACATGATTAC 1740
TCTGAACTTACTGGTTGAGTGTTTTCAGTACTCTTTTCTAACGAGTCTTTAGATGACATG 1800
TAGGTTCTGGCACAAACATGTGAAAGATGTATCTCGAGAGACAACAGGATCATAATGCTG 1860
CCTTGTTAACTGTTCTTCATCTTTAAGCAAGTAAGGCCTTTAGGTAGTGTCAGTCATTGT 1920
AAAGAGTTTGTTTTGAGAAAATGAAGGCATAAACACTAGGCTTAGTTGACTGGGGACTGT 1980
TCATCATTCAGAAATTTGTACAAAAAAAAAAAAGTTATGACTTGCTCTCTTAAGTAAATTC 2040
CTTGGCAACTAAAAGAAAAGAGGTGTTTTTAATAAGAATAAGATGATTGGGCATATAGAT 2100
TTATATGCTTTGTTACCTCCAGCCAGTAGAGGTAAATAAGATTATGCTAACAGCTTCTAT 2160
GTTCAACAGGATTATATTTTGATGTTGTAGTTGATTCACCCTTATAAACGTATGAAGAAA 2220
TGTTCATTTTAAAGCTTACGAGTTTTCATTTCTTATAAAAACTGATAAACAGAAAGTTGA 2280
ATGAGTGCTTCTCCCATGGCTTGATGGTTGCAACACTAGATGTCATATGATCAAGGCTTT 2340
CCTTCTTTTCACACATCAATGTTTGATAAATGGCAGTTGTGAAAGGAAGATCAAGGAAGC 2400
TTCCACTAATGGCTTAAAAGCCTGATAAGTGAGTGTATTCTTAACAAAGGGAACTCCCGA 2460
GGCAGCTGTTGCAGTGCTGGTGGTGTAAGGTCCCTGGGGGATGTGATCTTCATTGATGAT 2520
CTAGATTTATCATTATAAGGTATACTACTTTGTTATGTTCTTTTTGTTATTTTTCATATC 2580
CTTTATTTTCTATTTATTTATTTTCTTTTTTTTTCTTCCATATAGGGAGAATTTATATTT 2640
TGGACATCAGAGTTTCGTGAGCTGGAAATTCCTGTAATTGTGTGTGCATACCAGTTTTTG 2700
GCTAGCTATATCCAGCAATCTGTTCATTTGGCATTCCTGAAAGTTGCTCTTCAGGCTTTG 2760
CGTGAAGGAGTTGATTTAACTTTGTGATATTTAGAGAAGCAGAATTTGTATTTATATTTT 2820
TTACACATTGTCAATGTAGTTTGATAATATCCACATGGAACAAATGTTAAAAAAGGAAAA 2880
AAAAGTCTGTGTAATAAAGGAAAACTTCTCCAATAGATGAAAGTTTTCATTTATGTACAG 2940
TTGAGTGTTAAATATTGTTCTAAATGAAATCAATAAATTTACCCAATAAAAAAAAAAAAA 3000
AAAAAAAAAAAAAAAAAAA                                          3019
```

Figure 3.47 The full length cDNA and deduced protein sequences of the long form of *PGMRC1* (*PGMRC1-l*, 3019 bp; ORF 573 bp, 190 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A cytochrome b-5 like heme/steroid binding domain (Cyt-b5 domain; positions 68th - 166th of the deduced protein) is highlighted.

```
PGMRC1-s     TACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGACTCGACTATC 60
PGMRC1-m     TACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGACTCGACTATC 60
PGMRC1-l     TACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGACTCGACTATC 60
             ************************************************************

PGMRC1-s     ATGGCGGACGAGGGAGCGGACGCCGTCTCCATCGAAGAGTCCTTCCTGGGCTCACTACTC 120
PGMRC1-m     ATGGCGGACGAGGGAGCGGACGCCGTCTCCATCGAAGAGTCCTTCCTGGGCTCACTACTC 120
PGMRC1-l     ATGGCGGACGAGGGAGCGGACGCCGTCTCCATCGAAGAGTCCTTCCTGGGCTCACTACTC 120
             ************************************************************

PGMRC1-s     AAAGAAATATTCACCTCCCCACTTAATGTGTTCCTCTTGGGTGTCTGTACCGTCCTCATC 180
PGMRC1-m     AAAGAAATATTCACCTCCCCACTTAATGTGTTCCTCTTGGGTGTCTGTACCGTCCTCATC 180
PGMRC1-l     AAAGAAATATTCACCTCCCCACTTAATGTGTTCCTCTTGGGTGTCTGTACCGTCCTCATC 180
             ************************************************************

PGMRC1-s     TATAAGATATTCCGTTCGTCCGATGGCAGTGGAGGAGCAACAGGTCCAGTGGAACCTCCT 240
PGMRC1-m     TATAAGATATTCCGTTCGTCCGATGGCAGTGGAGGAGCAACAGGTCCAGTGGAACCTCCT 240
PGMRC1-l     TATAAGATATTCCGTTCGTCCGATGGCAGTGGAGGAGCAACAGGTCCAGTGGAACCTCCT 240
             ************************************************************

PGMRC1-s     GTGCCCAAGATGAAACGACAGGACATGACCTTGGAGCAGTTGAAGCAGTATGATGGCATG 300
PGMRC1-m     GTGCCCAAGATGAAACGACAGGACATGACCTTGGAGCAGTTGAAGCAGTATGATGGCATG 300
PGMRC1-l     GTGCCCAAGATGAAACGACAGGACATGACCTTGGAGCAGTTGAAGCAGTATGATGGCATG 300
             ************************************************************

PGMRC1-s     GGGGAGCATGGGCGTGTGTGTGCGGCAGTTAATGGCAAGATCTTTGATGTCACCCGAGGC 360
PGMRC1-m     GGGGAGCATGGGCGTGTGTGTGCGGCAGTTAATGGCAAGATCTTTGATGTCACCCGAGGC 360
PGMRC1-l     GGGGAGCATGGGCGTGTGTGTGCGGCAGTTAATGGCAAGATCTTTGATGTCACCCGAGGC 360
             ************************************************************

PGMRC1-s     TCCAAGTTCTATGGCCCAGGTGGGCCGTATTCTGCCTTTGCTGGCCGAGATGCAACAAGA 420
PGMRC1-m     TCCAAGTTCTATGGCCCAGGTGGGCCGTATTCTGCCTTTGCTGGCCGAGATGCAACAAGA 420
PGMRC1-l     TCCAAGTTCTATGGCCCAGGTGGGCCGTATTCTGCCTTTGCTGGCCGAGATGCAACAAGA 420
             ************************************************************

PGMRC1-s     GCTCTGGCAACCTTCAGTGTAAAGGATGTAAAGGAAGAGTACGATGACCTCAGTGACCTC 480
PGMRC1-m     GCTCTGGCAACCTTCAGTGTAAAGGATGTAAAGGAAGAGTACGATGACCTCAGTGACCTC 480
PGMRC1-l     GCTCTGGCAACCTTCAGTGTAAAGGATGTAAAGGAAGAGTACGATGACCTCAGTGACCTC 480
             ************************************************************

PGMRC1-s     TCCTCTATGCAGATGGACTCTGTCAGGGAATGGGAGATGCAGTTCACAGAAAAGTACGAT 540
PGMRC1-m     TCCTCTATGCAGATGGACTCTGTCAGGGAATGGGAGATGCAGTTCACAGAAAAGTACGAT 540
PGMRC1-l     TCCTCTATGCAGATGGACTCTGTCAGGGAATGGGAGATGCAGTTCACAGAAAAGTACGAT 540
             ************************************************************

PGMRC1-s     TATATTGGTAAATTTTTGAAACCAGGAGAACAGCCCACAGAGTACTCAGATGATGAGGAA 600
PGMRC1-m     TATATTGGTAAATTTTTGAAACCAGGAGAACAGCCCACAGAGTACTCAGATGATGAGGAA 600
PGMRC1-l     TATATTGGTAAATTTTTGAAACCAGGAGAACAGCCCACAGAGTACTCAGATGATGAGGAA 600
             ************************************************************

PGMRC1-s     GCAAAGGACACCAAAGCGAAGACGGATGATTAGATGTAGTTGAGGTGATTGCGCATTGCT 660
PGMRC1-m     GCAAAGGACACCAAAGCGAAGACGGATGATTAGATGTAGTTGAGGTGATTGCGCATTGCT 660
PGMRC1-l     GCAAAGGACACCAAAGCGAAGACGGATGATTAGATGTAGTTGAGGTGATTGCGCATTGCT 660
             ************************************************************

PGMRC1-s     GTATAGGTTAAGGCCTCTCGGTTCCACCAGACTCCAAAGCCCTTGAGCATGGTCTTAAGA 720
PGMRC1-m     GTATAGGTTAAGGCCTCTCGGTTCCACCAGACTCCAAAGCCCTTGAGCATGGTCTTAAGA 720
PGMRC1-l     GTATAGGTTAAGGCCTCTCGGTTCCACCAGACTCCAAAGCCCTTGAGCATGGTCTTAAGA 720
             ************************************************************

PGMRC1-s     TTAGGATGTGGACGTGAAAAAAAGTAAAAAAAAAAAAAAAGAACCCCACTCAATTAGTCA 780
PGMRC1-m     TTAGGATGTGGACGTGAAAAAAAGTAAAAAAAAAAAAAAAGAACCCCACTCAATTAGTCA 780
PGMRC1-l     TTAGGATGTGGACGTGAAAAAAAGTAAAAAAAAAAAAAAAGAACCCCACTCAATTAGTCA 780
             ************************************************************

PGMRC1-s     CTAATGATACGGTGTGATGGAAAAAGCCTACATTAGGTTGGGGGTTGGAGGTTTAAACTA 840
PGMRC1-m     CTAATGATACGGTGTGATGGAAAAAGCCTACATTAGGTTGGGGGTTGGAGGTTTAAACTA 840
PGMRC1-l     CTAATGATACGGTGTGATGGAATAAGCCTACATTAGGTTGGGGGTTGGAGGTTTAAACTA 840
             ********************** *************************************

PGMRC1-s     TATGTAAAACTACTACTTTATATTTTTTCTCATAAGGGGCTAATTAATCCCAAATATGTT 900
PGMRC1-m     TATGTAAAACTACTACTTTATATTTTTTCTCATAAGGGGCTAATTAATCCCAAATATGTT 900
PGMRC1-l     TATGTAAAACTACTACTTTATATTTTTTCTCATAAGGGGCTAATTAATCCCAAATATGTT 900
             ************************************************************

PGMRC1-s     CTCAATAAAGATTGTCTACTTTGAACAATTTATCGATATGTGGTGACTTTGGTTAGTCTG 960
PGMRC1-m     CTCAATAAAGATTGTCTACTTTGAACAATTTATCGATATGTGGTGACTTTGGTTAGTCTG 960
PGMRC1-l     CTCAATAAAGATTGTCTACTTTGAACAATTTATCGATATGTGGTGACTTTGGTTAGTCTG 960
             ************************************************************

PGMRC1-s     GGTGAGCCATGAAAGTTTGAGTAGGAGGAGCAGGAGGTGACAAGATCAGTCATTATCAGG 1020
PGMRC1-m     GGTGAGCCATGAAAGTTTGAGTAGGAGGAGCAGGAGGTGACAAGATCAGTCATTATCAGG 1020
PGMRC1-l     GGTGAGCCATGAAAGTTTGAGTAGGAGGAGCAGGAGGTGACAAGATCAGTCATTATCAGG 1020
             ************************************************************

PGMRC1-s     CTTATTGGGGTATTTCATAAGGTATAATCTTGCAGTTAAAATGGAAAATAAAGTCTCTTA 1080
PGMRC1-m     CTTATTGGGGTATTTCATAAGGTATAATCTTGCAGTTAAAATGGAAAATAAAGTCTCTTA 1080
PGMRC1-l     CTTATTGGGGTATTTCATAAGGTATAATCTTGCAGTTAAAATGGAAAATAAAGTCTCTTA 1080
             ************************************************************

PGMRC1-s     CAAAGGAGAGAGAAGGCTGATAGATATGCAGCTTTGTAGACCAATGCAAGCGACAAGTAT 1140
PGMRC1-m     CAAAGGAGAGAGAAGGCTGATAGATATGCAGCTTTGTAGACCAATGCAAGCAACAAGTAT 1140
PGMRC1-l     CAAAGGAGAGAGAAGGCTGATAGATATGCAGCTTTGTAGACCAATGCAAGCAACAAGTAT 1140
             *************************************************** ********
```

```
PGMRC1-s      GTGTATACAGATTAATATAATTATAGAAGTGATTATTGAAGGATTGGGTCCCATTGAAAC 1200
PGMRC1-m      GTGTATACAGATTAATATAATTATAGAAGTAATTATTGAAGGATTGGATCCCATTGAAAC 1200
PGMRC1-l      GTGTATACAGATTAATATAATTATAGAAGTAATTATTGAAGGATTGGATCCCATTGAAAC 1200
              ****************************** **************** ************

PGMRC1-s      ACAGCACCTACCAAACTTATCCTATTGTGTGATATATTTGTATAGATGGTTGAGATGTTG 1260
PGMRC1-m      ACAGCACCTACCAAACTTATCCTATTGTGTGATATATTTGTATAGATGGTTGAAATGTTG 1260
PGMRC1-l      ACAGCACCTACCAAACTTATCCTATTGTGTGATATATTTGTATAGATGGTTGAAATGTTG 1260
              ***************************************************** ******

PGMRC1-s      TTTTGTGTGGAATAAATGAATCATAGTAGTTTTGAAAATTGTTTTATGAGAATGATTGGA 1320
PGMRC1-m      TTTTGTGTGGAATAAATGAATCATAGTAGTTTTGAAAATTGTTTTATGAGAATGATTGGA 1320
PGMRC1-l      TTTTGTGTGGAATAAATGAATCATAGTAGTTTTGAAAATTGTTTTATGAGAATGATTGGA 1320
              ************************************************************

PGMRC1-s      TATAGTTTATGAATGAGCAGCCCAAAGATGATGAGTTGGGAAGAGTGCAAGTGCAAGGAA 1380
PGMRC1-m      TATAGTTTATGAATGAGCAGCCCAAAGATGATGAGTTGGGAAGAGTGCAAGTGCAAGGAA 1380
PGMRC1-l      TATAGTTTATGAATGAGCAGCCCAAAGATGATGAGTTGGGAAGAGTGCAAGTGCAAGGAA 1380
              ************************************************************

PGMRC1-s      TTCATCCTCAAATCAAACTTTCAGCCTTATAGAATACTGCAGAGGACTCATAATTGCTGG 1440
PGMRC1-m      TTCATCCTCAAATCAAACTTTCAGCCTTATAGAATACTGCAGAGGACTCATAATTGCTGG 1440
PGMRC1-l      TTCATCCTCAAATCAAACTTTCAGCCTTATAGAATACTGCAGAGGACTCATAATTGCTGG 1440
              ************************************************************

PGMRC1-s      TCTGACTCAGAGTTATTTTGATACCTAACCTCTTGCCAGCATGGCATGATCCCCATCTTT 1500
PGMRC1-m      TCTGACTCAGAGTTATTTTGATACCTAACCTCTTGCCAGCATGGCATGATCCCCATCTTT 1500
PGMRC1-l      TCTGACTCAGAGTTATTTTGATACCTAACCTCTTGCCAGCATGGCATGATCCCCATCTTT 1500
              ************************************************************

PGMRC1-s      TTCTAATCTGCCATGATTTATATTGTACTGTGGATACTCAGTGTGTGGATCCTTTATTCA 1560
PGMRC1-m      TTCTAATCTACCATGATTTATATTGTACTGTGGATACTCAGTGTGTGGATCCTTTATTCA 1560
PGMRC1-l      TTCTAATCTGCCATGATTTATATTGTACTGTGGATACTCAGTGTGTGGATCCTTTATTCA 1560
              ********* **************************************************

PGMRC1-s      GTCAATGTTTTAACATGTAAATATAGTGTATTCACCGTTGCCAAGTCCTGAAAAGACGTC 1620
PGMRC1-m      GTCAATGTTTTAACATGTAAATATAGTGTGTTCACCGTTGCCAAGTCCTGAAAAGACGTC 1620
PGMRC1-l      GTCAATGTTTTAACATGTAAATATAGTGTATTCACCGTTGCCAAGTCCTGAAAAGACGTC 1620
              ***************************** ******************************

PGMRC1-s      CTCCAAATCTGCCTGCCTATCACGTTTGGGAATGGTAAATGACTTAGATATTGGAATGAG 1680
PGMRC1-m      CTCCAAATCTGCCTGCCTATCACGTTTGGGAATGGTAAATGACTTAGATATTGGAATGAG 1680
PGMRC1-l      CTCCAAATCTGCCTGCCTATCACGTTTGGGAATGGTAAATGACTTAGATATTGGAATGAG 1680
              ************************************************************

PGMRC1-s      AGTGCAAGGGGATTATCTATTTTTTCTAGAAGTTTAGAGAGATAATGTTAACATGATTAC 1740
PGMRC1-m      AGTGCAAGGGGATTATCTATTTTTTCTAGAAGTTTAGAGAGATAATGTTAACATGATTAC 1740
PGMRC1-l      AGTGCAAGGGGATTATCTATTTTTTCTAGAAGTTTAGAGAGATAATGTTAACATGATTAC 1740
              ************************************************************

PGMRC1-s      TCTGAACTTACTGGTTGAGTGTTTTCAGTACTCTTTTCTAACGAGTCTTTAGATGACATG 1800
PGMRC1-m      TCTGAACTTACTGGTTGAGTGTTTTCAGTACTCTTTTCTAACGAGTCTTTAGATGACATG 1800
PGMRC1-l      TCTGAACTTACTGGTTGAGTGTTTTCAGTACTCTTTTCTAACGAGTCTTTAGATGACATG 1800
              ************************************************************

PGMRC1-s      TAGGTTCTGGCACAAACATGTGAAAGATGTATCTCGAGAGACAACAGGATCATAATGCTG 1860
PGMRC1-m      TAGGTTCTGGCACAAACATGTGAAAGATGTATCTCGAGAGACAACAGGATCATAATGCTG 1860
PGMRC1-l      TAGGTTCTGGCACAAACATGTGAAAGATGTATCTCGAGAGACAACAGGATCATAATGCTG 1860
              ************************************************************

PGMRC1-s      CCTTGTTAACTGTTCTTCATCTTTAAGCAAGTAAGGCCTTTAGGTAGTGTCAGTCATTGT 1920
PGMRC1-m      CCTTGTTAACTGTTCTTCATCTTTAAGCAAGTAAGGCCTTTAGGTAGTGTCAGTCATTGT 1920
PGMRC1-l      CCTTGTTAACTGTTCTTCATCTTTAAGCAAGTAAGGCCTTTAGGTAGTGTCAGTCATTGT 1920
              ************************************************************

PGMRC1-s      AAAGAGTTTGTTTTGAGAAAATGAAGGCATAAACACTAGGCTTAGTTGACTGGGGACTGT 1980
PGMRC1-m      AAAGAGTTTGTTTTGAGAAAATGAAGGCATAAACACTAGGCTTAGTTGACTGGGGACTGT 1980
PGMRC1-l      AAAGAGTTTGTTTTGAGAAAATGAAGGCATAAACACTAGGCTTAGTTGACTGGGGACTGT 1980
              ************************************************************

PGMRC1-s      TCATCATTCAGAAATTTGTACAAAAAAAAAAAAAAAAAA---------AAAAAAAAA--- 2028
PGMRC1-m      TCATCATTCAGAAATTTGTACAAAAAAAAAAAGTTATGACTTGCTCTCTTAAGTAAATTC 2040
PGMRC1-l      TCATCATTCAGAAATTTGTACAAAAAAAAAAAGTTATGACTTGCTCTCTTAAGTAAATTC 2040
              ********************************   *  *           **  ***

PGMRC1-s      ------------------------------------------------------------
PGMRC1-m      CTTGGCAACTAAAAGAAAAGAGGTGTTTTTAATAAGAATAAGATGATTGGGCATATAGAT 2100
PGMRC1-l      CTTGGCAACTAAAAGAAAAGAGGTGTTTTTAATAAGAATAAGATGATTGGGCATATAGAT 2100

PGMRC1-s      ------------------------------------------------------------
PGMRC1-m      TTATATGCTTTGTTACCTCCAGCCAGTAGAGGTAAATAAGATTATGCTAACAGCTTCTAT 2160
PGMRC1-l      TTATATGCTTTGTTACCTCCAGCCAGTAGAGGTAAATAAGATTATGCTAACAGCTTCTAT 2160

PGMRC1-s      ------------------------------------------------------------
PGMRC1-m      GTTCAACAGGATTATATTTTGATGTTGTAGTTGATTCACCCTTATAAACGTATGAAGAAA 2220
PGMRC1-l      GTTCAACAGGATTATATTTTGATGTTGTAGTTGATTCACCCTTATAAACGTATGAAGAAA 2220

PGMRC1-s      ------------------------------------------------------------
PGMRC1-m      TGTTCATTTTAAAGCTTACGAGTTTTCATTTCTTATAAAAACTGATAAACAGAAAGTTGA 2280
PGMRC1-l      TGTTCATTTTAAAGCTTACGAGTTTTCATTTCTTATAAAAACTGATAAACAGAAAGTTGA 2280
```

```
PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        ATGAGTGCTTCTCCCATGGCTTGATGGTTGCAACACTAGATGTCATATGATCAAGGCTTT 2340
PGMRC1-l        ATGAGTGCTTCTCCCATGGCTTGATGGTTGCAACACTAGATGTCATATGATCAAGGCTTT 2340

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        CCTTCTTTTCACACATCAATGTTTGATAAATGGCAGTTGTGAAAGGAAGATCCAGGAAGC 2400
PGMRC1-l        CCTTCTTTTCACACATCAATGTTTGATAAATGGCAGTTGTGAAAGGAAGATCAAGGAAGC 2400

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        TTCCACTAATGGCTTAAAAGCCTGATAAGTGAGTGTATTCTTAACAAAGGGAACTCCCGA 2460
PGMRC1-l        TTCCACTAATGGCTTAAAAGCCTGATAAGTGAGTGTATTCTTAACAAAGGGAACTCCCGA 2460

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        GGCAGCTGTTGCAGTGCTGGTGGTGTAAGGTCCCTGGGGGATGTGATCTTCATTGATGAT 2520
PGMRC1-l        GGCAGCTGTTGCAGTGCTGGTGGTGTAAGGTCCCTGGGGGATGTGATCTTCATTGATGAT 2520

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        CTAGATTTATCATTATAAGGTATACTACTTTGTTATGTTCATTTTTGTTATTTTTCATAT 2580
PGMRC1-l        CTAGATTTATCATTATAAGGTATACTACTTTGTTATGTTC-TTTTTGTTATTTTTCATAT 2579

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        CCTTTATTTTCTATTTATTTATTTTCTTTTTTTTTCTTCCATATAGGGAGAATTTATATT 2640
PGMRC1-l        CCTTTATTTTCTATTTATTTATTTTCTTTTTTTTTCTTCCATATAGGGAGAATTTATATT 2639

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        TTGGACATCAGAGTTTCGTGAGCTGGAAATTCCTGTAATTGTGTGTGCATACCAGTTTTT 2700
PGMRC1-l        TTGGACATCAGAGTTTCGTGAGCTGGAAATTCCTGTAATTGTGTGTGCATACCAGTTTTT 2699

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        GGCTAGCTATATCCAGCAATCTGTTCATTTGGCATTCCTGAAAGTTGCTCTTCAGGCTTT 2760
PGMRC1-l        GGCTAGCTATATCCAGCAATCTGTTCATTTGGCATTCCTGAAAGTTGCTCTTCAGGCTTT 2759

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        GCGTGAAGGAGTTGATTTAACTTTGTGATATTTAGAGAAGCAGAATTTGTATTTATATTT 2820
PGMRC1-l        GCGTGAAGGAGTTGATTTAACTTTGTGATATTTAGAGAAGCAGAATTTGTATTTATATTT 2819

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        TTTACACATTGTCAATGTAGTTTGATAATATCCACATGGAACAAATGTTAAAAAAAAAAA 2880
PGMRC1-l        TTTACACATTGTCAATGTAGTTTGATAATATCCACATGGAACAAATGTTAAAAAAGGAAA 2879

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        AAAAA--------AAAAAAAAAAA------------------------------------ 2896
PGMRC1-l        AAAAAGTCTGTGTAATAAAGGAAAACTTCTCCAATAGATGAAAGTTTTCATTTATGTACA 2939

PGMRC1-s        ------------------------------------------------------------
PGMRC1-m        ------------------------------------------------------------
PGMRC1-l        GTTGAGTGTTAAATATTGTTCTAAATGAAATCAATAAATTTACCCAATAAAAAAAAAAAA 2999

PGMRC1-s        --------------------
PGMRC1-m        --------------------
PGMRC1-l        AAAAAAAAAAAAAAAAAAAA 3019
```

Figure 3.48 Multiple alignment of three different forms of *PGMRC1* (*PGMRC1-s*, *PGMRC1-m* and *PGMRC1-l*) of *P. monodon*.

Innexin1 and *innexin2* were initially found in the conventional and reverse SSH libraries, respectively. The full length cDNA of *innexin1* and *innexin2* of *P. monodon* were successfully characterized. The amplification products of 5′RACE-PCR of *innexin1* were 1000 bp whereas those of 5′ and 3′RACE products of *innexin2* were 1000 and 800 bp, respectively (Fig. 3.49). The full length cDNAs of *innexin1* was 2505 bp in length with an ORF of 1143 bp corresponding to a polypeptide of 380

amino acids (Fig. 3.50) while the full length of *innexin2* was 1651 bp in length with an ORF of 1077 bp corresponding to a protein of 358 amino acids (Fig. 3.51).

The closest similarity of these transcripts were *innexin1* of *Schistocerca americana* (E-value = 6e-120) and *innexin2* of *Homarus gammarus* (E-value = 7e-161), respectively. Deduced innexin1 and innexin2 proteins contained an innexin domain (positions 22nd -363rd, E-value = 3.30e-84 and positions 20th -358th, E-value = 9.90e-77, respectively). The expected MW and p*I* of *innexin1* were 44.0 kDa and 6.42, respectively and those of *innexin2* of *P. monodon* were 41.68 kDa and 6.37, respectively.

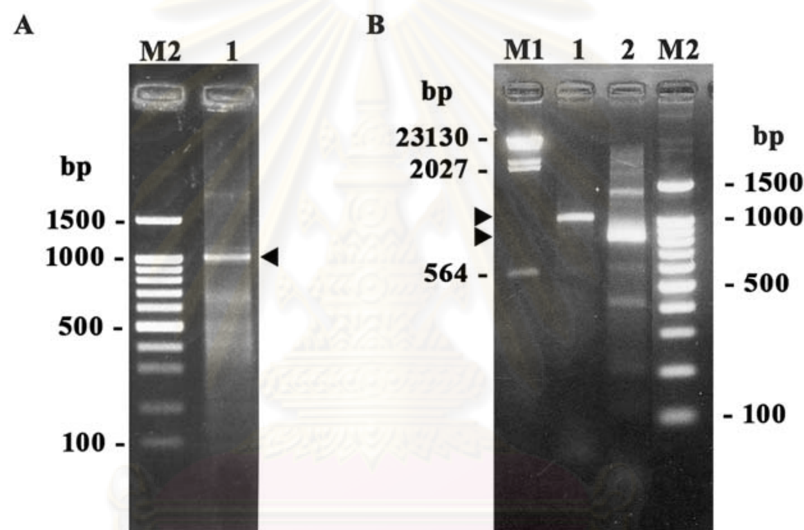


Figure 3.49 5′ RACE-PCR product of *innexin1* (A) and 5′ and 3′ RACE-PCR products *innexin2* (B) of *P. monodon.* A 100 bp marker (lanes M2) and λ-*Hind* III (lane M1) were used as the marker.

```
ACGCGGGGAGTATCCCCATGACCATCGAGCGGGACGATTCTGTGAGAGAGTGATTACGTG  60
TGTGACGGCTTGAACTCCGGGCACTATGTCTGCCATCAAGTACCTCTCGGGGCTGAAGTC  120
                         M  S  A  I  K  Y  L  S  G  L  K  S   12
CTATCTGGCCCGTGAAGAATGTGTGAACGAAAGCAGTGTGTTCCGACTGCACTACCAGAT  180
 Y  L  A  R  E  E  C  V  N  E  S  S  V  F  R  L  H  Y  Q  M   32
GACCGTGGTGCTGCTGATCGGCTCCTCCGTCCTCCTCACGGCCGCGGAATTCTTCGGGAA  240
 T  V  V  L  L  I  G  S  S  V  L  L  T  A  A  E  F  F  G  N   52
TCCCATCGACTGCATCACGGGCCTCGGGGCCAGGCACGTCATCAACACGTACTGTTGGAT  300
 P  I  D  C  I  T  G  L  G  A  R  H  V  I  N  T  Y  C  W  I   72
ACACTCGACGTTCACTATACAGGACTACTATTTAAGGGAAAGGGGATCAGAGGTAGCACA  360
 H  S  T  F  T  I  Q  D  Y  Y  L  R  E  R  G  S  E  V  A  Q   92
GCCAGGAGTGGGATCTCCCAATGGATACGATGAGGAGGAGGTTGAAGAAAGGTGGCGCTT  420
 P  G  V  G  S  P  N  G  Y  D  E  E  E  V  E  E  R  W  R  F   112
```

```
CCATAACTACTACCAGTGGGTTGTCTTCTTCCTTTTCTTCCAGGCTATTTTGTGCTATAT 480
 H  N  Y  Y  Q  W  V  V  F  F  L  F  F  Q  A  I  L  C  Y  M  132
GCCAAAATTCATCTGGAATAATGCTGAAGGAGGCTTGATGAAGACCATTGCTAATGGCCT 540
 P  K  F  I  W  N  N  A  E  G  G  L  M  K  T  I  A  N  G  L  152
GAACCCTGGATTGTTCAGGGAGGACGAAGTCAGCTCTCGCAAGAAGGTCATCATTGACTA 600
 N  P  G  L  F  R  E  D  E  V  S  S  R  K  K  V  I  I  D  Y  172
CATTGTGAAGCATATTCGGATGCACAATGGTTATGTATTCAAGTACTGGTTTTGCGAGTT 660
 I  V  K  H  I  R  M  H  N  G  Y  V  F  K  Y  W  F  C  E  L  192
GTTATGTTTTGTCAACATTGTTGGGCAGCTCTTCCTGATTGACAGGTTCCTTGGTGGAGA 720
 L  C  F  V  N  I  V  G  Q  L  F  L  I  D  R  F  L  G  G  E  212
GTTCTTGACCTATGGACCAAGAGTTGTAGAGTACAGTGAAATGGACCAGGAGGAGCGTGT 780
 F  L  T  Y  G  P  R  V  V  E  Y  S  E  M  D  Q  E  E  R  V  232
GGACCCCATGATCTATGTCTTCCCACGAATGACCAAGTGTCACTTCCACAAATTTGGTCC 840
 D  P  M  I  Y  V  F  P  R  M  T  K  C  H  F  H  K  F  G  P  252
TTCAGGAACTCTTGAGCGTCATGATGCATTCTGCCTTTTGCCTCTCAATATTCTGAATGA 900
 S  G  T  L  E  R  H  D  A  F  C  L  L  P  L  N  I  L  N  E  272
GAAGGTGTTCATCATGATCTGGTTCTGGTTTGTCATCTTGGCTTGCCTGCTTGGTGCCCT 960
 K  V  F  I  M  I  W  F  W  F  V  I  L  A  C  L  L  G  A  L  292
CATCCTGTACCGAGTTGCCCTCTTCACACTCCCTGGCCTTCGCCCCAGGGCCATGCACAA 1020
 I  L  Y  R  V  A  L  F  T  L  P  G  L  R  P  R  A  M  H  K  312
GCATAACAAGTCTATTTCTCTGGAGACCGTACAAGCCATCACCAACAAGACTTCCATTGG 1080
 H  N  K  S  I  S  L  E  T  V  Q  A  I  T  N  K  T  S  I  G  332
AGACTGGTGGATCCTGTATGTACTATCAACCAATATTGACCCTCTTATCTACCGTGACAT 1140
 D  W  W  I  L  Y  V  L  S  T  N  I  D  P  L  I  Y  R  D  I  352
CATGACTACCCTGGCCAAAGAAATCGAGACTGCGAATAGCAACAATCCTTACAACAGTGC 1200
 M  T  T  L  A  K  E  I  E  T  A  N  S  N  N  P  Y  N  S  A  372
AGGGTTATACTCATCATCATCAGTATGAATGTTAGAGGACATGGATTGAGGGCTGTCACT 1260
 G  L  Y  S  S  S  S  V  *                                   380
CAGCTGAAGGTGTAATAAAGACTTTTAATGAAGTAAAGGACAAAAACAACTGGGACCTTT 1320
TATACTGTGTTTTCTGCCCACACTTGTTTCTTTGATTGGGCGAGTGATGGCTAAATTATT 1380
TTGAGTACTTTTGTTTTGGTGCTATGTATGGTGTGCTTGGGAACAGCTACTGGTATGTGT 1440
CATTTAATTTACGGACCCTTGCTTTTGATGAACTTATCAGTTTATAGCACTTGCACAGAG 1500
ATATGGAACCCATTGGAAAAGAAATTATTATTTTGTTTAAATCTGTTTAATGCACCAATC 1560
CTTCAGCTGATCCATGATGTGACATTAATGTTCTCATGTTTCAGTTTTTTCTAAGATTTA 1620
GTTCTCCTTACAATAGAGGTGGACTAGCCATGTTGCATGAGGGTGACAGCCACCAGTTCC 1680
TGTCTTGAATAGAATATATCTAGATTGTTTTTCCTCTTTCACTGGACTAGAATCTTTCAG 1740
TACTTCTTCGTTGTTGAAATGTGAAGTAAGCCTTGATCTTATGATGTATATGTGCTTTTA 1800
CGTCTAAAAGGTCTTCAAATATTTATTTATTCTTAATGTGTCAAGGTAATGCACTATTAT 1860
ATTCAGCATTGATTTGCATTTCTGCTGGATAGCCTTTTCAATAATAAACAAAAGAAATGT 1920
AAATAGATGTTGGCATTTTAATGTTATCTATCAAGACTGACAGTGTATGAATTATGGAAC 1980
GACAACTTTGCTGGTCAAAGATACTTCTTAAAGTGCCATTATCTAAATTGACGCGAGAGC 2040
ATTTAGTTACTAATGAACAACACAATCATCTTGACTATGTCCTGTAATGGCAAACCATGT 2100
ACTGTAACGATCATGCTCTCTACGGAAGCAGCTGTTCTTTTATGGATTGAAAATATGCGA 2160
ATCTTGACTTTGCAGGAAGGATTGTTTCTCCGATATTATAAACTGAAGGAAGCAAGTTTT 2220
ATTTGTGAAAGGCTTTATTTGAATCTTATTTCCCCTCCCCAGTTGATAGATGTATCACAA 2280
AAAGTTGTCCGAAAAGTGCCAAGTGAGAATCGGGTCTATCTGCTCACGTGACTGATTGTT 2340
GGATGAACCAGATCAGGTCGTTTGGAATGTGCAGTTAAAGGACTTCTGTTTGTGATTTCA 2400
CTGTATAATACAGTCCTAAGTTTAAGATATATTTTGTATGTAAGGTTCATGTATTTTGTA 2460
GCTACATTAAAGAGATTTTTATAAAAAAAAAAAAAAAAAAAAAAA                2505
```

Figure 3.50 The full length cDNA and deduced protein sequences of *innexin1* (2505 bp; ORF 1143 bp, 380 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An innexin domain (positions 22th - 363th of the deduced protein) is highlighted.

```
ACGCGGGGAGTGCCACTTGGACGTCCTAGCGAGGAGGAGTGCCTTACCTCCCGTTGCTGG 60
ATCCTTTATAGGATATTTTCTCTCTGCGTGGCTAGTCTCGAAGCCCCCTCGCCTCTCCAC 120
GATGCGTGATGTTTTCGACTCAATCCGGGGTCTGCTCAAAGTGGACTCCCTCAGCGTGGA 180
 M  R  D  V  F  D  S  I  R  G  L  L  K  V  D  S  L  S  V  D  20
CAACAAGATCTTCCAAATGCACTACAAAGTCACGATGTTCTTTCTCCTGGCGTGTAGCTT 240
 N  K  I  F  Q  M  H  Y  K  V  T  M  F  F  L  L  A  C  S  L  40
GCTCGTGACACAGCGGCAGTACTTCGGCGACCCCATCGACTGCATCGTGGAAACGGTGGA 300
```

```
 L  V  T  Q  R  Q  Y  F  G  D  P  I  D  C  I  V  E  T  V  D   60
TCAGGAGGTGATGGATACCTACTGCTGGATCCACGCAACGTTCACCATCCCCGAAATGAA 360
 Q  E  V  M  D  T  Y  C  W  I  H  A  T  F  T  I  P  E  M  N   80
CGGCGCCGCCGTCGGGCACGAGGTGGCGCACCCGGGCATTGCGAACCCAAACGTCCCCGG 420
 G  A  A  V  G  H  E  V  A  H  P  G  I  A  N  P  N  V  P  G  100
GGAAGAGAAACGGGAAATCAAACACCACAAGTACTACCAGTGGGTGACCCTCGTCCTGGC 480
 E  E  K  R  E  I  K  H  H  K  Y  Y  Q  W  V  T  L  V  L  A  120
GATACAGGCCCTCTTCTTCTACGTCCCGCGCTACCTGTGGAAGATGTGGGAAGGAGGCAA 540
 I  Q  A  L  F  F  Y  V  P  R  Y  L  W  K  M  W  E  G  G  K  140
GATTAAGATGCTGGTGATGCAGCTTGACTCCCCCATTGTTGACGACGACGTGAAGAAGGA 600
 I  K  M  L  V  M  Q  L  D  S  P  I  V  D  D  D  V  K  K  E  160
GCGCAAGGACATGCTCGTCTCTTACTTCAGGATGAATATGAATAACCATAACTTCTACGC 660
 R  K  D  M  L  V  S  Y  F  R  M  N  M  N  N  H  N  F  Y  A  180
GTTTAAGTATTTCTCCTGCGAGGTTCTGAACTTCATCAATGTGATCGCACAGATTTACGT 720
 F  K  Y  F  S  C  E  V  L  N  F  I  N  V  I  A  Q  I  Y  V  200
GACGGATGCGTTCCTCGGCCACAGTTTCTCGAGGTACGGCCGAGAGGTGATCGAGTTCAG 780
 T  D  A  F  L  G  H  S  F  S  R  Y  G  R  E  V  I  E  F  S  220
CCAGCAGGAGATCACCTCGCGGGACGACCCGATGGACCGCGTGTTCCCCAAGGTGGCCAA 840
 Q  Q  E  I  T  S  R  D  D  P  M  D  R  V  F  P  K  V  A  K  240
GTGCACCTTCCACATGTCAGGGGCTTCGGGTAGCCTGGAGAAGCACGACGGACTCTGCGT 900
 C  T  F  H  M  S  G  A  S  G  S  L  E  K  H  D  G  L  C  V  260
TCTCCCGCTCAACATCTTCAACGAGAAGATCTACATTTTCCTGTGGTTCTGGTTCATCAT 960
 L  P  L  N  I  F  N  E  K  I  Y  I  F  L  W  F  W  F  I  I  280
CGTAGCTGTGATCACGGCCGTCGGTCTGCTCTACCGCATCGCCACCTTCCTCCCCGGCTT 1020
 V  A  V  I  T  A  V  G  L  L  Y  R  I  A  T  F  L  P  G  F  300
CAGGCAGATCCTCCTGAAGACCAAGTCGCGCCTCGCCTCCTCAGGGACGGTCGAGGCAGT 1080
 R  Q  I  L  L  K  T  K  S  R  L  A  S  S  G  T  V  E  A  V  320
CACTCGCCGCTGCGAAATCGGCGACTGGTTCCTGCTCTATCAGCTGGCCAAGAACATGGA 1140
 T  R  R  C  E  I  G  D  W  F  L  L  Y  Q  L  A  K  N  M  D  340
TCCTCTCATATACAAAGAATTCCTGAGCGAACTTGCCTACAAACTGGACGAAAGCTGATT 1200
 P  L  I  Y  K  E  F  L  S  E  L  A  Y  K  L  D  E  S  *     358
TCCGAAGCGTCACAGAAAAAAAATCACATTCCTTTTCTTGTGCCAAAGACCATAAGGTCT 1260
TCCCATATCAAAAAACTTTGAATTATGGAGAGGTGGGTATAGCCTGTAGATATTTCATAC 1320
TTTTTGTTGTTACTTAGAAAAGTATGGGAATGAATTTCTCTGTGGAATAGGAGGAAAATA 1380
AGAAAATATTTGTTGTACAGTTATCCCATCACCAAAACACCCCACATTCCTTAAACTACA 1440
ATACTGATGTAATATGAGTCCCAGTTTATATGTGCTTTCCTTACATTGTATAATAGAGTA 1500
TGTGTGTGGTCAATAAATACCGGATATTGTGTGTCGGATTCGCGTACGATGTGTATTATT 1560
TTTTTTCATTTTAATAAAGGCTTATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA     1615
```

Figure 3.51 The full length cDNA and deduced protein sequences of *innexin2* (1615 bp; ORF 1077 bp, 358 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An innexin domain (positions positions 20th - 358th of the deduced protein) is highlighted.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

A homologue of *saposin* was initially found in the heat stress hemocyte cDNA library. The full length cDNA of *saposin* were successfully characterized in testis of *P. monodon*. 5′ RACE-PCR of *saposin* was carried out twice. A fragment of 1100 bp obtained from the first 5′ RACE-PCR was cloned and sequenced (Fig. 3.52). Nevertheless, the full length cDNA was still not obtained. Another primer was designed and the second 5′ RACE-PCR was carried out and the amplification product of 1700 bp was cloned and sequenced.

The full length cDNAs of *saposin* was 3034 bp in length with an ORF of 2589 bp corresponding to a polypeptide of 862 amino acids (Fig. 3.53). The closest similarity of this transcript was *saposin* of *Aedes aegypti* (E-value = 5e-140). A deduced saposin protein contained 2 domains of saposin/surfactant protein-B A-type domain (SAPA domains, positions 25th - 58th and 823rd - 856th, E-value = 2.74e-12 and 1.56e-07, respectively) and 7 domains of Saposin (B) domains (SAPB domains, positions 68th -144th, 178th – 251st, 272nd - 346th, 437th - 512th, 531st - 606th, 646th – 721st, and 738th - 813th, E-value = 1.32e-22, 5.32e-09, 7.28e-16, 4.34e-23, 4.61e-27, 2.63e-22, and 5.83e-15, respectively). The expected MW and p*I* of the deduced saposin were 95.63 kDa and 4.65, respectively.

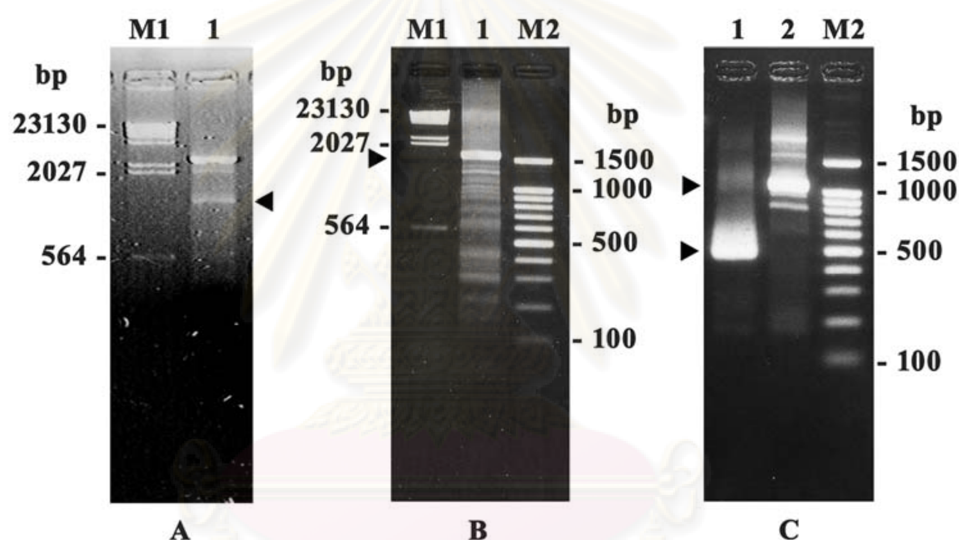


Figure 3.52 RACE-PCR products of 5′-*saposin* (A and B), 3′-*cdk7* (C; lane 1) and 3′ *dihydrolipoamide dehydrogenase* (C; lane 2) of *P. monodon*. The 100 bp marker (lane M2) and λ-*Hind* III (lane M1) was used as the marker.

3′ RACE-PCR of *cdk7* was successfully carried out and the amplification product of 500 bp was identified (Fig. 3.52). After sequence assembly, the full length cDNA of *cdk7* was obtained. It was 1431 bp in length with an ORF of 1062 bp corresponding to a polypeptide of 353 amino acids (Fig. 3.54). The closest similarity of this transcript was *cdk7* of *Drosophila melanogaster* (E-value = 1e-122). A deduced cdk7 protein contained serine/threonine protein kinases, catalytic domain

(S_TKc domain, positions 20th - 304th of the deducede protein, E-value = 2.83e-96). The expected MW and p*I* of this gene product were 95.63 kDa and 4.65, respectively.

```
ATACATCACATGACGACTCTCGGCTGATCTCTTCCGTTTTGCTTCACTTGTGTGTGTTCT 60
CTCTGGTGCTGGGGTTGTGGTCGTCCTTACTGCGTAGGTCAAGTTCCTAAGCCACCATGA 120
                                                        M  K 2
AGCTGCCTGGCTTTGGGTTGGCTCTCCTCCTGGCAGCATGTGTTGTGCACAGTGAAGCTA 180
 L  P  G  F  G  L  A  L  L  L  A  A  C  V  V  H  S  E  A  T 22
CCTTGTTAGGATCCAGGAAGTGCACATTTGGTCCCAGTTACTGGTGCCACTCTATTCAAA 240
 L  L  G  S  R  K  C  T  F  G  P  S  Y  W  C  H  S  I  Q  N 42
ATGCTAAGGAGTGCAATGCAGTTAATCACTGTATTCAGACGATATGGAAAACCTGGAAT 300
 A  K  E  C  N  A  V  N  H  C  I  Q  T  I  W  E  N  L  E  L 62
TGCCTGAAGACAATGATGATATTTGCACCTTGTGTAAGAACATGGTGAAGGAGGCCAGAG 360
 P  E  D  N  D  D  I  C  T  L  C  K  N  M  V  K  E  A  R  D 82
ACCAACTTCTCAGTAATGAAACCCAGGAAGAAATTCGTGAGGTGTTTGATGGGTCATGCC 420
 Q  L  L  S  N  E  T  Q  E  E  I  R  E  V  F  D  G  S  C  R 102
GTCTCATCCCCATTAAGATAATCTCTGACGAATGTGTTGATATTGCCAACGACTTCATTC 480
 L  I  P  I  K  I  I  S  D  E  C  V  D  I  A  N  D  F  I  P 122
CCGAATTGATTGACACACTGGCATCACAGATGAACCCCCAGTTGGTCTGTGCCACTGCAG 540
 E  L  I  D  T  L  A  S  Q  M  N  P  Q  L  V  C  A  T  A  G 142
GTCTTTGCAACTCTGCCAGAGTTGACAAACTCATCAGTGAAAATCAGGCTGCTCTTCAAG 600
 L  C  N  S  A  R  V  D  K  L  I  S  E  N  Q  A  A  L  Q  G 162
GATTCAACCCTAATGCATTAAAGCACTCGGGGGAGCACCCACAACCAGGAGACTGTGAAT 660
 F  N  P  N  A  L  K  H  S  G  E  H  P  Q  P  G  D  C  E  S 182
CTTGCAGGGACTTCATTGCACGAACCATCCGCCTTGTCAAGACCCACTCTCGAGCTGAGC 720
 C  R  D  F  I  A  R  T  I  R  L  V  K  T  H  S  R  A  E  L 202
TCGTGGACAGATTGATAGCTATATGTGGACGTTTTGGATCTCTCTCTGATGGCTGCAGTG 780
 V  D  R  L  I  A  I  C  G  R  F  G  S  L  S  D  G  C  S  A 222
CCTTGGTTGAAGCTAACTTCGATGATATTTACAATTTCTTAACAGATCAGCTTACACCAG 840
 L  V  E  A  N  F  D  D  I  Y  N  F  L  T  D  Q  L  T  P  E 242
AAGACTTCTGTGACCTAGTTGAAATGTGTGAGAACAGAATGCACCAGAGTGGACAGTACA 900
 D  F  C  D  L  V  E  M  C  E  N  R  M  H  Q  S  G  Q  Y  T 262
CTCGTCNTGCACTTTCACACTCTGGTGATGAACCTTGCGACTTCTGTGAAGCAATTGTGC 960
 R  X  A  L  S  H  S  G  D  E  P  C  D  F  C  E  A  I  V  Q 282
AACACTGGAGAGAGGTTCTTACAGCAAATACCACTGAAGAAGAATTCAAAGAGATCCTAG 1020
 H  W  R  E  V  L  T  A  N  T  T  E  E  E  F  K  E  I  L  D 302
ATGGCTTGTGTCGTCAGACTGGCAGGTTCAGCAAGAACTGCCTTGCTTTGGTAGATGAAT 1080
 G  L  C  R  Q  T  G  R  F  S  K  N  C  L  A  L  V  D  E  Y 322
ATTACCTGATTGTGTACAGCTTTCTGGTCTCTGAAATTCAACCCAAGGAAATCTGTGAAG 1140
 Y  L  I  V  Y  S  F  L  V  S  E  I  Q  P  K  E  I  C  E  A 342
CTGTAGGACTTTGTGGTTCTAATTCAGTTTTCAGTGGAGAGCATCCAGCTTGGACAGTGC 1200
 V  G  L  C  G  S  N  S  V  F  S  G  E  H  P  A  W  T  V  L 362
TTGATGTAAGCCAGAGAATTCCACAAACTCCTCTTAGGCCATCATTGATGGTTGGGCAGC 1260
 D  V  S  Q  R  I  P  Q  T  P  L  R  P  S  L  M  V  G  Q  H 382
ACTTGATTGGTGGAGATGAAAGCAGTGCCATCAGATTTGAGAGTGAAAACCAAGGTTCAA 1320
 L  I  G  G  D  E  S  S  A  I  R  F  E  S  E  N  Q  G  S  N 402
ATTTGCCACGTGTGAAGCTATCAAAAAGTGGCATTGGTGTGTCGAATGCTGGTAGGAATG 1380
 L  P  R  V  K  L  S  K  S  G  I  G  V  S  N  A  G  R  N  G 422
GGATGTTGGCTGCACCTGTTGGAAAGGGACGCGTTGGGGATGACAACAAGTGTGTGATGT 1440
 M  L  A  A  P  V  G  K  G  R  V  G  D  D  N  K  C  V  M  C 442
GCGAGTTTGCTCTTCATTTCCTGCAGAATATGCTTGAGCAGAAGGACACTCGTAAGGACA 1500
 E  F  A  L  H  F  L  Q  N  M  L  E  Q  K  D  T  R  K  D  I 462
TCGAAGATGCTGTTGAGAGGCTGTGCACCATGATGCCCCATTCACTGGCAGAGGAGTGTG 1560
 E  D  A  V  E  R  L  C  T  M  M  P  H  S  L  A  E  E  C  E 482
AGGACTATGTAGATGCCTATGGTGACCAAGTTATTGAGTTGCTGGCTCAAGAGATTGACC 1620
 D  Y  V  D  A  Y  G  D  Q  V  I  E  L  L  A  Q  E  I  D  P 502
CATCCCAGATCTGTCCTATGCTGCATCTCTGCCCATCTGAAGGAGAGTCAGAGGAGGCAG 1680
 S  Q  I  C  P  M  L  H  L  C  P  S  E  G  E  S  E  E  A  E 522
AGCAGGTCACATCTGAAAAACCTGATGTATCTTGTGTTGTGTGTGAGTATGCTTTGACCC 1740
 Q  V  T  S  E  K  P  D  V  S  C  V  V  C  E  Y  A  L  T  Q 542
AGCTGGAGGACATGTTGGAAGATAACAGAACTGAAGCAGGCATTGAGAGTGCTCTAGAAA 1800
 L  E  D  M  L  E  D  N  R  T  E  A  G  I  E  S  A  L  E  R 562
GGTTGTGTGCCCTTTTGCCCAAGTCAGCACGTAAAGAATGTGATATGTTTGTTGAAATGT 1860
 L  C  A  L  L  P  K  S  A  R  K  E  C  D  M  F  V  E  M  Y 582
```

```
ACACTGATCAGGTCATACAGATGTTGCTCAACAACTTGTCCCCTGATGAAATATGCACTA 1920
  T  D  Q  V  I  Q  M  L  L  N  N  L  S  P  D  E  I  C  T  N 602
ACCTGGGATTGTGTAAGCAAACAGAAAGTGCATTGCCTGTGCTTGATGCCTCTCACCAGT 1980
  L  G  L  C  K  Q  T  E  S  A  L  P  V  L  D  A  S  H  Q  L 622
TGCCAGTGTCTCGCATGTTCGTTCCAGCTGTATCATCACAAACAACAAACAACTTGGAGA 2040
  P  V  S  R  M  F  V  P  A  V  S  S  Q  T  T  N  N  L  E  M 642
TGACACAGTCTGCAGCATGTGTGTTGTGCGAGTTTGCTATGGTTCAGGTTGATGACTTGC 2100
  T  Q  S  A  A  C  V  L  C  E  F  A  M  V  Q  V  D  D  L  L 662
TCTCAGAAAATGCTACTGAGGATGAAATCATTGAAGTTGTGGACTTCATTTGTGCTCATA 2160
  S  E  N  A  T  E  D  E  I  I  E  V  V  D  F  I  C  A  H  M 682
TGCCAGGTGTTCTTGCTGATGATTGTATTGGCTTTGTTGAACAGTATGCTGATGCCATCA 2220
  P  G  V  L  A  D  D  C  I  G  F  V  E  Q  Y  A  D  A  I  I 702
TCAAGTTATTGGTTCATGAACTTGGTCCGAAGACTGTTTGCCAACAGATCAAGCTCTGCA 2280
  K  L  L  V  H  E  L  G  P  K  T  V  C  Q  Q  I  K  L  C  K 722
AACCTCCATCATTTGAAAGTATGAGAGCCCTTATTAATATGAGAATGGACAAATGCCAAG 2340
  P  P  S  F  E  S  M  R  A  L  I  N  M  R  M  D  K  C  Q  V 742
TTTGTGAGGGAGTTGTTAACTATATTGACAAGAAGCTGAAGGATGGAGATGCAACGACCA 2400
  C  E  G  V  V  N  Y  I  D  K  K  L  K  D  G  D  A  T  T  T 762
CCATTGACACGGTTCTTGAAGAAGTTTATCGACTCTTCCCGAATAATGCAAAGGACACGT 2460
  I  D  T  V  L  E  E  V  Y  R  L  F  P  N  N  A  K  D  T  C 782
GCCGCAGTATGATCGAAGTTTATGGGCCCTATGTAGTGAACCTCCTGGCTGAACTAGGGG 2520
  R  S  M  I  E  V  Y  G  P  Y  V  V  N  L  L  A  E  L  G  D 802
ACTCGAAGCGTGTGTGCCAGGCCATAAAGTTCTGCCCCCACCACACTTCGGAACCACTCC 2580
  S  K  R  V  C  Q  A  I  K  F  C  P  H  H  T  S  E  P  L  L 822
TGGGAGCTGAAAAGTGCACTTGGGGTCCTTCTTATTGGTGCCAGACCAAGATGCATGCAA 2640
  G  A  E  K  C  T  W  G  P  S  Y  W  C  Q  T  K  M  H  A  T 842
CAGCCTGTAAGGCAACTGTTCACTGTGAAACGAAGGTGTGGAAGGGAGTTGTTCCTGCCA 2700
  A  C  K  A  T  V  H  C  E  T  K  V  W  K  G  V  V  P  A  I 862
TTTAAAGATAAGAGGATTGATTGCAAGAAGAGGAGGAGGAAGAAAAGAAAAGAAAAAAAA 2760
  *
AAAGAAATGAAGATGAAGACAGAGTGGATGTAGATATTGGGAGGGCTTATTTTTTTCTCA 2820
AAAAAGCAGCAGTGCAAAAAAAAAGATAAATGGTTTCAAATATACACTTGTCATTATGAGA 2880
ACTACTGTCTGTGATGAATATTTTTTAATGCTTAATGATTATGATTCCTCCAGTTTGACT 2940
TATATTTTGTCTTGTTAGTTGTAAGGAATATTGGAATAACCAGGATATGTAAGAATATAT 3000
AGAAGAATAATTTTAAAAAAAAAAAAAAAAAAAA                           3034
```

Figure 3.53 The full length cDNA and deduced protein sequences of *saposin* (3034 bp; ORF 2589 bp, 862 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. Two domains of saposin/surfactant protein-B A-type domain (SAPA domain, positions 25th - 58th and 823rd - 856th) and 7 domains of saposin (B) Domains (SAPB domains, positions 68th - 144th, 178th – 251st, 272nd - 346th, 437th - 512th, 531st - 606th, 646th – 721st, and 738th - 813th of the deduced protein) are highlighted.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
GGCACGAGGGCGGAAGACAGGATGGAAGTAGAACAAGAGAAGAAAGGAAGGATTAGAATA 60
                     M  E  V  E  Q  E  K  K  G  R  I  R  I   13
GAAGAAAAATTGAAGAGATATGAGAAGATCGATTTCTTGGGAGAAGGACAGTTTGCCACT 120
E  E  K  L  K  R  Y  E  K  I  D  F  L  G  E  G  Q  F  A  T   23
GTATATAAGGCTCTTGATGTGGAGACCAAGCAGATAGTAGCTGTCAAAAAGATCAAACTA 180
V  Y  K  A  L  D  V  E  T  K  Q  I  V  A  V  K  K  I  K  L   53
GGTAGCAGAGAGGAGGCAAGGGATGGCATCAACCGTACGGCTCTCCGAGAGATCAAGCTC 240
G  S  R  E  E  A  R  D  G  I  N  R  T  A  L  R  E  I  K  L   73
TTGCAGGAGGTCCACCACCCAAACCTCATTGGCCTCCTCGATGTCTTTGGCTACAAGTCA 300
L  Q  E  V  H  H  P  N  L  I  G  L  L  D  V  F  G  Y  K  S   93
AATGTGTCGCTGGTGTTTGATTTCATGGATACAGATTTAGAGGTGATCATCAAGGACACA 360
N  V  S  L  V  F  D  F  M  D  T  D  L  E  V  I  I  K  D  T   113
GACAACATCATCCTCACGCCCTCCAACATCAAAGCATATATGATCCAAACATTAAAAGGC 420
D  N  I  I  L  T  P  S  N  I  K  A  Y  M  I  Q  T  L  K  G   133
```

```
TTGGAATTCCTGCATCTTCACTGGATCCTACACAGAGATCTGAAACCAAACAACCTACTA 480
L  E  F  L  H  L  H  W  I  L  H  R  D  L  K  P  N  N  L  L   153
GTCAATTCAGATGGCATACTTAAAATAGGAGATTTTGGTCTGGCAAGATTCTTTGGCTCT 540
V  N  S  D  G  I  L  K  I  G  D  F  G  L  A  R  F  F  G  S   173
CCCAACAGACAGTATTCACATCAAGTAGTTACAAGATGGTACAGGAGTCCAGAATTGCTG 600
P  N  R  Q  Y  S  H  Q  V  V  T  R  W  Y  R  S  P  E  L  L   193
TTTGGCGCGAGATCCTACGGCACAGGGGTAGACATGTGGGCGATTGGCTGTATCCTGGCG 660
F  G  A  R  S  Y  G  T  G  V  D  M  W  A  I  G  C  I  L  A   213
GAGATGTTGGTTCGCTGTCCCTACTTCCCGGGTGACTCTGATCTAGACCAGCTTACCAGG 720
E  M  L  V  R  C  P  Y  F  P  G  D  S  D  L  D  Q  L  T  R   233
ATCTTCACTGCCCTAGGGACTCCTGGTGATGACGACTGGCCGGACATGACGAAACTTCCC 780
I  F  T  A  L  G  T  P  G  D  D  D  W  P  D  M  T  K  L  P   253
GACTACGTATCATTCAAGCACTTCGAGGGTTCCCCACTGCGAGACCTCTTTCCTGCTGCC 840
D  Y  V  S  F  K  H  F  E  G  S  P  L  R  D  L  F  P  A  A   273
AGTGATGACCTTCTCCAGCTATTGGGGTCTTTGCTCACTATTAATCCTATGAAACGATGC 900
S  D  D  L  L  Q  L  L  G  S  L  L  T  I  N  P  M  K  R  C   293
AGCTGTACTGAGGCTCTGAAGATGGAGTATTTCAGCAATAAGCCTGTCCCGACACCAGGA 960
S  C  T  E  A  L  K  M  E  Y  F  S  N  K  P  V  P  T  P  G   313
CCTCTTCTTCCTCTCCCACCAACCATTAGACAGAGAAGTGAGGCAGAAAAACCGTCTCTC 1020
P  L  L  P  L  P  P  T  I  R  Q  R  S  E  A  E  K  P  S  L   333
AAGCGAAAGATTATTGAAGAGTCTGGCTTTGGAGGTTCCTTAGCAAAGAAGCTTCAATTC 1080
K  R  K  I  I  E  E  S  G  F  G  G  S  L  A  K  K  L  Q  F   353
TAGTCCATACAAGGAAGAGAAAAAGAAGCACTACCTCCTGATGTGCCAGCGTGACCTCGG 1140
*
TACTTAGTTGTAAAGGTGAAAATAAATTCCTTTAAATGAGGCAAGAAAGAGAGAGGCTGA 1200
TTTATGAAAGTGAATGATATTTATGTTGATGTTAATTTTTTTTCCTTTGACATGTTAAAG 1260
ACAAAGGGATGACAGTATTGTCCTTTTATTTCTTTATTTACTTGAGTACTGTGTAATATA 1320
TTGATGTTCACAGAAAGTGTGTTATTTTTTAAACCAGTGGTATTATTGTAAATCATTTCA 1380
GAATAAATTTTCATATAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA          1431
```

Figure 3.54 The full length cDNA and deduced protein sequences of *cdk7* (1431 bp; ORF 1062 bp, 353 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A serine/threonine protein kinases catalytic domain (S_TKc domain, positions 20th - 304th of the deduced protein) is highlighted.

3′ RACE-PCR of *troponinT* homologue were carried out. 3′ RACE products (600 bp) were cloned and sequenced (Fig. 3.55). After sequences of the RACR-PCR product and EST were assembled, the full length cDNA of *troponinT* was successfully isolated. It was 1430 bp in length with an ORF of 1107 bp encoding a polypeptide of 368 amino acids (Fig. 3.56). The closest similarity of this transcript was *TPA_inf: troponin T isoform 3* of *Drosophila pseudoobscura* (E-value = 3e-133). No functionall protein domain was found in the deduced troponin T of *P. monodon*. The expected MW and p*I* of the deduced troponinT protein were 43.92 kDa and 4.94, respectively.

5′ and 3′ RACE-PCR of *ERO1* homologue were carried out. The resulting products of 650 bp and 1100 bp were cloned and sequenced (Fig. 3.55). Two forms of the full length cDNA of *ERO1* were successfully characterized. The short form of

ERO1 was 1756 bp in length with an ORF of 1413 bp (*ERO1*-s; Fig. 3.57) corresponding to 470 amino acids. The long form of *ERO1* was 1774 bp in length with an ORF of 1434 bp corresponding to 477 amino acids (*ERO1*-l; Fig. 3.58). The closest similarity of these transcripts were *Ero1L CG1333-PB, isoform B* of *Apis mellifera* (E-value = 8e-145 and 4e-143, respectively). A deduced ERO1 protein contained an ERO1 domain (*ERO1*-s at positions 66th - 468th, E-value = 1.90e-158 and *ERO1*-l at positions 66th – 462nd, E-value = 8.20e-160). Amino acid and neucleotide sequences of these two isoform of *ERO1* were aligned (Fig. 3.59-3.60). Two isoforms of *ERO1* were different due to the presence/absence of a CCAATGATTTTCTGAAAG motif encoding a ANDFLK hexapeptide. The expected MW and p*I* of these gene products were 53.81 kDa and 6.34 and 54.5 kDa and 6.34, respectively.

5′ and 3′ RACE-PCR of *dihydrolipoamide dehydrogenase* (*DLADH*) homologue were carried out. 5′ RACE product (650 bp, Fig. 3.55) and 3′ RACE products (1100 and 800 bp, Fig. 3.52) were cloned and sequenced. Nucleotide sequences were assembled, 2 different isoforms of *DLADH* were successfully characterized. They were 1770 and 2050 bp in length with an identical ORF of 1521 bp, 506 amino acids but different 3′ UTR regions (Fig. 3.61 and 3.62). Nucleotide sequences of two isoforms of *DLADH* were aligned and length polymorphism rather than sequence variation at the 3′ UTRs (Fig. 3.63).

The closest similarity of these transcripts were *dihydrolipoamide dehydrogenase* of *Mus musculus* (E-value = 0.0). A deduced DLADH protein contained a pyridine nucleotide-disulphide oxidoreductase domain (Pyr_redox domain; positions 213th - 311th, E-value = 1.30e-25), a pyridine nucleotide-disulphide oxidoreductase 2 domain (Pyr_redox 2 domain; positions 41st - 358th, E-value = 8.90e-69), and a pyridine nucleotide-disulphide oxidoreductase dimerization domain (Pyr_redox_dim domain; positions 386th - 495th, E-value = 3.10e-63). The expected MW and p*I* of this gene product were 53.82 kDa and 6.96, respectively.

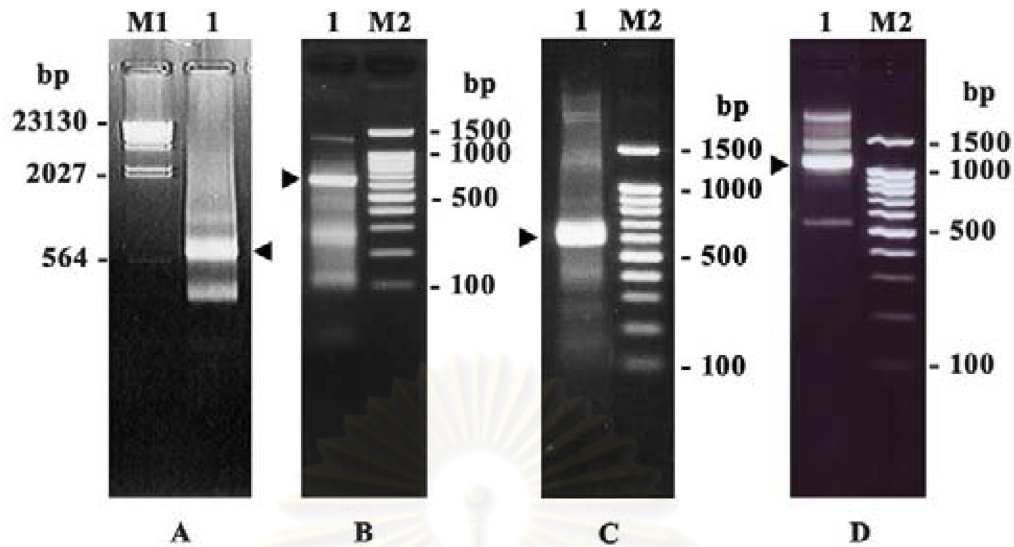


Figure 3.55 3′RACE-PCR of *troponin T* (A), 5′RACE-PCR of *dihydrolipoamide dehydrogenase* (B) and *ERO1* (C) and 3′RACE-PCR of *ERO1* (D) of *P. monodon*. A 100 bp marker (lane M2) and λ-*Hind* III (lane M1) was used as the marker.

```
CACGAGGCTCTCTCTCCTGCTCGTACTCTCCATCTCTTCGAAGACAGTAAATAACACACC 60
AGCACCATGTCTGACGACGAATCAGCATATTCGGATGCCGAAAAGAGGAGGAAGAAGGGA 120
      M  S  D  D  E  S  A  Y  S  D  A  E  K  R  R  K  K  G   18
GATGAGGGGGCCAACTTCCTTAAGAACCGCCAGCAGATGAAGATGTCCGAGCTGGACGAA 180
D  E  G  A  N  F  L  K  N  R  Q  Q  M  K  M  S  E  L  D  E   38
CAGCTGGCGGAATACATCGCTGAGTGGAGGAAGCAGAGATCCAAGGAAGAGGATGAGCTC 240
Q  L  A  E  Y  I  A  E  W  R  K  Q  R  S  K  E  E  D  E  L   58
AGGAAACTCAAGGAGAAGCAGGCCAAGAGGAAGATCCTCCGCGCCGAAGAGGAGAAGAAG 300
R  K  L  K  E  K  Q  A  K  R  K  I  L  R  A  E  E  E  K  K   78
CTGACCGAGCAGAAGAAGGCCGAGGAGGAGCGCAAGCTTCGCGAGGAATCCGAGAGGAAG 360
L  T  E  Q  K  K  A  E  E  E  R  K  L  R  E  E  S  E  R  K   98
CAGAAGGAACAGGAAGAGAAGAGGAAACGCCTCGAGGAAGCAGAGAAGAAGCGCCAGGCC 420
Q  K  E  Q  E  E  K  R  K  R  L  E  E  A  E  K  K  R  Q  A   118
ATGATGAAGGGATCCAGCGATGGCGTCAAGAAGTTCGGCGTTAAGAGCGGTGGTGACAAG 480
M  M  K  G  S  S  D  G  V  K  K  F  G  V  K  S  G  G  D  K   138
TTCTCCAACATCCAGGCGGCCAAGGGCGAACTGGGCAAGACCCGCGAGCAGCTGGCCGAG 540
F  S  N  I  Q  A  A  K  G  E  L  G  K  T  R  E  Q  L  A  E   158
GAGAAGAAGATTGCCCTGTCCATCCGCGTGAAGCCCCTCTGCGTTGACGGCGTTGGTGCG 600
E  K  K  I  A  L  S  I  R  V  K  P  L  C  V  D  G  V  G  A   178
TCCGCCCTCCGCCAGAAGGCCGAGGAAATGTGGAACCTCATCATCAAGCTGGAGACCGAG 660
S  A  L  R  Q  K  A  E  E  M  W  N  L  I  I  K  L  E  T  E   198
AAATACGACATGGAGGAGAGGATGAAGCGACAGGACTACGATCTGAAGGAGCTTAGGGAA 720
K  Y  D  M  E  E  R  M  K  R  Q  D  Y  D  L  K  E  L  R  E   218
CGTCAGAAGCAACAGCTCAGGCAAAAGGCCTTGAAGAAGGGTCTTGATCCCGAGGCTCTC 780
R  Q  K  Q  Q  L  R  Q  K  A  L  K  K  G  L  D  P  E  A  L   238
ACCGGAAAACACCCGCCCAAGATCCAGACTGCCTCCAAGTTCGAGCGACGCACAGACAGG 840
T  G  K  H  P  P  K  I  Q  T  A  S  K  F  E  R  R  T  D  R   258
AGGACCTATGACGACAAGAAGAAACTTTTCGAGGGTGGCTGGGAAGTTGTACACAACGAG 900
R  T  Y  D  D  K  K  K  L  F  E  G  G  W  E  V  V  H  N  E   278
GAGCTGGAAAGGTACTGGAAGGACAAATACGAAGAGTTTGTCAACAGGACGAAGTCCAAG 960
E  L  E  R  Y  W  K  D  K  Y  E  E  F  V  N  R  T  K  S  K   298
CTTCCCAAGTGGTTCGGCGAGCGACCGGGCAAGAAGGCCGGCGACCCCGAATCCCCCGAA 1020
L  P  K  W  F  G  E  R  P  G  K  K  A  G  D  P  E  S  P  E   318
```

```
GAGGAAGAGGGCGGTGAGGAAGACGAGATGCCACCCGAGCCTGTGTACGAGGAGGAAGAA 1080
E  E  E  G  G  E  E  D  E  M  P  P  E  P  V  Y  E  E  E  E   338
GAAGAAGAAGAGGAGGAGGAAGAAGAGGAGGAAGAGGAAGAGGAAGAGGAGGAGGAAGAG 1140
E  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E  E   358
GAAGAGGAGGAGGAGGAAGAGGAAGAGGAGTGATGAGTAAGTCTACCGTATAACTGTCCT 1200
E  E  E  E  E  E  E  E  E  E  *                              368
AGAAATAGCCGAATCGATGTGAGAAACTGGATCCCGTTTGAGGACTTCATGGTGTCTGAT 1260
TGTGATATGATGATAAAAGATGTCTTTGGATGTCGTCGCGTTATTTTGATTGCGTCATAG 1320
CATGTGAGCATACTCCGTGGGGTGCATCGTGTCATGTGCACCTCGTGACTGTACATCAGC 1380
GAGCAATAAAAGAGATCCAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAA           1430
```

Figure3.56 The full length cDNA and deduced protein sequences of *troponin T isoform 3* (1430 bp; ORF 1107 bp, 368 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined.

```
CTAATTACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAGTCAT 60
GACAGTTACGTCAGAGTAGTAGTATATATCGTCTACACTCCAGTCCCGGGATTCAGTATC 120
TGTTCTCTCGAGTGTCGCGAGTGTTTAAGAGCGAGTCATGATGTGGTCGAGGACTTCCGC 180
                                     M  M  W  S  R  T  S  A  8
GTGGGTGTTTGTCATTTTGGCAATTTCCTTTCACACTTGCACGGCGATTTGGTATGGGAT 240
 W  V  F  V  I  L  A  I  S  F  H  T  C  T  A  I  W  Y  G  I  28
TAAGAAGACCCCAAACGGAGACACAACTGCTGGTGAGAGGTGCTTCTGTCAGTTAAAAGG 300
 K  K  T  P  N  G  D  T  T  A  G  E  R  C  F  C  Q  L  K  G  48
AGTAATAGATGACTGTTCCTGTTCTGTGGAGACTTTAGACAGCTTCAACAACCTGAAGCT 360
 V  I  D  D  C  S  C  S  V  E  T  L  D  S  F  N  N  L  K  L  68
GTATCCACGCCTGAATAGCCTCCTGCAGTATGACTACTTCAGGTACTGGAAGGTGAACTT 420
 Y  P  R  L  N  S  L  L  Q  Y  D  Y  F  R  Y  W  K  V  N  L  88
GAAAAAAGAATGTCCTTTTTGGGAAGATGACAGCAAGTGTGCCATTCGTTACTGCAGTGT 480
 K  K  E  C  P  F  W  E  D  D  S  K  C  A  I  R  Y  C  S  V  108
GAAGCCATGTACTGATGTCCCAGAGGGTATAAAGGGAGCTTCCATAGACAAAATTGAGAA 540
 K  P  C  T  D  V  P  E  G  I  K  G  A  S  I  D  K  I  E  K  128
GGAAAAGAAGGAAAAGTCCCACATGGTGACTGGACATTGTGATGGAGAGAATGACCTTGG 600
 E  K  K  E  K  S  H  M  V  T  G  H  C  D  G  E  N  D  L  G  148
ATACTTGAATACTACTCTCAGCAAGGAATCCAAAGTTGGTTTTAAGCGCTGGGCAGCCCA 660
 Y  L  N  T  T  L  S  K  E  S  K  V  G  F  K  R  W  A  A  H  168
TGATGATGCACAGCTGAACTTTTGCAAAATTGATGACGACAGTTCTGAGGATAGTGAATA 720
 D  D  A  Q  L  N  F  C  K  I  D  D  D  S  S  E  D  S  E  Y  188
TGTTGATCTATTGCTGAATCCTGAGCGGTACACAGGTTATGCAGGACCTTCAGCACATAG 780
 V  D  L  L  L  N  P  E  R  Y  T  G  Y  A  G  P  S  A  H  R  208
AATATGGAGAACAATATACCAAGAAAATTGCTTTAAGCCATCTAGGGCAATTGGCCGTTA 840
 I  W  R  T  I  Y  Q  E  N  C  F  K  P  S  R  A  I  G  R  Y  228
CACAGACTTCAGTAGTATTGGAGAAATGTGCTTGGAAAAAAGAACATTTTACAGAGCTAT 900
 T  D  F  S  S  I  G  E  M  C  L  E  K  R  T  F  Y  R  A  I  248
TTCCGGTCTCCATACCAGTATCAACATTCACCTGAGTGCTAACTACCTCCTGTCAGACCA 960
 S  G  L  H  T  S  I  N  I  H  L  S  A  N  Y  L  L  S  D  Q  268
GAATGGCTTTGAAATGTCAAAGGATGGCCAGTGGGGTCCGAATGTACAGGAATTCCAGAC 1020
 N  G  F  E  M  S  K  D  G  Q  W  G  P  N  V  Q  E  F  Q  T  288
GAGGTTTGACCCAGAACTAACTGGCGGAGAGGGAACCCACCGCCTGAAGAACCTCTACTT 1080
 R  F  D  P  E  L  T  G  G  E  G  T  H  R  L  K  N  L  Y  F  308
TGTATACCTTCTTGAACTAAGAGCTTTAGCTAAAGCTGCACCATACCTTGAAAGCTTAGA 1140
 V  Y  L  L  E  L  R  A  L  A  K  A  A  P  Y  L  E  S  L  E  328
ATACTACACAGGAAATGAAAATGAAGATAATGATGTATCAAAGGCAGTTAAAGACTTATT 1200
 Y  Y  T  G  N  E  N  E  D  N  D  V  S  K  A  V  K  D  L  L  348
AACTGTTGTTAAGAGTTTTCCAGAGCACTTTGATGAAAGCTCCATGTTCTCTGGCGGCCA 1260
 T  V  V  K  S  F  P  E  H  F  D  E  S  S  M  F  S  G  G  Q  368
GCATGCAGCTAAGTTGAAAGAGGAATTTCGGCAACATTTTTGGAATGTGTCTCGTATTAT 1320
 H  A  A  K  L  K  E  E  F  R  Q  H  F  W  N  V  S  R  I  M  388
GGACTGCGTTGGATGTGACAAGTGTCGTCTTTGGGGCAAGCTGCAGGTAACTGGCCTTGG 1380
 D  C  V  G  C  D  K  C  R  L  W  G  K  L  Q  V  T  G  L  G  408
TACAGCACTCAAGATTCTGTTTTCAGGCAATCTAGACCCAGAAGGTAACCAAAACTTAGA 1440
 T  A  L  K  I  L  F  S  G  N  L  D  P  E  G  N  Q  N  L  D  428
TCTTCCAGCCGTAAGACACACAAAGTTCCAGCTGTCTCGTATTGAGATTGTTGCTCTTAT 1500
 L  P  A  V  R  H  T  K  F  Q  L  S  R  I  E  I  V  A  L  I  448
```

```
CAATGCCTTTGGTCGGTTGTCCAGCAGCATCATGGCGCTAGAGAACTTTCGGCAGAGCAT 1560
 N  A  F  G  R  L  S  S  S  I  M  A  L  E  N  F  R  Q  S  I  468
AACCAGAAGATAATGACAAGATTGGCTACTTTGAGCTGAACTTCTGTGATATTTGTTATA 1620
 T  R  R  *                                                  471
TGTTACCCAAATACATAGTTGTATGCTAGATTATGACAGGTGGGCTGCCAACTGCCAGTG 1680
GTAGCATCATATATTCAAAGAAGGAACATGACAGATGTAAGACAAGGAAAAAAAAAAAAA 1740
AAAAAAAAAAAAAAAA                                             1756
```

Figure 3.57 The full length cDNA and deduced protein sequences of the short form of *Ero1L, isoform B* (*ERO1-s*, 1756 bp; ORF 1413 bp, 470 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An ERO1 domain (*ERO1*-s, positions 66th - 468th of the deduced protein) is highlighted.

```
CTAATTACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAGTCAT 60
GACAGTTACGTCAGAGTAGTAGTATATATCGTCTACACTCCAGTCCCGGGATTCAGTATC 120
TGTTCTCTCGAGTGTCGCGAGTGTTTAAGAGCGAGTCATGATGTGGTCGAGGACTTCCGC 180
                                     M  M  W  S  R  T  S  A  8
GTGGGTGTTTGTCATTTTGGCAATTTCCTTTCACACTTGCACGGCGATTTGGTATGGGAT 240
 W  V  F  V  I  L  A  I  S  F  H  T  C  T  A  I  W  Y  G  I  28
TAAGAAGACCCCAAACGGAGACACAACTGCTGGTGAGAGGTGCTTCTGTCAGTTAAAAGG 300
 K  K  T  P  N  G  D  T  T  A  G  E  R  C  F  C  Q  L  K  G  48
AGTAATAGATGACTGTTCCTGTTCTGTGGAGACTTTAGACAGCTTCAACAACCTGAAGCT 360
 V  I  D  D  C  S  C  S  V  E  T  L  D  S  F  N  N  L  K  L  68
GTATCCACGCCTGAATAGCCTCCTGCAGTATGACTACTTCAGGTACTGGAAGGTGAACTT 420
 Y  P  R  L  N  S  L  L  Q  Y  D  Y  F  R  Y  W  K  V  N  L  88
GAAAAAAGAATGTCCTTTTTGGGAAGATGACAGCAAGTGTGCCATTCGTTACTGCAGTGT 480
 K  K  E  C  P  F  W  E  D  D  S  K  C  A  I  R  Y  C  S  V  108
GAAGCCATGTACTGATGTCCCAGAGGGTATAAAGGGAGCTTCCATAGACAAAATTGAGAA 540
 K  P  C  T  D  V  P  E  G  I  K  G  A  S  I  D  K  I  E  K  128
GGAAAAGAAGGAAAAGTCCCACATGGTGACTGGACATTGTGATGGAGAGAATGACCTTGG 600
 E  K  K  E  K  S  H  M  V  T  G  H  C  D  G  E  N  D  L  G  148
ATACTTGAATACTACTCTCAGCAAGGAATCCAAAGTTGGTTTTAAGCGCTGGGCAGCCCA 660
 Y  L  N  T  T  L  S  K  E  S  K  V  G  F  K  R  W  A  A  H  168
TGATGATGCACAGCTGAACTTTTGCAAAATTGATGACGACAGTTCTGAGGATAGTGAATA 720
 D  D  A  Q  L  N  F  C  K  I  D  D  D  S  S  E  D  S  E  Y  188
TGTTGATCTATTGCTGAATCCTGAGCGGTACACAGGTTATGCAGGACCTTCAGCACATAG 780
 V  D  L  L  L  N  P  E  R  Y  T  G  Y  A  G  P  S  A  H  R  208
AATATGGAGAACAATATACCAAGAAAATTGCTTTAAGCCATCTAGGGCAATTGGCCGTTA 840
 I  W  R  T  I  Y  Q  E  N  C  F  K  P  S  R  A  I  G  R  Y  228
CACAGACTTCAGTAGTATTGGAGCCAATGATTTTCTGAAAGAAATGTGCTTGGAAAAAAG 900
 T  D  F  S  S  I  G  A  N  D  F  L  K  E  M  C  L  E  K  R  248
AACATTTTACAGAGCTATTTCCGGTCTCCATACCAGTATCAACATTCACCTGAGTGCTAA 960
 T  F  Y  R  A  I  S  G  L  H  T  S  I  N  I  H  L  S  A  N  268
CTACCTCCTGTCAGACCAGAATGGCTTTGAAATGTCAAAGGATGGCCAGTGGGGTCCGAA 1020
 Y  L  L  S  D  Q  N  G  F  E  M  S  K  D  G  Q  W  G  P  N  288
TGTACAGGAATTCCAGACGAGGTTTGACCCAGAACTAACTGGCGGAGAGGGAACCCACCG 1080
 V  Q  E  F  Q  T  R  F  D  P  E  L  T  G  G  E  G  T  H  R  308
CCTGAAGAACCTCTACTTTGTATACCTTCTTGAACTAAGAGCTTTAGCTAAAGCTGCACC 1140
 L  K  N  L  Y  F  V  Y  L  L  E  L  R  A  L  A  K  A  A  P  328
ATACCTTGAAAGCTTAGAATACTACACAGGAAATGAAAATGAAGATAATGATGTATCAAA 1200
 Y  L  E  S  L  E  Y  Y  T  G  N  E  N  E  D  N  D  V  S  K  348
GGCAGTTAAAGACTTATTAACTGTTGTTAAGAGTTTTCCAGAGCACTTTGATGAAAGCTC 1260
 A  V  K  D  L  L  T  V  V  K  S  F  P  E  H  F  D  E  S  S  368
CATGTTCTCTGGCGGCCAGCATGCAGCTAAGTTGAAAGAGGAATTTCGGCAACATTTTTG 1320
 M  F  S  G  G  Q  H  A  A  K  L  K  E  E  F  R  Q  H  F  W  388
GAATGTGTCTCGTATTATGGACTGCGTTGGATGTGACAAGTGTCGTCTTTGGGGCAAGCT 1380
 N  V  S  R  I  M  D  C  V  G  C  D  K  C  R  L  W  G  K  L  408
GCAGGTAACTGGCCTTGGTACAGCACTCAAGATTCTGTTCTCAGGCAATCTAGACCCAGA 1440
 Q  V  T  G  L  G  T  A  L  K  I  L  F  S  G  N  L  D  P  E  428
AGGTAACCAAAACTTAGATCTTCCAGCCGTGAGACACACAAAGTTCCAGCTGTCTCGTAT 1500
 G  N  Q  N  L  D  L  P  A  V  R  H  T  K  F  Q  L  S  R  I  448
```

```
TGAGATTGTTGCTCTTATCAATGCCTTTGGTCGGTTGTCCAGCAGCATCATGGCGCTAGA 1560
 E  I  V  A  L  I  N  A  F  G  R  L  S  S  S  I  M  A  L  E  468
GAACTTTCGGCAGAGCATAACCAGAAGATAATGACAAGATTGGCTACTTTGAGCTGAACT 1620
 N  F  R  Q  S  I  T  R  R  *                                477
TCTGTGATATTTGTTATATGTTACCCAAATACATAGTTGTATGCTAGATTATGACAGGTG 1680
GGCTGCCAACTGCCAGTGGTAGCATCATATATTCAAAGAAGGAACATGACAGATGTAAGA 1740
CAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAA                            1774
```

Figure3.58 The full length cDNA and deduced protein sequences of the long form of *Ero1L, isoform B* (*ERO1-l*; 1774 bp; 1434 bp, 477 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. An ERO1 domain (positions 66th - 462th of the deduced protein) is highlighted.

```
ERO1-s      MMWSRTSAWVFVILAISFHTCTAIWYGIKKTPNGDTTAGERCFCQLKGVIDDCSCSVETL 60
ERO1-l      MMWSRTSAWVFVILAISFHTCTAIWYGIKKTPNGDTTAGERCFCQLKGVIDDCSCSVETL 60
            ************************************************************

ERO1-s      DSFNNLKLYPRLNSLLQYDYFRYWKVNLKKECPFWEDDSKCAIRYCSVKPCTDVPEGIKG 120
ERO1-l      DSFNNLKLYPRLNSLLQYDYFRYWKVNLKKECPFWEDDSKCAIRYCSVKPCTDVPEGIKG 120
            ************************************************************

ERO1-s      ASIDKIEKEKKEKSHMVTGHCDGENDLGYLNTTLSKESKVGFKRWAAHDDAQLNFCKIDD 180
ERO1-l      ASIDKIEKEKKEKSHMVTGHCDGENDLGYLNTTLSKESKVGFKRWAAHDDAQLNFCKIDD 180
            ************************************************************

ERO1-s      DSSEDSEYVDLLLNPERYTGYAGPSAHRIWRTIYQENCFKPSRAIGRYTDFSSIG----- 235
ERO1-l      DSSEDSEYVDLLLNPERYTGYAGPSAHRIWRTIYQENCFKPSRAIGRYTDFSSIGANDFL 240
            *******************************************************

ERO1-s      -EMCLEKRTFYRAISGLHTSINIHLSANYLLSDQNGFEMSKDGQWGPNVQEFQTRFDPEL 294
ERO1-l      KEMCLEKRTFYRAISGLHTSINIHLSANYLLSDQNGFEMSKDGQWGPNVQEFQTRFDPEL 300
             ***********************************************************

ERO1-s      TGGEGTHRLKNLYFVYLLELRALAKAAPYLESLEYYTGNENEDNDVSKAVKDLLTVVKSF 354
ERO1-l      TGGEGTHRLKNLYFVYLLELRALAKAAPYLESLEYYTGNENEDNDVSKAVKDLLTVVKSF 360
            ************************************************************

ERO1-s      PEHFDESSMFSGGQHAAKLKEEFRQHFWNVSRIMDCVGCDKCRLWGKLQVTGLGTALKIL 414
ERO1-l      PEHFDESSMFSGGQHAAKLKEEFRQHFWNVSRIMDCVGCDKCRLWGKLQVTGLGTALKIL 420
            ************************************************************

ERO1-s      FSGNLDPEGNQNLDLPAVRHTKFQLSRIEIVALINAFGRLSSSIMALENFRQSITRR 471
ERO1-l      FSGNLDPEGNQNLDLPAVRHTKFQLSRIEIVALINAFGRLSSSIMALENFRQSITRR 477
            *********************************************************
```

Figure 3.59 Pairwise alignments of deduced amino acid sequences of the short and long forms of *P. monodon Ero1 isoform B.*

```
ERO1-s          CTAATTACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAGTCAT 60
ERO1-l          CTAATTACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGGAGTCAT 60
                ************************************************************
ERO1-s          GACAGTTACGTCAGAGTAGTAGTATATATCGTCTACACTCCAGTCCCGGGATTCAGTATC 120
ERO1-l          GACAGTTACGTCAGAGTAGTAGTATATATCGTCTACACTCCAGTCCCGGGATTCAGTATC 120
                ************************************************************
ERO1-s          TGTTCTCTCGAGTGTCGCGAGTGTTTAAGAGCGAGTCATGATGTGGTCGAGGACTTCCGC 180
ERO1-l          TGTTCTCTCGAGTGTCGCGAGTGTTTAAGAGCGAGTCATGATGTGGTCGAGGACTTCCGC 180
                ************************************************************
ERO1-s          GTGGGTGTTTGTCATTTTGGCAATTTCCTTTCACACTTGCACGGCGATTTGGTATGGGAT 240
ERO1-l          GTGGGTGTTTGTCATTTTGGCAATTTCCTTTCACACTTGCACGGCGATTTGGTATGGGAT 240
                ************************************************************
ERO1-s          TAAGAAGACCCCAAACGGAGACACAACTGCTGGTGAGAGGTGCTTCTGTCAGTTAAAAGG 300
ERO1-l          TAAGAAGACCCCAAACGGAGACACAACTGCTGGTGAGAGGTGCTTCTGTCAGTTAAAAGG 300
                ************************************************************
ERO1-s          AGTAATAGATGACTGTTCCTGTTCTGTGGAGACTTTAGACAGCTTCAACAACCTGAAGCT 360
ERO1-l          AGTAATAGATGACTGTTCCTGTTCTGTGGAGACTTTAGACAGCTTCAACAACCTGAAGCT 360
                ************************************************************
ERO1-s          GTATCCACGCCTGAATAGCCTCCTGCAGTATGACTACTTCAGGTACTGGAAGGTGAACTT 420
ERO1-l          GTATCCACGCCTGAATAGCCTCCTGCAGTATGACTACTTCAGGTACTGGAAGGTGAACTT 420
                ************************************************************
```

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

```
ERO1-s          GAAAAAAGAATGTCCTTTTTGGGAAGATGACAGCAAGTGTGCCATTCGTTACTGCAGTGT 480
ERO1-l          GAAAAAAGAATGTCCTTTTTGGGAAGATGACAGCAAGTGTGCCATTCGTTACTGCAGTGT 480
                ************************************************************

ERO1-s          GAAGCCATGTACTGATGTCCCAGAGGGTATAAAGGGAGCTTCCATAGACAAAATTGAGAA 540
ERO1-l          GAAGCCATGTACTGATGTCCCAGAGGGTATAAAGGGAGCTTCCATAGACAAAATTGAGAA 540
                ************************************************************

ERO1-s          GGAAAAGAAGGAAAAGTCCCACATGGTGACTGGACATTGTGATGGAGAGAATGACCTTGG 600
ERO1-l          GGAAAAGAAGGAAAAGTCCCACATGGTGACTGGACATTGTGATGGAGAGAATGACCTTGG 600
                ************************************************************

ERO1-s          ATACTTGAATACTACTCTCAGCAAGGAATCCAAAGTTGGTTTTAAGCGCTGGGCAGCCCA 660
ERO1-l          ATACTTGAATACTACTCTCAGCAAGGAATCCAAAGTTGGTTTTAAGCGCTGGGCAGCCCA 660
                ************************************************************

ERO1-s          TGATGATGCACAGCTGAACTTTTGCAAAATTGATGACGACAGTTCTGAGGATAGTGAATA 720
ERO1-l          TGATGATGCACAGCTGAACTTTTGCAAAATTGATGACGACAGTTCTGAGGATAGTGAATA 720
                ************************************************************

ERO1-s          TGTTGATCTATTGCTGAATCCTGAGCGGTACACAGGTTATGCAGGACCTTCAGCACATAG 780
ERO1-l          TGTTGATCTATTGCTGAATCCTGAGCGGTACACAGGTTATGCAGGACCTTCAGCACATAG 780
                ************************************************************

ERO1-s          AATATGGAGAACAATATACCAAGAAAATTGCTTTAAGCCATCTAGGGCAATTGGCCGTTA 840
ERO1-l          AATATGGAGAACAATATACCAAGAAAATTGCTTTAAGCCATCTAGGGCAATTGGCCGTTA 840
                ************************************************************

ERO1-s          CACAGACTTCAGTAGTATTGGAG------------------AAATGTGCTTGGAAAAAAG 882
ERO1-l          CACAGACTTCAGTAGTATTGGAGCCAATGATTTTCTGAAAGAAATGTGCTTGGAAAAAAG 900
                ***********************                  *******************

ERO1-s          AACATTTTACAGAGCTATTTCCGGTCTCCATACCAGTATCAACATTCACCTGAGTGCTAA 942
ERO1-l          AACATTTTACAGAGCTATTTCCGGTCTCCATACCAGTATCAACATTCACCTGAGTGCTAA 960
                ************************************************************

ERO1-s          CTACCTCCTGTCAGACCAGAATGGCTTTGAAATGTCAAAGGATGGCCAGTGGGGTCCGAA 1002
ERO1-l          CTACCTCCTGTCAGACCAGAATGGCTTTGAAATGTCAAAGGATGGCCAGTGGGGTCCGAA 1020
                ************************************************************

ERO1-s          TGTACAGGAATTCCAGACGAGGTTTGACCCAGAACTAACTGGCGGAGAGGGAACCCACCG 1062
ERO1-l          TGTACAGGAATTCCAGACGAGGTTTGACCCAGAACTAACTGGCGGAGAGGGAACCCACCG 1080
                ************************************************************

ERO1-s          CCTGAAGAACCTCTACTTTGTATACCTTCTTGAACTAAGAGCTTTAGCTAAAGCTGCACC 1122
ERO1-l          CCTGAAGAACCTCTACTTTGTATACCTTCTTGAACTAAGAGCTTTAGCTAAAGCTGCACC 1140
                ************************************************************

ERO1-s          ATACCTTGAAAGCTTAGAATACTACACAGGAAATGAAAATGAAGATAATGATGTATCAAA 1182
ERO1-l          ATACCTTGAAAGCTTAGAATACTACACAGGAAATGAAAATGAAGATAATGATGTATCAAA 1200
                ************************************************************

ERO1-s          GGCAGTTAAAGACTTATTAACTGTTGTTAAGAGTTTTCCAGAGCACTTTGATGAAAGCTC 1242
ERO1-l          GGCAGTTAAAGACTTATTAACTGTTGTTAAGAGTTTTCCAGAGCACTTTGATGAAAGCTC 1260
                ************************************************************

ERO1-s          CATGTTCTCTGGCGGCCAGCATGCAGCTAAGTTGAAAGAGGAATTTCGGCAACATTTTTG 1302
ERO1-l          CATGTTCTCTGGCGGCCAGCATGCAGCTAAGTTGAAAGAGGAATTTCGGCAACATTTTTG 1320
                ************************************************************

ERO1-s          GAATGTGTCTCGTATTATGGACTGCGTTGGATGTGACAAGTGTCGTCTTTGGGGCAAGCT 1362
ERO1-l          GAATGTGTCTCGTATTATGGACTGCGTTGGATGTGACAAGTGTCGTCTTTGGGGCAAGCT 1380
                ************************************************************

ERO1-s          GCAGGTAACTGGCCTTGGTACAGCACTCAAGATTCTGTTTTCAGGCAATCTAGACCCAGA 1422
ERO1-l          GCAGGTAACTGGCCTTGGTACAGCACTCAAGATTCTGTTCTCAGGCAATCTAGACCCAGA 1440
                *************************************** ********************

ERO1-s          AGGTAACCAAAACTTAGATCTTCCAGCCGTAAGACACACAAAGTTCCAGCTGTCTCGTAT 1482
ERO1-l          AGGTAACCAAAACTTAGATCTTCCAGCCGTGAGACACACAAAGTTCCAGCTGTCTCGTAT 1500
                ****************************** *****************************

ERO1-s          TGAGATTGTTGCTCTTATCAATGCCTTTGGTCGGTTGTCCAGCAGCATCATGGCGCTAGA 1542
ERO1-l          TGAGATTGTTGCTCTTATCAATGCCTTTGGTCGGTTGTCCAGCAGCATCATGGCGCTAGA 1560
                ************************************************************

ERO1-s          GAACTTTCGGCAGAGCATAACCAGAAGATAATGACAAGATTGGCTACTTTGAGCTGAACT 1602
ERO1-l          GAACTTTCGGCAGAGCATAACCAGAAGATAATGACAAGATTGGCTACTTTGAGCTGAACT 1620
                ************************************************************

ERO1-s          TCTGTGATATTTGTTATATGTTACCCAAATACATAGTTGTATGCTAGATTATGACAGGTG 1662
ERO1-l          TCTGTGATATTTGTTATATGTTACCCAAATACATAGTTGTATGCTAGATTATGACAGGTG 1680
                ************************************************************

ERO1-s          GGCTGCCAACTGCCAGTGGTAGCATCATATATTCAAAGAAGGAACATGACAGATGTAAGA 1722
ERO1-l          GGCTGCCAACTGCCAGTGGTAGCATCATATATTCAAAGAAGGAACATGACAGATGTAAGA 1740
                ************************************************************

ERO1-s          CAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1756
ERO1-l          CAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1774
                **********************************
```

Figure 3.60 Pairwise alignments of different isoforms of *Ero1, isoform B* of *P. monodon.*

```
TTCCTCTGAGGAGTAAGAGCCGCCAGCTGATCGAGAAAATGCAAGCGAACATTTGGACTC 60
                                      M  Q  A  N  I  W  T  R 8
GAATTTCTCACATTTCGAAGGTGCCCTTGGCCAGGTGTCCAGGTGCATCTGTGGGCCTCT 120
 I  S  H  I  S  K  V  P  L  A  R  C  P  G  A  S  V  G  L  S  28
CCTCTGTGGGACAACGTCGCCATGCATCCCATGAAGCCGACCTTGTGGTCATTGGATCAG 180
 S  V  G  Q  R  R  H  A  S  H  E  A  D  L  V  V  I  G  S  G  48
GACCAGGGGGCTACGTAGCTGCCATCAAGGCTGCCCAGCTGGGAATGAAGACTGTATGCG 240
 P  G  G  Y  V  A  A  I  K  A  A  Q  L  G  M  K  T  V  C  V  68
TGGAGAAGAATGCAACATATGGTGGCACCTGCCTAAATGTGGGCTGTATCCCCTCCAAGT 300
 E  K  N  A  T  Y  G  G  T  C  L  N  V  G  C  I  P  S  K  S  88
CGCTGCTCAACAATTCCCATTACTATCATATGGCCAAAGGAAAGGAGTTTGCTGACCGAG 360
 L  L  N  N  S  H  Y  Y  H  M  A  K  G  K  E  F  A  D  R  G  108
GCATTGAGGTTGACAATGTGAGGCTTAACTTGGACAAGCTGATGGGAGCCAAAGAGAAGG 420
 I  E  V  D  N  V  R  L  N  L  D  K  L  M  G  A  K  E  K  A  128
CAGTGAAGGCACTCACTGGTGGCATTGCTCATCTCTTTAAGAACAACAAGATTGTTGGGC 480
 V  K  A  L  T  G  G  I  A  H  L  F  K  N  N  K  I  V  G  L  148
TCAGTGGCCATGGCAAGATCACAGGGCCCAATGAAGTGACCGTCCTCAAAGAAGACGGCT 540
 S  G  H  G  K  I  T  G  P  N  E  V  T  V  L  K  E  D  G  S  168
CTAATGACACTGTCAAGACCAAGAACATTCTGATTGCCACTGGCTCTGAGGTTACTCCCT 600
 N  D  T  V  K  T  K  N  I  L  I  A  T  G  S  E  V  T  P  F  188
TCCCAGGCATCCCTGTAGATGAGGAGCAGATTGTATCCTCCACTGGTGCGCTGAAGCTCA 660
 P  G  I  P  V  D  E  E  Q  I  V  S  S  T  G  A  L  K  L  K  208
AGAGTGTTCCTGAGAAGTTGATTCTCATTGGGGCTGGTGTCATTGGCCTTGAGCTTGGAT 720
 S  V  P  E  K  L  I  L  I  G  A  G  V  I  G  L  E  L  G  S  228
CTGTGTGGTCACGTCTTGGAGCCCATGTGACAGCGGTAGAGTTTCTGGGCTCCATTGGTG 780
 V  W  S  R  L  G  A  H  V  T  A  V  E  F  L  G  S  I  G  G  248
GCTTGGGGATTGATGCAGAGATCTCAAAGAACTTCCAGCGAATCCTAACCAAGCAGGGAC 840
 L  G  I  D  A  E  I  S  K  N  F  Q  R  I  L  T  K  Q  G  L  268
TCAGGTTCAAGCTGAGCACAAAGGTGATGAGTGCCAGCAAGCAAGGCGACAAGATCATGG 900
 R  F  K  L  S  T  K  V  M  S  A  S  K  Q  G  D  K  I  M  V  288
TCTCTGTTGAGGGAGTCAAAAATGGAAAGAAAGAGGAGCTTGAATGTGACACCCTCCTTG 960
 S  V  E  G  V  K  N  G  K  K  E  E  L  E  C  D  T  L  L  V  308
TCTGTGTGGGACGACGACCCTACACCACCAACCTTGGCCTGGAGGAGCTTGGGATTGAGA 1020
 C  V  G  R  R  P  Y  T  T  N  L  G  L  E  E  L  G  I  E  K  328
AAGACGAGAAAGGTCGCATTCCTGTCAACTCTCGCTTCCAGACCATCATCCCCAATATCT 1080
 D  E  K  G  R  I  P  V  N  S  R  F  Q  T  I  I  P  N  I  F  348
TTGCCATTGGGGACTGCATTCATGGCCCCATGCTGGCCCACAAGGCAGAAGATGAGGGCA 1140
 A  I  G  D  C  I  H  G  P  M  L  A  H  K  A  E  D  E  G  I  368
TTGTGTGTGTAGAGGGCATTGCTGGTGGCCCTGTCCACATCGACTACAACTGTGTACCAT 1200
 V  C  V  E  G  I  A  G  G  P  V  H  I  D  Y  N  C  V  P  S  388
CTGTTATCTACACTCATCCTGAGGTGGCCTGGGTTGGCAAGACAGAGGAAGACCTGAAGG 1260
 V  I  Y  T  H  P  E  V  A  W  V  G  K  T  E  E  D  L  K  A  408
CTGAGGGTGTGGAGTATGCAGTTGGCAAGTTCCCATTTGCAGCCAATTCTCGTGCTAAGT 1320
 E  G  V  E  Y  A  V  G  K  F  P  F  A  A  N  S  R  A  K  C  428
GTAATGACGACACTGATGGCCTGGTCAAGATCTTGGCAGACAAGCACACAGATCGGCTGT 1380
 N  D  D  T  D  G  L  V  K  I  L  A  D  K  H  T  D  R  L  L  448
TGGGCGCACACATCATTGGTCCAGGTGCAGGCGAGATGATCAATGAAGCAGCATTGGCCA 1440
 G  A  H  I  I  G  P  G  A  G  E  M  I  N  E  A  A  L  A  M  468
TGGAGTACGGTGCTAGTTGTGAGGATGTAGCGCGTGTATGCCATGCCCACCCCACCTGCT 1500
 E  Y  G  A  S  C  E  D  V  A  R  V  C  H  A  H  P  T  C  S  488
CAGAGGCCTTCCGTGAGGCTAACCTGGCTGCATACTTCGGAAAGCCCATCAACTTCTAAT 1560
 E  A  F  R  E  A  N  L  A  A  Y  F  G  K  P  I  N  F  *     506
GAATAGGCCGTTATTTATAAAAGAAACGATAAAAAATAACAAGATATATGTATTGTCTAT 1620
ATTTTTTCCGTGCTTTTAGTCTTGTTTAGGGATGAGGATCTTTGATGCATTTCATGAGCT 1680
ATTTGAACCCATTCTTTCTTTCTCTTTTCTTTCTTTTACACACTTAGGTAGCACAATCAT 1740
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA                               1770
```

Figure3.61 The full length cDNA sequences of the short form of *dihydrolipoamide dehydrogenase* (*DLADH-s*, 1770 bp; ORF 1521 bp, 506 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A pyridine nucleotide-disulphide oxidoreductase domain (Pyr_redox domain; positions 213th - 311th), a pyridine nucleotide-disulphide oxidoreductase 2 domain (Pyr_redox 2 domain;

positions 41th - 358th), and a pyridine nucleotide-disulphide oxidoreductase, dimerisation domain (Pyr_redox_dim domain; positions 386th-495th of the deduced protein) are highlighted.

```
TTCCTCTGAGGAGTAAGAGCCGCCAGCTGATCGAGAAAATGCAAGCGAACATTTGGACTC 60
                                      M  Q  A  N  I  W  T  R 8
GAATTTCTCACATTTCGAAGGTGCCCTTGGCCAGGTGTCCAGGTGCATCTGTGGGCCTCT 120
  I  S  H  I  S  K  V  P  L  A  R  C  P  G  A  S  V  G  L  S 28
CCTCTGTGGGACAACGTCGCCATGCATCCCATGAAGCCGACCTTGTGGTCATTGGATCAG 180
  S  V  G  Q  R  R  H  A  S  H  E  A  D  L  V  V  I  G  S  G 48
GACCAGGGGGCTACGTAGCTGCCATCAAGGCTGCCCAGCTGGGAATGAAGACTGTATGCG 240
  P  G  G  Y  V  A  A  I  K  A  A  Q  L  G  M  K  T  V  C  V 68
TGGAGAAGAATGCAACATATGGTGGCACCTGCCTAAATGTGGGCTGTATCCCCTCCAAGT 300
  E  K  N  A  T  Y  G  G  T  C  L  N  V  G  C  I  P  S  K  S 88
CGCTGCTCAACAATTCCCATTACTATCATATGGCCAAAGGAAAGGAGTTTGCTGACCGAG 360
  L  L  N  N  S  H  Y  Y  H  M  A  K  G  K  E  F  A  D  R  G 108
GCATTGAGGTTGACAATGTGAGGCTTAACTTGGACAAGCTGATGGGAGCCAAAGAGAAGG 420
  I  E  V  D  N  V  R  L  N  L  D  K  L  M  G  A  K  E  K  A 128
CAGTGAAGGCACTCACTGGTGGCATTGCTCATCTCTTTAAGAACAACAAGATTGTTGGGC 480
  V  K  A  L  T  G  G  I  A  H  L  F  K  N  N  K  I  V  G  L 148
TCAGTGGCCATGGCAAGATCACAGGGCCCAATGAAGTGACCGTCCTCAAAGAAGACGGCT 540
  S  G  H  G  K  I  T  G  P  N  E  V  T  V  L  K  E  D  G  S 168
CTAATGACACTGTCAAGACCAAGAACATTCTGATTGCCACTGGCTCTGAGGTTACTCCCT 600
  N  D  T  V  K  T  K  N  I  L  I  A  T  G  S  E  V  T  P  F 188
TCCCAGGCATCCCTGTAGATGAGGAGCAGATTGTATCCTCCACTGGTGCGCTGAAGCTCA 660
  P  G  I  P  V  D  E  E  Q  I  V  S  S  T  G  A  L  K  L  K 208
AGAGTGTTCCTGAGAAGTTGATTCTCATTGGGGCTGGTGTCATTGGCCTTGAGCTTGGAT 720
  S  V  P  E  K  L  I  L  I  G  A  G  V  I  G  L  E  L  G  S 228
CTGTGTGGTCACGTCTTGGAGCCCATGTGACAGCGGTAGAGTTTCTGGGCTCCATTGGTG 780
  V  W  S  R  L  G  A  H  V  T  A  V  E  F  L  G  S  I  G  G 248
GCTTGGGGATTGATGCAGAGATCTCAAAGAACTTCCAGCGAATCCTAACCAAGCAGGGAC 840
  L  G  I  D  A  E  I  S  K  N  F  Q  R  I  L  T  K  Q  G  L 268
TCAGGTTCAAGCTGAGCACAAAGGTGATGAGTGCCAGCAAGCAAGGCGACAAGATCATGG 900
  R  F  K  L  S  T  K  V  M  S  A  S  K  Q  G  D  K  I  M  V 288
TCTCTGTTGAGGGAGTCAAAAATGGAAAGAAAGAGGAGCTTGAATGTGACACCCTCCTTG 960
  S  V  E  G  V  K  N  G  K  K  E  E  L  E  C  D  T  L  L  V 308
TCTGTGTGGGACGACGACCCTACACCACCAACCTTGGCCTGGAGGAGCTTGGGATTGAGA 1020
  C  V  G  R  R  P  Y  T  T  N  L  G  L  E  E  L  G  I  E  K 328
AAGACGAGAAAGGTCGCATTCCTGTCAACTCTCGCTTCCAGACCATCATCCCCAATATCT 1080
  D  E  K  G  R  I  P  V  N  S  R  F  Q  T  I  I  P  N  I  F 348
TTGCCATTGGGGACTGCATTCATGGCCCCATGCTGGCCCACAAGGCAGAAGATGAGGGCA 1140
  A  I  G  D  C  I  H  G  P  M  L  A  H  K  A  E  D  E  G  I 368
TTGTGTGTGTAGAGGGCATTGCTGGTGGCCCTGTCCACATCGACTACAACTGTGTACCAT 1200
  V  C  V  E  G  I  A  G  G  P  V  H  I  D  Y  N  C  V  P  S 388
CTGTTATCTACACTCATCCTGAGGTGGCCTGGGTTGGCAAGACAGAGGAAGACCTGAAGG 1260
  V  I  Y  T  H  P  E  V  A  W  V  G  K  T  E  E  D  L  K  A 408
CTGAGGGTGTGGAGTATGCAGTTGGCAAGTTCCCATTTGCAGCCAATTCTCGTGCTAAGT 1320
  E  G  V  E  Y  A  V  G  K  F  P  F  A  A  N  S  R  A  K  C 428
GTAATGACGACACTGATGGCCTGGTCAAGATCTTGGCAGACAAGCACACAGATCGGCTGT 1380
  N  D  D  T  D  G  L  V  K  I  L  A  D  K  H  T  D  R  L  L 448
TGGGCGCACACATCATTGGTCCAGGTGCAGGCGAGATGATCAATGAAGCAGCATTGGCCA 1440
  G  A  H  I  I  G  P  G  A  G  E  M  I  N  E  A  A  L  A  M 468
TGGAGTACGGTGCTAGTTGTGAGGATGTAGCGCGTGTATGCCATGCCCACCCCACCTGCT 1500
  E  Y  G  A  S  C  E  D  V  A  R  V  C  H  A  H  P  T  C  S 488
CAGAGGCCTTCCGTGAGGCTAACCTGGCTGCATACTTCGGAAAGCCCATCAACTTCTAAT 1560
  E  A  F  R  E  A  N  L  A  A  Y  F  G  K  P  I  N  F  *    506
GAATAGGCCGTTATTTATAAAAGAAACGATAAAAAATAACAAGATATATGTATTGTCTAT 1620
ATTTTTTCCGTGCTTTTAGTCTTGTTTAGGGATGAGGATCTTTGATGCATTTCATGAGCT 1680
ATTTGAACCCATTCCTTCTTTCTCTTTTCTTTCTTTTACACACCTAGGTAGCACAATCCT 1740
GAAAAAAAAGAGATAAAAATTTTGGCGGGGGATTTACTGTTCAGTTCTTTTCAGAGTAAA 1800
ATTTATTTTTATTTATCTTGCAATGAAGTTTGTTGACATTTAATAAATTATTTCTCTGTC 1860
ACTTTTTCTTGCGGTCTTTTCTAACCCAAAGATGTAATGCATCATATCTGTACATAGTTG 1920
GGCATATGAAACTTGTCGTATGTGGAAGATGCAATGTAAGTATATTATACTGGGGGAGAA 1980
```

```
TTTCTCAATTTGTAAGTAAAACCCACTGTCTAAAAAAACCAAAAAAAAAAAAAAAAAAAA 2040
AAAAAAAAAA                                                   2050
```

Figure3.62 The full length cDNA sequences of the long form of *dihydrolipoamide dehydrogenase* (*DLADH-l*, 2050 bp; ORF 1521 bp, 506 aa) of *P. monodon*. Start and stop codons were illustrated in boldface and underlined. A pyridine nucleotide-disulphide oxidoreductase domain (Pyr_redox domain; positions 213th -311th), a pyridine nucleotide-disulphide oxidoreductase 2 domain (Pyr_redox 2 domain; positions 41th -358th), and a pyridine nucleotide-disulphide oxidoreductase, dimerisation domain (Pyr_redox_dim domain; positions 386th -495th of the deduced protein) are highlighted.

```
DLADH-s      TTCCTCTGAGGAGTAAGAGCCGCCAGCTGATCGAGAAAATGCAAGCGAACATTTGGACTC 60
DLADH-l      TTCCTCTGAGGAGTAAGAGCCGCCAGCTGATCGAGAAAATGCAAGCGAACATTTGGACTC 60
             ************************************************************
DLADH-s      GAATTTCTCACATTTCGAAGGTGCCCTTGGCCAGGTGTCCAGGTGCATCTGTGGGCCTCT 120
DLADH-l      GAATTTCTCACATTTCGAAGGTGCCCTTGGCCAGGTGTCCAGGTGCATCTGTGGGCCTCT 120
             ************************************************************
DLADH-s      CCTCTGTGGGACAACGTCGCCATGCATCCCATGAAGCCGACCTTGTGGTCATTGGATCAG 180
DLADH-l      CCTCTGTGGGACAACGTCGCCATGCATCCCATGAAGCCGACCTTGTGGTCATTGGATCAG 180
             ************************************************************
DLADH-s      GACCAGGGGGCTACGTAGCTGCCATCAAGGCTGCCCAGCTGGGAATGAAGACTGTATGCG 240
DLADH-l      GACCAGGGGGCTACGTAGCTGCCATCAAGGCTGCCCAGCTGGGAATGAAGACTGTATGCG 240
             ************************************************************
DLADH-s      TGGAGAAGAATGCAACATATGGTGGCACCTGCCTAAATGTGGGCTGTATCCCCTCCAAGT 300
DLADH-l      TGGAGAAGAATGCAACATATGGTGGCACCTGCCTAAATGTGGGCTGTATCCCCTCCAAGT 300
             ************************************************************
DLADH-s      CGCTGCTCAACAATTCCCATTACTATCATATGGCCAAAGGAAAGGAGTTTGCTGACCGAG 360
DLADH-l      CGCTGCTCAACAATTCCCATTACTATCATATGGCCAAAGGAAAGGAGTTTGCTGACCGAG 360
             ************************************************************
DLADH-s      GCATTGAGGTTGACAATGTGAGGCTTAACTTGGACAAGCTGATGGGAGCCAAAGAGAAGG 420
DLADH-l      GCATTGAGGTTGACAATGTGAGGCTTAACTTGGACAAGCTGATGGGAGCCAAAGAGAAGG 420
             ************************************************************
DLADH-s      CAGTGAAGGCACTCACTGGTGGCATTGCTCATCTCTTTAAGAACAACAAGATTGTTGGGC 480
DLADH-l      CAGTGAAGGCACTCACTGGTGGCATTGCTCATCTCTTTAAGAACAACAAGATTGTTGGGC 480
             ************************************************************
DLADH-s      TCAGTGGCCATGGCAAGATCACAGGGCCCAATGAAGTGACCGTCCTCAAAGAAGACGGCT 540
DLADH-l      TCAGTGGCCATGGCAAGATCACAGGGCCCAATGAAGTGACCGTCCTCAAAGAAGACGGCT 540
             ************************************************************
DLADH-s      CTAATGACACTGTCAAGACCAAGAACATTCTGATTGCCACTGGCTCTGAGGTTACTCCCT 600
DLADH-l      CTAATGACACTGTCAAGACCAAGAACATTCTGATTGCCACTGGCTCTGAGGTTACTCCCT 600
             ************************************************************
DLADH-s      TCCCAGGCATCCCTGTAGATGAGGAGCAGATTGTATCCTCCACTGGTGCGCTGAAGCTCA 660
DLADH-l      TCCCAGGCATCCCTGTAGATGAGGAGCAGATTGTATCCTCCACTGGTGCGCTGAAGCTCA 660
             ************************************************************
DLADH-s      AGAGTGTTCCTGAGAAGTTGATTCTCATTGGGGCTGGTGTCATTGGCCTTGAGCTTGGAT 720
DLADH-l      AGAGTGTTCCTGAGAAGTTGATTCTCATTGGGGCTGGTGTCATTGGCCTTGAGCTTGGAT 720
             ************************************************************
DLADH-s      CTGTGTGGTCACGTCTTGGAGCCCATGTGACAGCGGTAGAGTTTCTGGGCTCCATTGGTG 780
DLADH-l      CTGTGTGGTCACGTCTTGGAGCCCATGTGACAGCGGTAGAGTTTCTGGGCTCCATTGGTG 780
             ************************************************************
DLADH-s      GCTTGGGGATTGATGCAGAGATCTCAAAGAACTTCCAGCGAATCCTAACCAAGCAGGGAC 840
DLADH-l      GCTTGGGGATTGATGCAGAGATCTCAAAGAACTTCCAGCGAATCCTAACCAAGCAGGGAC 840
             ************************************************************
DLADH-s      TCAGGTTCAAGCTGAGCACAAAGGTGATGAGTGCCAGCAAGCAAGGCGACAAGATCATGG 900
DLADH-l      TCAGGTTCAAGCTGAGCACAAAGGTGATGAGTGCCAGCAAGCAAGGCGACAAGATCATGG 900
             ************************************************************
DLADH-s      TCTCTGTTGAGGGAGTCAAAAATGGAAAGAAAGAGGAGCTTGAATGTGACACCCTCCTTG 960
DLADH-l      TCTCTGTTGAGGGAGTCAAAAATGGAAAGAAAGAGGAGCTTGAATGTGACACCCTCCTTG 960
             ************************************************************
DLADH-s      TCTGTGTGGGACGACGACCCTACACCACCAACCTTGGCCTGGAGGAGCTTGGGATTGAGA 1020
DLADH-l      TCTGTGTGGGACGACGACCCTACACCACCAACCTTGGCCTGGAGGAGCTTGGGATTGAGA 1020
             ************************************************************
```

```
DLADH-s     AAGACGAGAAAGGTCGCATTCCTGTCAACTCTCGCTTCCAGACCATCATCCCCAATATCT 1080
DLADH-l     AAGACGAGAAAGGTCGCATTCCTGTCAACTCTCGCTTCCAGACCATCATCCCCAATATCT 1080
            ************************************************************
DLADH-s     TTGCCATTGGGGACTGCATTCATGGCCCCATGCTGGCCCACAAGGCAGAAGATGAGGGCA 1140
DLADH-l     TTGCCATTGGGGACTGCATTCATGGCCCCATGCTGGCCCACAAGGCAGAAGATGAGGGCA 1140
            ************************************************************
DLADH-s     TTGTGTGTGTAGAGGGCATTGCTGGTGGCCCTGTCCACATCGACTACAACTGTGTACCAT 1200
DLADH-l     TTGTGTGTGTAGAGGGCATTGCTGGTGGCCCTGTCCACATCGACTACAACTGTGTACCAT 1200
            ************************************************************
DLADH-s     CTGTTATCTACACTCATCCTGAGGTGGCCTGGGTTGGCAAGACAGAGGAAGACCTGAAGG 1260
DLADH-l     CTGTTATCTACACTCATCCTGAGGTGGCCTGGGTTGGCAAGACAGAGGAAGACCTGAAGG 1260
            ************************************************************
DLADH-s     CTGAGGGTGTGGAGTATGCAGTTGGCAAGTTCCCATTTGCAGCCAATTCTCGTGCTAAGT 1320
DLADH-l     CTGAGGGTGTGGAGTATGCAGTTGGCAAGTTCCCATTTGCAGCCAATTCTCGTGCTAAGT 1320
            ************************************************************
DLADH-s     GTAATGACGACACTGATGGCCTGGTCAAGATCTTGGCAGACAAGCACACAGATCGGCTGT 1380
DLADH-l     GTAATGACGACACTGATGGCCTGGTCAAGATCTTGGCAGACAAGCACACAGATCGGCTGT 1380
            ************************************************************
DLADH-s     TGGGCGCACACATCATTGGTCCAGGTGCAGGCGAGATGATCAATGAAGCAGCATTGGCCA 1440
DLADH-l     TGGGCGCACACATCATTGGTCCAGGTGCAGGCGAGATGATCAATGAAGCAGCATTGGCCA 1440
            ************************************************************
DLADH-s     TGGAGTACGGTGCTAGTTGTGAGGATGTAGCGCGTGTATGCCATGCCCACCCCACCTGCT 1500
DLADH-l     TGGAGTACGGTGCTAGTTGTGAGGATGTAGCGCGTGTATGCCATGCCCACCCCACCTGCT 1500
            ************************************************************
DLADH-s     CAGAGGCCTTCCGTGAGGCTAACCTGGCTGCATACTTCGGAAAGCCCATCAACTTCTAAT 1560
DLADH-l     CAGAGGCCTTCCGTGAGGCTAACCTGGCTGCATACTTCGGAAAGCCCATCAACTTCTAAT 1560
            ************************************************************
DLADH-s     GAATAGGCCGTTATTTATAAAAGAAACGATAAAAAATAACAAGATATATGTATTGTCTAT 1620
DLADH-l     GAATAGGCCGTTATTTATAAAAGAAACGATAAAAAATAACAAGATATATGTATTGTCTAT 1620
            ************************************************************
DLADH-s     ATTTTTTCCGTGCTTTTAGTCTTGTTTAGGGATGAGGATCTTTGATGCATTTCATGAGCT 1680
DLADH-l     ATTTTTTCCGTGCTTTTAGTCTTGTTTAGGGATGAGGATCTTTGATGCATTTCATGAGCT 1680
            ************************************************************
DLADH-s     ATTTGAACCCATTCTTTCTTTCTCTTTTCTTTCTTTTACACACTTAGGTAGCACAATCAT 1740
DLADH-l     ATTTGAACCCATTCCTTCTTTCTCTTTTCTTTCTTTTACACACCTAGGTAGCACAATCCT 1740
            ************** *************************** *************** *
DLADH-s     AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA------------------------------ 1770
DLADH-l     GAAAAAAAAGAGATAAAAATTTTGGCGGGGGATTTACTGTTCAGTTCTTTTCAGAGTAAA 1800
             ******** * * *****

DLADH-s     ------------------------------------------------------------
DLADH-l     ATTTATTTTTATTTATCTTGCAATGAAGTTTGTTGACATTTAATAAATTATTTCTCTGTC 1860


DLADH-s     ------------------------------------------------------------
DLADH-l     ACTTTTTCTTGCGGTCTTTTCTAACCCAAAGATGTAATGCATCATATCTGTACATAGTTG 1920


DLADH-s     ------------------------------------------------------------
DLADH-l     GGCATATGAAACTTGTCGTATGTGGAAGATGCAATGTAAGTATATTATACTGGGGGAGAA 1980


DLADH-s     ------------------------------------------------------------
DLADH-l     TTTCTCAATTTGTAAGTAAAACCCACTGTCTAAAAAAACCAAAAAAAAAAAAAAAAAAAA 2040


DLADH-s     ----------
DLADH-l     AAAAAAAAAA 2050
```

Figure 3.63 Pairwise alignments of different isoforms of *Dihydrolipoamide dehydrogenase* of *P. monodon*

Table 3.9 Summary of the full lengths of gene homologues from testis of *P. monodon* using RACE-PCR

Clone No.	Transcripts	Species	E-value	Full length
TT-N-S01-0071-W	*PMTST1*	-	-	2201bp and 2413 bp (ORF 112 amino acid, 339 bp)
TT-N-S01-0004-W	*Multiple inositol polyphosphate phosphatase*	*Nasonia vitripennis*	9e-35	1644 bp (ORF 457 amino acids, 1374 bp)*
TT-N-S01-0678-W	*Multiple inositol polyphosphate phosphatase 2*	*Apis mellifera*	4e-31	1746 bp (ORF 397 amino acids, 1194 bp)
TT-N-S01-0090-W	*Prohibitin-2 (a repressor of estrogen receptor activity)*	*Tribolium castaneum*	2e-128	1382 bp (ORF 296 amino acids, 891 bp)
TT-N-S01-0169-W	*Cdk7*	*Drosophila melanogaster*	1e-122	1431 bp (ORF 353 amino acids, 1062 bp)
TT-N-S01-0189-W	*Growth factor receptor-bound protein*	*Aedes aegypti*	5e-91	1188 bp and 1883 bp (ORF 211 amino acids, 636 bp)
TT-N-S01-0246-W	*Innexin 1*	*Schistocerca americana*	6e-120	2505 bp (ORF 380 amino acids, 1143 bp)
TT-N-S01-0985-W	*Splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila)*	*Apis mellifera*	1e-40	2673 and 2658 bp (ORF 267 and 262 amino acids, 804 and 789 bp)
TT-N-S01-1036-W	*Rac GTPase activating protein 1*	*Apis mellifera*	3e-99	2838 bp (ORF 603 amino acids, 1812 bp)
TT-N-S01-1055-W	*Flotillin 2*	*Drosophila melanogaster*	0.0	1937 bp (ORF 439 amino acids, 1320 bp)
TT-N-ST01-0019-W	*Meiotic recombination protein DMC1/LIM15 homolog isoform 1*	*Canis familiaris*	1e-146	1661 bp (ORF 341 amino acids, 1026 bp)
TT-N-ST02-0009-W	*Innexin 2*	*Homarus gammarus*	7e-161	1651 bp (ORF 358 amino acids, 1077 bp)
TT-N-ST02-0022-W	*Progestin receptor membrane component 1*	*Oryzias latipes*	1e-41	2028, 2896, and 3019 bp (ORF 190 amino acids, 573 bp)
HC-H-S01-0460-LF	*Saposin*	*Aedes aegypti*	5e-140	3034 bp (ORF 862 amino acids, 2589 bp)
HT-N-S01-0128-W HT-N-S01-0200-W	*TPA_inf: troponin T isoform 3*	*Drosophila pseudoobscura*	3e-133	1430 bp (ORF 368 amino acids, 1107 bp)
HT-N-S01-0244-W	*Ero1L CG1333-PB, isoform B*	*Apis mellifera*	8e-145 4e-143	1756 and 1774 bp (ORF 470 and 477 amino acids, 1413 and 1434 bp)
HT-N-S01-0382-W	*Dihydrolipoamide dehydrogenase*	*Mus musculus*	0.0	1770 and 2050 bp (ORF 506 amino acids, 1521 bp)

* Full length cDNA not including poly A

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Phylogenetic analysis indicated that the newly identified *PMTra-2* was allocated in an invertebrate *Tra2* group but it is distantly diverse from other invertebrate *Tra2*. Therefore, *PMTra-2* should be regarded as a new subgroup of invertebrate *Tra-2* (Fig. 3.64).

Phylogenetic analysis also indicated that vertebrate *Dmc1* is conserved and can be allocated to 2 different groups. Nevertheless, *Dmc1* is more diverse in invertebrate species and *Dmc1* of *P. monodon* should be regarded as a new member of invertebrate *Dmc1* proteins (Fig. 3.65).

On the basis of phylogenetic analysis, different subtypes of *PGMRC* (*PGMRC1* and *PGMRC2*) should have arisen from the gene duplication process. *P. monodon PGMRC1* was clustered with *PGMRC1* of the sea urchin, *Strongylocentrotus purpuratus* but distantly related with *PGMRC1* of vertebrates and fish. Accordingly, it should be regarded as a new member of the invertebrate group (Fig. 3.66). It is also interesting to identify whether different isoforms of *P. monodon PGMRC1* are transcribed from a single locus through the alternative splicing process or encoded from different loci.

For *cyclophillin A*, a bootstraspped neighbor-joining tree revealed 3 different groups of cyclophilin (Fig 3.67). *Cyclophilin A* of *P. monodon* clustered to that of *Artemia franciscana* and allocated to be the same group as that of fish species like *Ictalurus punctatus* and *Danio rerio*. Other invertebrate *cyclophilin A* were phylogenetically allocated in different groups. Topology of the phylogenetic tree suggested that different groups of cyclophilin A should be arisen from the gene duplication process.

SUMO-1 (also called suppressor of *mif two*) plays an important role in diverse reproductive functions such as spermatogenesis and modulation of steroid receptor activity. Phylogenetic analysis indicated evolutionarily separate lineages of *SUMO-1* in vertebrates and invertebrates. Members of the former group were more closely related than those of the latter. *SUMO-1* of *P. monodon* clustered to that of *Artemia franciscana* and should be regarded as a new member of invertebrate *SUMO-1* (Fig. 3.68).

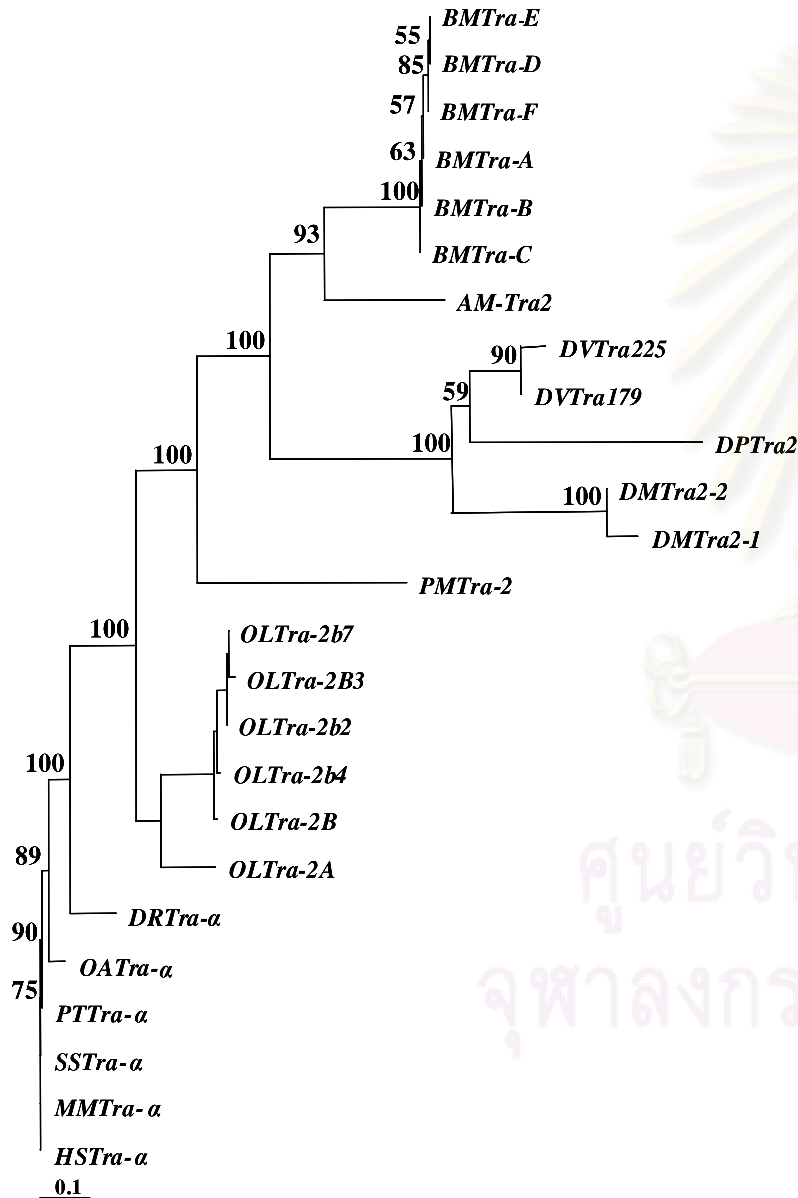


Figure 3.64 A bootstrapped neighbor-joining tree illustrating relationships between *Tra-2* of various taxa. Values at the node represent the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original data. Protein sequences of different isoforms of *Tra-2* from *Bombyx mori* (*BMTra-2A*, AAT42220.2; *BMTra-2B*, AAX47000.1; *BMTra-2C*, AAX47001.1; *BMTra-2D*, AAX47002.1; *BMTra-2E*, AAX47003.1 and *BMTra-2F*, AAX47004.1), *A. mellifera* (AMTra-2, XP 001121070), *D. melanogaster* (*DMTra-2A*, NP 476764 and *DMTra-2C*, XP 001360605), *D. virilis* (*DVTra-2-225*, AAB58112 and *DVTra-2-179*, AAB58114) and *D. pseudoobscura* (*DPTra-2*, XP001360605), *Oryzias latipes* (*OLTra-2a*, BAC06513.1; *OLTra-2b*, BAC06514.1; *OLTra-2b3*, BAD24702.1; *OLTra-2b4*, BAD24703.1 and *OLTra-2b7*, BAD24706.1;), *Pan troglodytes* (*PTTra-2α*, XP_518996.2), *Ornithorhynchus anatinus* (*OATra-2α*, XP_001513669.1), *Sus scrofa* (*SSTra-2α*, NP_001070691.1), *Mus musculus* (*MMTra-2α*, NP_932770.2), *Homo sapiens* (*HSTra-2α*, NP_037425.1), *D. rerio* (*DRTra-2α*, NP_956710.1) were retrieved from the GenBank and compared with that of *P. monodon*.

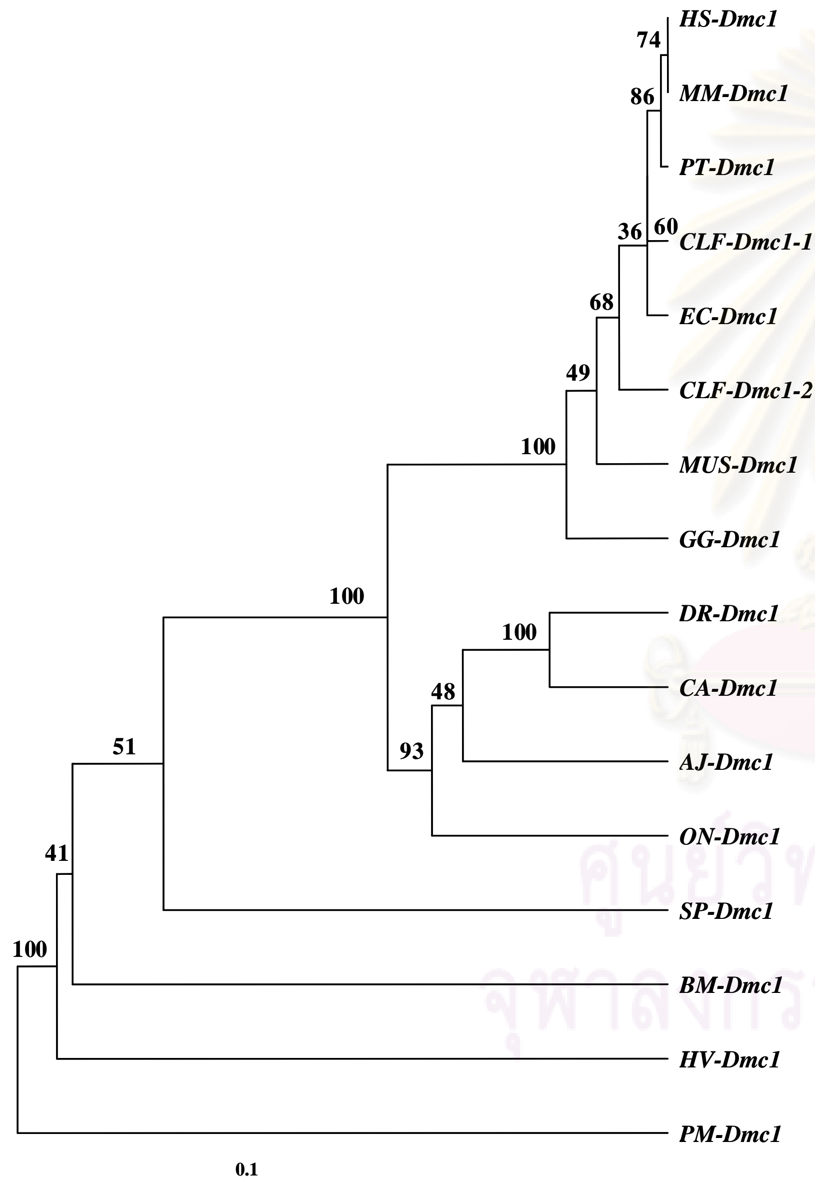


Figure 3.65 A bootstrapped neighbor-joining tree illustrating relationships between *Dmc1* of various taxa. Values at the node represent the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original aligned protein sequences. Protein sequences of different isoforms of *Dmc1* from various species; *Homo sapiens* (*HS-Dmc1*, NM_007068), *Equus caballus* (*EC-Dmc1*, XM_001501584), *Macaca mulatta* (*MM-Dmc1*, XM_001094012), *Pan troglodytes* (*PT-Dmc1*, XM_515130), *Canis lupus familiaris* (*CLF-Dmc1-1*, XM_844891 and *CLF-Dmc1-2*, XM_855217), *Gallus gallus* (*GG-Dmc1*, XM_425477), *Mus musculus* (*MM-Dmc1*, NM_010059), *Anguilla japonica* (*AJ-Dmc1*, AB182645), *Oreochromis niloticus* (*ON-Dmc1*, AB182646), *Danio rerio* (*DR-Dmc1*, NM_001020782), *Carassius auratus* (*CA-Dmc1*, EF545136), *Bombyx mori* (*BM-Dmc1*, NM_001044087), *Hydra vulgaris* (*HV-Dmc1*, AB047581) and *Strongylocentrotus purpuratus* (*SP-Dmc1*, XM_786187), were retrieved from the GenBank and compared with *Dmc1* of *P. monodon* (*PM-Dmc1*).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

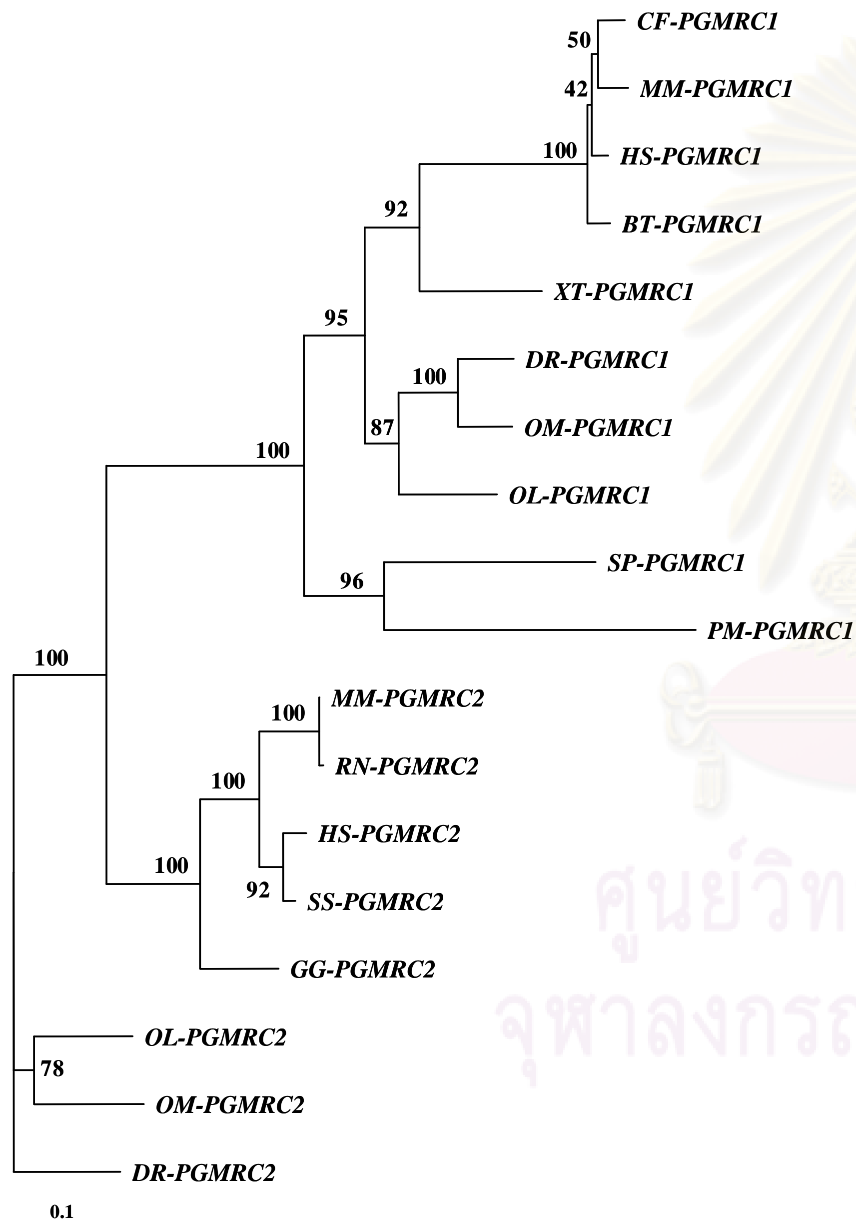


Figure 3.66 A bootstrapped neighbor-joining tree illustrating relationships between *PGMRC* of various taxa. Values at the node represent the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original aligned protein sequences. Protein sequences of different isoforms of *PGMRC1* and *PGMRC2* from various species; *Oryzias latipes* (*OL-PGMRC1*, BAE47967.1), *Danio rerio* (*DR-PGMRC1*, NP_001007393.1), *Oncorhynchus mykiss* (*OM-PGMRC1*, AAL49963.1), *Strongylocentrotus purpuratus* (*SP-PGMRC1*, XP_783332.1), *Canis familiaris* (*CF-PGMRC1*, XP_538151.1), *Homo sapiens* (*HS-PGMRC1*, NP_006658.1), *Mus musculus* (*MM-PGMRC1*, AAB97466.1), *Bos taurus* (*BT-PGMRC1*, NP_001068601.1), *Xenopus tropicalis* (*XT-PGMRC1*, NP_001006842.1) and *Rattus norvegicus* (*RN-PGMRC2*, NP_001008375.1), *Mus musculus* (*MM-PGMRC2*, AAH44759.1), *Sus scrofa* (*SS-PGMRC2*, ABX45132.1) *Homo sapiens* (*HS-PGMRC2*, NP_006311.1), *Gallus gallus* (*GG-PGMRC2*, NP_001006441.1), *Oryzias latipes* (*OL-PGMRC2*, NP_001098199.1), *Oncorhynchus mykiss* (*OM-PGMRC2*, ABD58973.1), *Danio rerio* (*DR-PGMRC2*, NP_998269.1) were retrieved from the GenBank and compared with *PGMRC1* of *P. monodon* (*PM-PGMRC1*).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

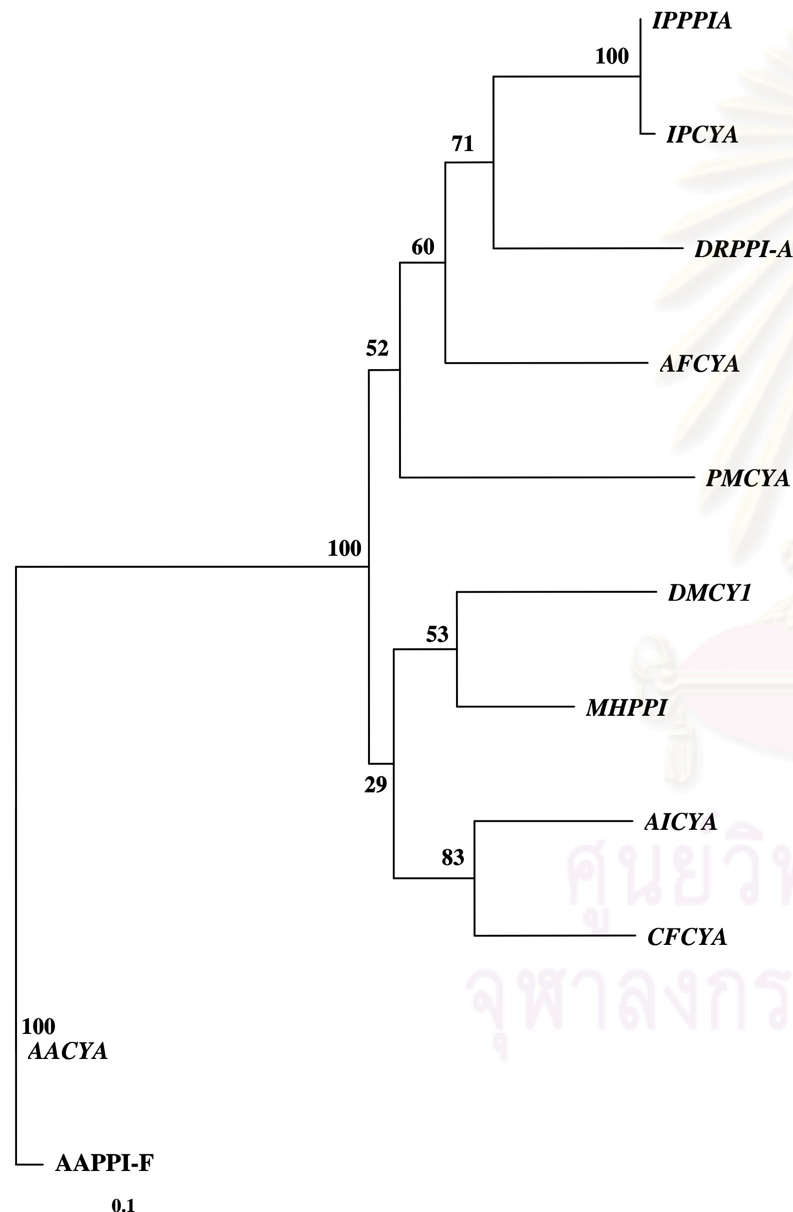


Figure 3.67 A bootstrapped neighbor-joining tree illustrating relationships between a homologue of *cyclophilin A* from *P. monodon* (*PMCYA*), and that of various taxa. Values at the node represent the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original aligned protein sequences. A scale bar indicates 10% of protein sequence divergence. Protein sequences of different isoforms of cyclophilin A (or peptidyl-prolyl isomerase, PPI) from various species; *Maconellicoccus hirsutus* (MHPPI, ABM55516.1), *Chlamys farreri* (CFCYA, AAR11779.1) *Ictalurus punctatus* (IPPPIA, AAY86951.1), *Drosophila melanogaster* (DMCY1, NP_523366.2), *Ictalurus punctatus* (IPCYA, ABO15709.1), *Aedes aegypti* (AAPPI-F, XP_001655662.1), *Aedes aegypti* (AACYA, ABF18058.1), *Danio rerio* (PPI-A, NP_956251.1), *Argopecten irradians* (AICYA, ABM92916.1), *Artemia franciscana* (AFCYA, ABN13586.1) were retrieved from the GenBank and compared with that of *P. monodon*.

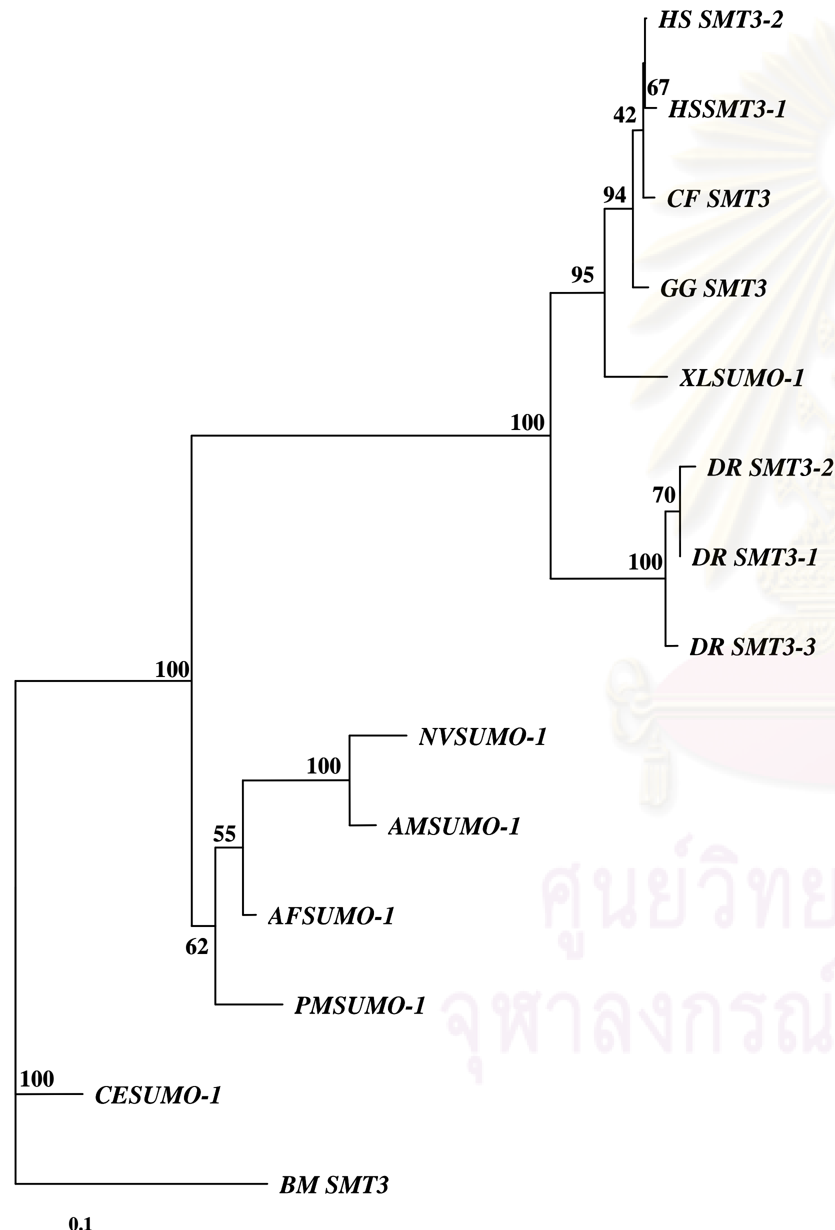


Figure 3.68 A bootstrapped neighbor-joining tree illustrating relationships between a homologue of small ubiquitin modifier-1 from *P. monodon* (PMSUMO-1), and that of various taxa. Values at the node represent the percentage of times that the particular node occurred in 1000 trees generated by bootstrapping the original aligned protein sequences. A scale bar indicates 10% of protein sequence divergence. Protein sequences of different isoforms of SUMO-1 (or SMT3) from various species; *Apis mellifera* (AMSUMO-1, XP_392826.1), *Artemia franciscana* (AFSUMO-1, ABQ41279.1), *Brugia malayi* (BMSMT3, EDP30718.1), *Nasonia vitripennis* (NVSUMO-1, XP_001607301.1), *Caenorhabditis elegans* (CESUMO-1, NP_490842.1), *Xenopus laevis* (XLSUMO-1, NP_001083717.1), *Danio rerio* (DRSMT3-1, NP_998324.1), *Homo sapiens* (HSSMT3, AAH66306.1), *Canis familiaris* (CESMT3, XP_536034.1), *Homo sapiens* (HSSMT3-A, NP_003343.1), *Danio rerio* (DRSMT3, AAI52218.1), *Gallus gallus* (GGSMT3, NP_989466.1), *Danio rerio* (DRSMT3-2, XP_001343161.1) were retrieved from the GenBank and compared with that of *P. monodon*.

3.5 Examination of expression patterns and tissue distribution analysis of genes related to testicular development by RT-PCR

Expression patterns of 59 gene homologues were non-quantitatively examined using the cDNA template from gonads of juvenile and broodstock of male and female *P. monodon* (N = 4 for each group) using RT-PCR analysis.

An EST that marginally matched *low molecular weight neurofilament protein XNF-L* (E-value = 8e-04, called *PMTST1* after further characterization by RACE-PCR) was only expressed in testes (N = 8) but not ovaries (N = 8) of *P. monodon* when analyzed by RT-PCR for 25, 30, and 35 cycles (Fig. 3.69 and Table 3.11).

Four ESTs representing *multiple inositol polyphosphate phosphatase* (*MIPP*) homologues were found in the conventional testis cDNA library, and can be divided to isoform 1 (TT-N-S01-0004-W and TT-N-S01-0459-W) and isoform 2 (TT-N-S01-0678-W and TT-N-S01-0730-W) owing similarity analysis of the original EST clones. Primers were designed from a representative of each isoform (TT-N-S01-0004-W and TT-N-S01-0678-W). RT-PCR was carried out and revealed that both *MIPP* showed a trend of more abundant expression in testes than ovaries of *P. monodon* (Fig. 3.70 and Table 3.11). Likewise, *HSP70-2* and *Dmc1* also exhibited more preferential expression in testes than ovaries of *P. monodon* broodstock (Fig. 3.71 and Table 3.11).

Table 3.10 Summary of expression patterns of various gene homologues analyzed by RT-PCR

Category	Number of genes
Testis-specific expression	1
Trands of preferential expression	40
-Testis	4
-Ovary	36
Comparable expression levels in testes and ovaries	14
No products or nonspecific amplification products	4
Total	**59**

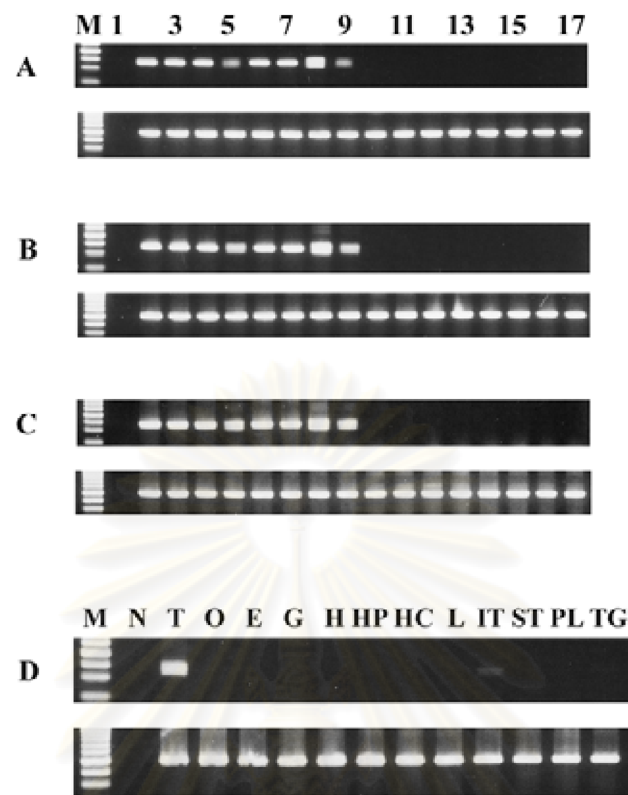


Figure 3.69 RT-PCR of *PMTST1* (374 bp) using the first strand cDNA of testes of broodstock (lanes 2-5) and juveniles (lanes 6-9) and ovaries of broodstock (lanes 10-13) and juveniles (lanes 14-17) of *P. monodon* for 25 (A), 30 (B) and 35 (C) cycles. *EF 1-α* (500 bp) was used as the positive control. Lanes 1 were the negative control (without the cDNA template).

Tissue distribution analysis (D) was carried out using the template from testes (T), ovaries (O), eyestalks (E), gills (G), heart (H), hepatopancreas (HP), hemocytes (HC), lymphoid organs (LO), intestine (IT), stomach (ST), pleopods (PL), and thoracic ganglion (TG).

In contrast, thirty-six genes exhibited a trend of more abundantly expression patterns in ovaries than testes (Table 3.12). These included *SUMO-1, cyclophilin A* (*CYA*), *transformer-2 protein B* (*Tra-2*), *B-cell receptor-associated protein 37* (*prohibitin-2*), *thyroid hormone receptor-associated protein complex 240 kDa component (Trap240), growth factor receptor-bound protein, and dynein light intermediate chain* (*D2LIC*) (Fig. 3.72).

Table 3.11 Gene homologues that specifically expressed in testes but not ovaries and those showing a trend of preferentially expressed in testes to ovaries of *P. monodon*

Gene Homologues	Expected size (bp)	Tissue distribution
1. *Low molecular weight neurofilament protein (PMTST1)*	374	T and IT
2. *Multiple inositol polyphosphate phosphatase ; MIPP*	174	T and O
3. *Multiple inositol polyphosphate phosphatase 2; MIPP2*	396	T and O
4. *Meiotic recombination protein DMC1/LIM15 homolog isoform 1 (Dmc1)*	425	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
5. *Testis-specific heat shock-related protein 22 (HST70-2)*	321	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG

In addition, *carbonyl reductase 1*, *cyclin dependent kinase 2* (*Cdk2*), *ubiquitin carboxyl-terminal hydrolase 14, ubiquitin carboxyl-terminal hydrolase* 5, *cell division cycle 2* (*cdc2*), *E1B-55kDa-associated protein 5 isoform 5*, and *oncoprotein nm23* also showed a treand of preferential expressed in ovaries to testes of *P. monodon* (Fig. 3.73).

Fourteen gene homologues did not revealed differential expression levels in testes and ovaries of *P. monodon* (Table 3.13). Although *PGMRC1* and *Inx2* did not revealed differential expression between ovaries and testes of *P. monodon* when their expression patterns in juvenile and broodstock of each sex were considered together, *PGRMC1* showed a trend of more abundantly expressed in testes of juveniles than broodstock of *P. monodon* (Fig. 3.74). In contrast, *Inx2* exhibited a trend of preferentially expressed in ovaries of juveniles to broodstock of *P. monodon* (Fig. 3.74).

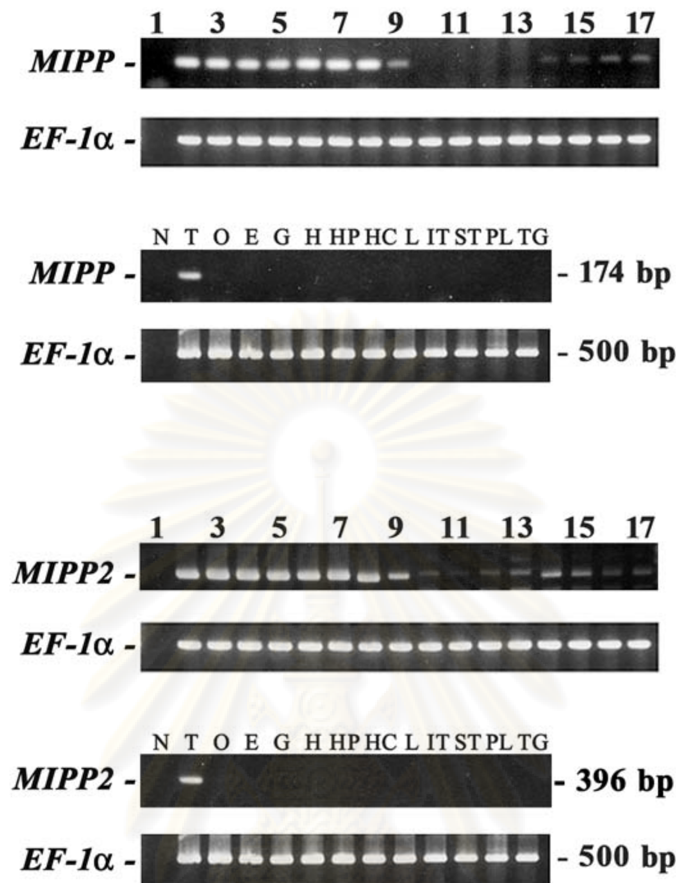


Figure 3.70 RT-PCR of *MIPP* (A) and *MIPP2* (B), using the first strand cDNA of testes of broodstock (lanes 2–5) and juveniles (lanes 6–9) and ovaries of broodstock (lanes 10–13) and juveniles (lanes 14–17) of *P. monodon* for 25 cycles. *EF 1-α* (500 bp) was used as the positive control. Lanes 1 were the negative control (without the cDNA template).

Tissue distribution analysis (D) was carried out using the template from testes (T), ovaries (O), eyestalks (E), gills (G), heart (H), hepatopancreas (HP), hemocytes (HC), lymphoid organs (LO), intestine (IT), stomach (ST), pleopods (PL), and thoracic ganglion (TG).

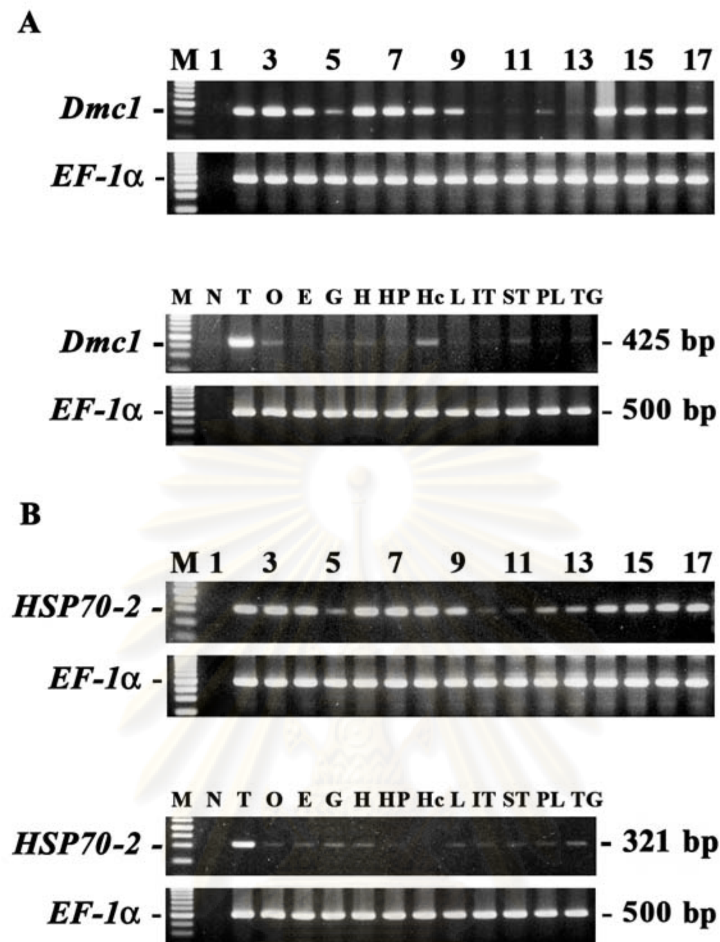


Figure 3.71 RT-PCR of *Dmc1* (A) and *HSP70-2* (B) using the first strand cDNA of testes of broodstock (lanes 2-5) and juveniles (lanes 6-9) and ovaries of broodstock (lanes 10-13) and juveniles (lanes 14-17) of *P. monodon* for 25 cycles. *EF 1-α* (500 bp) was used as the positive control. Lanes 1 were negative control (without the cDNA template).

Tissue distribution analysis (D) was carried out using the template from testes (T), ovaries (O), eyestalks (E), gills (G), heart (H), hepatopancreas (HP), hemocytes (HC), lymphoid organs (LO), intestine (IT), stomach (ST), pleopods (PL), and thoracic ganglion (TG).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.12 Gene homologues showing a trend of more abundant expression patterns in ovaries than testes of *P. monodon*

Gene Homologues	size (bp)	Tissue distribution
1. *B-cell receptor-associated protein 37* (*Prohibitin 2*)	563	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
2. *26S Proteasome non-ATPase regulatory subunit 3 (Diphenol oxidase A2 component)*	140	Not determined
3. *Actin-binding protein anillin*	199	Not determined
4. *Acyl-CoA oxidase (ACX3)*	250	Not determined
5. *Arginyl-tRNA--protein transferase 1*	292	Not determined
6. *Budding uninhibited by benzimidazoles 3 homolog (Mitotic checkpoint)*	257	Not determined
7. *Carbonyl reductase 1*	457	Not determined
8. *Cell division cycle 2 (cdc2)*	510	Not determined
9. *COP9 constitutive photomorphogenic homolog subunit 5 isoform 1*	190	Not determined
10. *Cyclin dependent kinase 2 (Cdk2)*	348	Not determined
11. *Cyclophilin A*	495	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
12. *Dynactin 5*	300	Not determined
13. *Dynein light intermediate chain*	324	T, O, G, H, HC, LO, IT, PL, and TG
14. *E1B-55kDa-associated protein 5 isoform 5*	196	Not determined
15. *Eukaryotic translation initiation factor 4 gamma, 2*	337	Not determined
16. *Growth factor receptor-bound protein*	299	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
17. *Histone H2AV (H2A.F/Z)*	329	Not determined
18. *Karyopherin (importin) alpha 2*	295	Not determined
19. *Laminin beta chain*	363	Not determined
20. *Metaxin 2*	212	Not determined
21. *Multiprotein bridging factor 1*	235	Not determined
22. *Nucleoside diphosphate linked moiety X-type motif 9*	341	Not determined
23. *Oncoprotein nm23*	229	Not determined
24. *PCTAIRE protein kinase 2 (PCTK2)*	250	Not determined
25. *Profilin*	259	Not determined
26. *Proteasome alpha 3 subunit*	250	Not determined
27. *Rac GTPase activating protein 1 isoform 1*	248	Not determined
28. *Serine/threonine-protein kinase 23 (Muscle-specific serine kinase 1, MSSK-1)*	272	Not determined
29. *Small ubiquitin-like modifier(SUMO-1)*	362	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
30. *Spermidine synthase*	223	Not determined
31. *Thyroid hormone receptor-associated protein complex 240 kDa component (Trap240)*	335	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
32. *Transformer-2 protein B*	229	T, O, G, H, HC, LO, IT, and ST
33. *Ubiquitin carboxyl-terminal hydrolase 14*	240	Not determined
34. *Ubiquitin carboxyl-terminal hydrolase 5*	528	Not determined
35. *Ubiquitin-conjugating enzyme E2*	232	Not determined
36. *USO-1*	314	Not determined

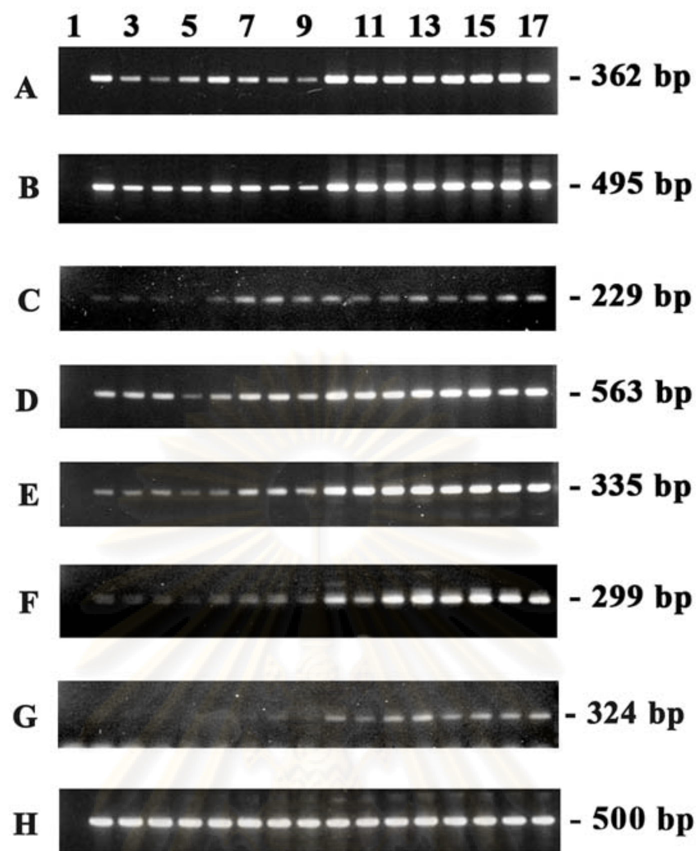


Figure 3.72 RT-PCR of *SUMO-1* (A), *CYA* (B), *Tra-2* (C), *prohibitin-2* (D), *Trap240* (E), *growth factor receptor-bound protein* (F), and *D2LIC* (G) using the first strand cDNA of testes of broodstock (lanes 2-5) and juveniles (lanes 6-9) and ovaries of broodstock (lanes 10-13) and juveniles (lanes 14-17) of *P. monodon* for 25 cycles. *EF 1-α* (H) was used as the positive control. Lane 1 was the negative control (without the cDNA template).

Tissue distribution analysis of 15 gene homologues of *P. monodon* was carried out. *PMTST1* was abundantly expressed in testis and expressed at a low level in intestine but was not expressed in other tissues of *P. monodon* broodstock (Fig. 3.69). This suggested that *PMTST1* may play an important role for testicular development of *P. monodon*.

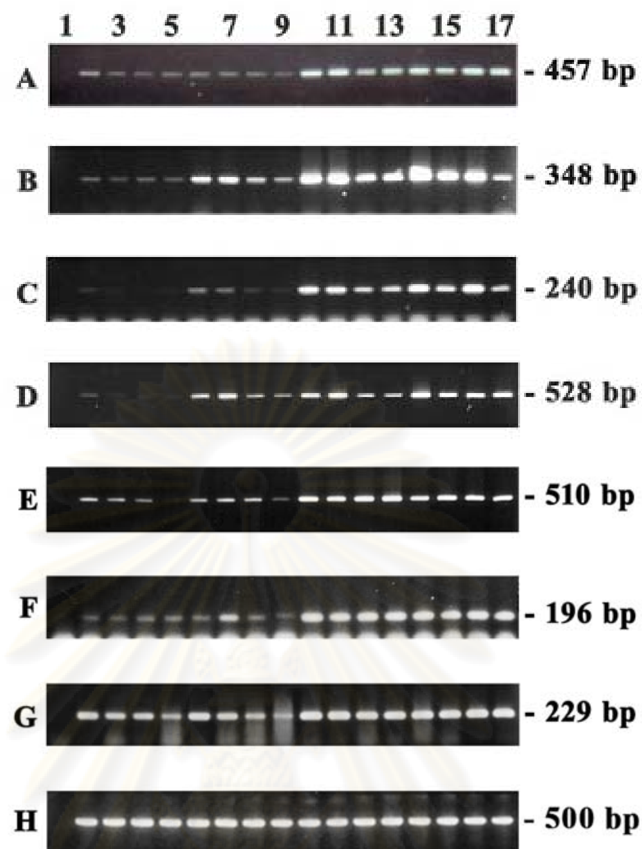


Figure 3.73 RT-PCR of *carbonyl reductase 1* (A), *Cdk2* (B), *ubiquitin carboxyl-terminal hydrolase 14* (C), *ubiquitin carboxyl-terminal hydrolase* 5 (D), *cdc2* (E), *E1B-55kDa-associated protein isoform 5* (F), and *oncoprotein nm23* (G) using the first strand cDNA of testes of broodstock (lanes 2-5) and juveniles (lanes 6-9) and ovaries of broodstock (lanes 10-13) and juveniles (lanes 14-17) of *P. monodon* for 25 cycles. *EF 1-α* (H) was used as the positive control. Lanes 1 were the negative control (without the cDNA template).

Tissue distribution analysis revealed that both *MIPP* (TT-N-S01-0004-W) and *MIPP2* (TT-N-S01-0678-W) were more abundantly expressed in testis than ovaries and these transcripts were not expressed in other tissues of *P. monodon* broodstock (Fig. 3.70). In contrast, *HSP70-2* and *Dmc1* were constitutively expressed in all tissues of *P. monodon* broodstock. *HSP70-2* and *Dmc1* was more abundantly expressed in testis than other tissues. Nevertheless, the former was low abundantly

expressed in hemocyte whereas low expression of the latter was observed in eyestalk and lymphoid organs (Fig. 3.71).

On the other hand, *prohibitin-2*, *Trap240*, *innexin 2*, *PGRMC1*, *SUMO-1*, *CYA*, *growth factor receptor-bound protein*, and *saposin* were expressed in all examined tissues whereas *Tra-2* was expressed in all tissues except eyestalks, hepatopancreas, pleopods and thoracic ganglion. *D2LIC* was not expressed in eyestalks, hepatopancreas, and stomach (Fig. 3.74-3.75).

Table 3.13 Gene homologues that did not exhibit differential expression in testes and ovaries of *P. monodon*

Gene Homologues	Expected size (bp)	Tissue distribution
1. *Cell division protein kinase 7 (Cdk7)*	381	Not determined
2. *Cystathionine-γ-lyase*	145	Not determined
3. *Degradation in endoplasmic reticulum protein 1 (DER1)*	206	Not determined
4. *Flotillin 2*	854	Not determined
5. *Importin 7*	297	Not determined
6. *Innexin 1*	422	Not determined
7. *Innexin2*	391	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
8. *Polyadenylate binding protein II*	187	Not determined
9. *Progestin receptor membrane component 1 (PGRMC1)*	122	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
10. *Programmed cell death protein 7*	181	Not determined
11. *Prosaposin isoform 3*	614	T, O, E, G, H, HP, HC, LO, IT, ST, PL, and TG
12. *Proteasome (prosome, macropain) 26S subunit, ATPase, 2, isoform CRA_a*	187	Not determined
13. *Synaptobrevin-like protein 1*	291	Not determined
14. *Thyroid hormone receptor coactivating protein 120kDa (TrCP120)*	392	Not determined

Table 3.14 Gene homologues that exhibit non-specific amplification from RT-PCR analysis

Gene Homologues	Expected size (bp)	Remarks
1. *Adaptor-related protein complex 1, beta 1 subunit*	290	Non-specific product
2. *BCS-2*	372	Non-specific product
3. *Inhibitor of apoptosis protein (IAP)*	238	Non-specific product
4. *Serine palmitoyl transferase LCB2 subunit*	178	Non-specific product

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

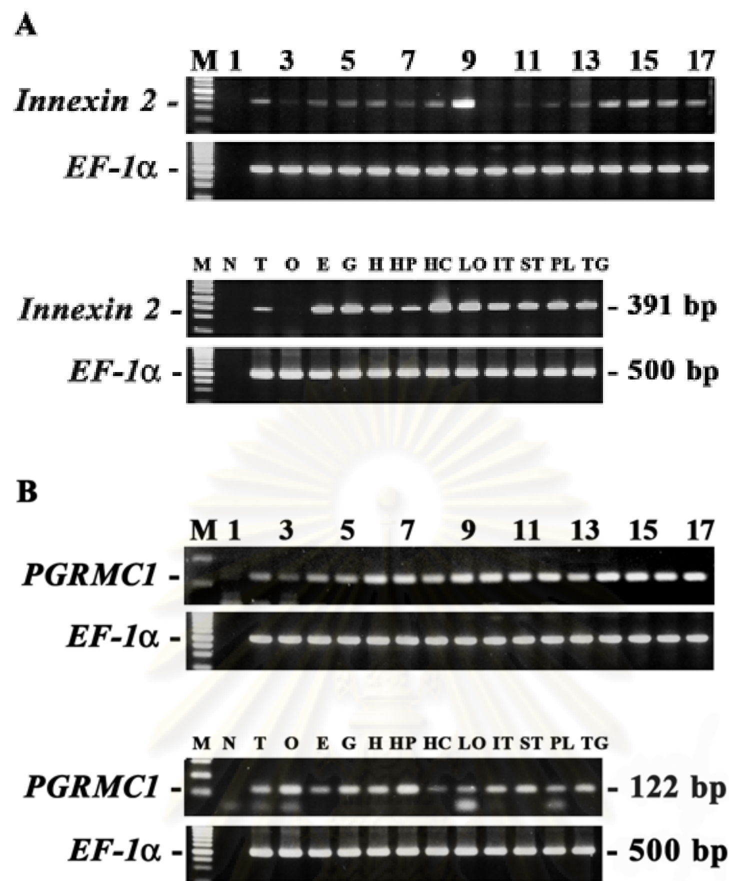


Figure 3.74 RT-PCR of *innexin 2* (A), *PGRMC1* (B) using the first strand cDNA of testes of broodstock (lanes 2-5) and juveniles (lanes 6-9) and ovaries of broodstock (lanes 10-13) and juveniles (lanes 14-17) of *P. monodon* for 25 cycles. *EF 1-α* (H) was used as the positive control. Lanes 1 were the negative control (without the cDNA template.

Tissue distribution analysis (D) was carried out using the template from testes (T), ovaries (O), eyestalks (E), gills (G), heart (H), hepatopancreas (HP), hemocytes (HC), lymphoid organs (LO), intestine (IT), stomach (ST), pleopods (PL), and thoracic ganglion (TG).

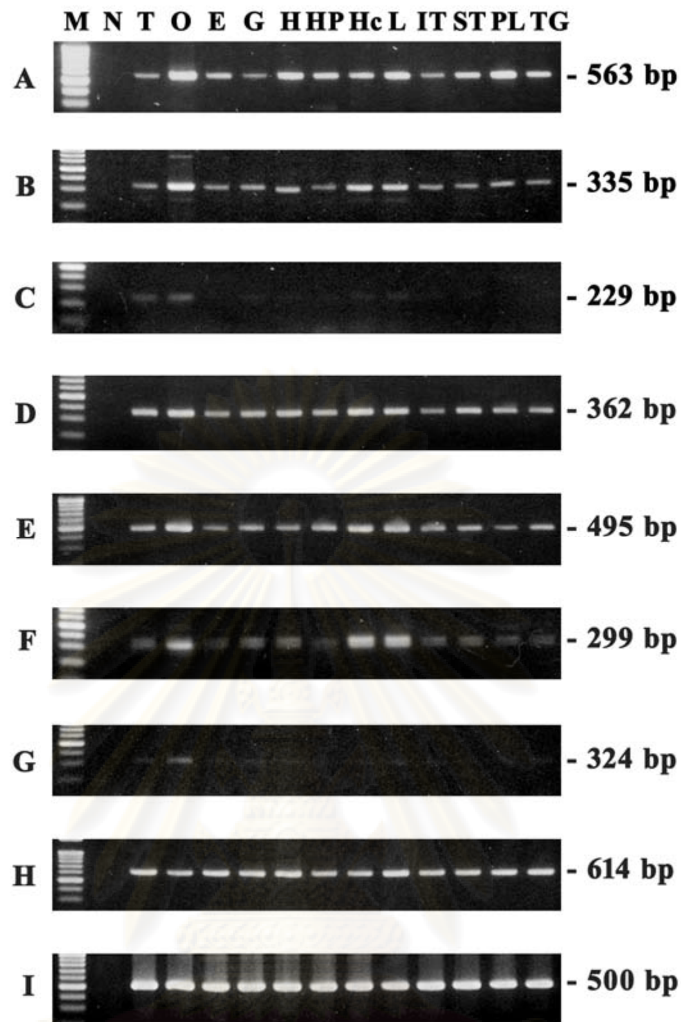


Figure 3.75 Tissue distribution analysis of *prohibitin-2* (A), *Trap240* (B), *Tra-2* (C), *SUMO-1* (D), *CYA* (E), *growth factor receptor-bound protein* (F), *D2LIC* (G) and *saposin* (H) using the first strand cDNA of testes (T), ovaries (O), eyestalks (E), gills (G), heart (H), hepatopancreas (HP), hemocytes (HC), lymphoid organs (LO), intestine (IT), stomach (ST), pleopods (PL), and thoracic ganglion (TG) of *P. monodon* broodstock for 25 cycles. *EF 1-α* was used as the positive control (I). Lane N was the negative control (without the cDNA template).

Non-specific amplification products were obtained from amplification of *adaptor-related protein complex 1*, *beta 1 subunit*, *BCS-2*, *inhibitor of apoptosis protein* (*IAP*), and *serine palmitoyl transferase LCB2 subunit*; Table 3.14).

3.6 Semiquantitative RT-PCR of functionally important gene homologues in testes and ovaries of broodstock and juvenile *P. monodon*

Expression levels of 6 gene homologues, *low molecular weight neurofilament protein XNF-L (PMTST1)*, *multiple inositol polyphosphate phosphatase* (*MIPP*), *transformer-2* (*Tra-2*), *prohibitin-2*, *thyroid hormone receptor-associated protein complex 240 kDa component (Trap240),* and *heat shock-related protein 2* (*HSP70-2*), were determined by semiquantitative RT-PCR. *EF-1α* was used as an internal control.

The first strand cDNA of testes or ovaries from juvenile males and females (4 months old with a body weight of approximately 25-30 g), broodstock-sized domesticated shrimp (F1 generation, an average body weight, ABW = 88.66 ± 9.02 g; the gonadosomatic index, GSI = 0.34 ± 0.15%, N = 9), wild male (ABW = 81.69 ± 15.63 g, GSI = 0.84 ± 0.32%, N = 8), wild female (ABW = 142.98 ± 28.37 g, GSI = 2.98 ± 2.02%, N = 4) broodstock and wild shrimp at 1-5 (ABW = 104.79 ± 12.40 g, GSI = 0.40 ± 0.13%, N = 6), 6-9 (ABW = 99.19 ± 14.94 g, GSI = 0.41 ± 0.29%, N = 4) and 10-16 days (ABW = 102.73 ± 19.23 g, GSI = 0.48 ± 0.16%, N = 8) after molting were used as the template for semiquantitative RT-PCR analysis. This technique requires optimization of several parameters including concentration of primers, $MgCl_2$, and the number of PCR cycles.

3.6.1 Optimization of the primer concentration, $MgCl_2$ concentration, and cycle numbers

RT-PCR of each gene was carried out with fixed components except primer concentration. Lower concentrations may result in non-quantitative amplification whereas higher concentrations of primer may leave a large amount of unused primers which could give rise to non-specific amplification products. The suitable concentration of primers for each gene is shown by Table 3.15.

The optimal concentration of $MgCl_2$ for each primer pair was carefully examined using the amplification conditions with the optimized primer concentration. The concentration of $MgCl_2$ that gave the highest yields and specificity for each PCR product was chosen (Table 3.15)

The number of amplification cycles was important because the product reflecting the expression level should be measured quantitatively before reaching a plateau amplification phase. At the plateau stage, transcripts initially present at different levels may give equal intensity of the amplification products.

In this experiment, the cycler numbers of RT-PCR of each gene was performed using the conditions that primer and $MgCl_2$ concentrations were optimized. The number of cycles that gave the highest yield before the product reached a plateau phase of amplification was chosen (Table 3.15).

Table 3.15 Optimal primer and $MgCl_2$ concentrations and the number of PCR cycles for semiquantitative RT-PCR analysis of gene homologues in *P. monodon*

Transcripts	**Expected size (bp)**	**Primer concentration**	**MgCl2 concentration**	**PCR cycler**
Low molecular weight neurofilament protein XNF-L (PMTT1)	347	0.075	1.5	26
Transformer-2 (*Tra*-2)	499	0.15	2.0	33
Prohibitin-2	563	0.075	1.0	29
Thyroid hormone receptor-associated protein complex 240 kDa component (Trap240)	335	0.075	1.0	30
Multiple inositol polyphosphate phosphatase (*MIPP*)	174	0.15	2.0	26
Heat shock-related protein 2 (*HSP70-2*)	321	0.25	1.0	26
EF-1α	500	0.15	1.5	21

3.6.2 Semiquantitative RT-PCR analysis

Result of semiquantitative RT-PCR further confirmed that *PMTST1* was only expressed in testes but not ovaries of *P. monodon*. The expression level of *PMTST1* in testes *P. monodon* was not different in all three groups including CJ-TT (0.8644 ± 0.0612), DB-TT (0.7729 ± 0.0784), and WB-TT (0.8240 ± 0.0559) ($P > 0.05$) whereas the expression level of this gene in males at 10-16 DAM (1.1262 ± 0.0448) were greater than that at 1-5 DAM (0.9377 ± 0.1128) and 6-9 DAM (0.9586 ± 0.0589) ($P < 0.05$). Nevertheless, the expression levels of *PMTST1* in males at 1-5 DAM and 6-9 DAM were not different ($P > 0.05$) (Fig.3.76-3.77). In addition, the expression levels of this gene at 6-9 DAM and 10-16 DAM were higher than that of WB-TT ($P < 0.05$).

Expression levels of *MIPP* in ovaries of *P. monodon* (CJ-OV and WB-OV) were extremely low. In the testis groups, the expression level of *MIPP* in CJ-TT (0.9972 ± 0.0532) was higher than in that in DB-TT and WB-TT (0.8575 ± 0.07323 and 0.8876 ± 0.0740, respectively; $P < 0.05$). In addition, the expression level of this gene in males at 10-16 DAM (0.7702 ± 0.0792) was also higher than that at 6-9 DAM (0.5838 ± 0.0273, $P < 0.05$). However, the expression levels of *MIPP* at 1-5 DAM (0.6560 ± 0.0923) and 6-9 DAM were not different ($P>0.05$) (Fig.3.78-3.79). In addition, *MIPP* was more abundantly expressed in WB-TT than 1-5 DAM (0.6560 ± 0.0923) and 6-9 DAM ($P < 0.05$).

The expression levels of *Tra-2* in testes and ovaries *P. monodon* were not different in all examined groups; CJ-TT (0.0687 ± 0.0046), DB-TT (0.0589 ± 0.0118), WB-TT (0.0647 ± 0.0056), CJ-OV (0.0663 ± 0.0061), and WB-OV (0.0625 ± 0.0067) ($P > 0.05$). However, the expression level of this gene in males at 10-16 DAM (0.0585 ± 0.0053) was higher than that of 1-5 DAM (0.0472 ± 0.0053) and 6-9 DAM (0.0468±0.0033) ($P < 0.05$) but the expression levels of *Tra-2* at 1-5 DAM and 6-9 DAM were not different ($P > 0.05$) (Fig.3.80-3.81). In addition, the expression level of this gene in WB-TT was greater than that of 1-5 DAM and 6-9 DAM ($P < 0.05$).

The expression level of *prohibitin2* in WB-TT (1.0563 ± 0.1105) was less than that in CJ-TT, DB-TT, CJ-OV, and WB-OV (1.3046 ± 0.0711, 1.2368 ± 0.0719, 1.2868 ± 0.0516 and 1.2879 ± 0.0506, respectively; $P < 0.05$). In contrast, expression

levels of this gene in males at 1-5 DAM (0.9185 ± 0.1406), 6-9 DAM (0.8609 ± 0.0937) and 10-16 DAM (0.9939 ± 0.0834) were not different ($P > 0.05$) (Fig.3.82-3.83). In addition, the expression level of *prohibitin2* in WB-TT was greater than that of 6-9 DAM ($P<0.05$).

The expression levels of *Trap240* in CJ-TT (0.0643 ± 0.0030), CJ-OV (0.0705 ± 0.0022) and WB-OV (0.0678 ± 0.0054) were more abundant than those in WB-TT and DB-TT (0.0557 ± 0.0028 and 0.0550 ± 0.0072, respectively; $P > 0.05$). Therefore, *Trap240* was preferentially expressed in ovaries of *P. monodon* ($P < 0.05$). In addition, the expression level of this gene in males at 10-16 DAM (0.0596 ± 0.0057) was more abundant than at 1-5 DAM (0.0503 ± 0.0060) ($P<0.05$). Expressions of *Trap240* at 6-9 DAM and 10-16 DAM were not different ($P > 0.05$) (Fig.3.84-3.85). The expression levels of this gene in WB-TT and molting groups were not different ($P > 0.05$).

The expression levels of *HSP70-2* in testes were more abundant than those in ovaries of *P. monodon* ($P < 0.05$). Expression levels of *HSP70-2* in the testes groups were not different in all examined groups; CJ-TT (0.9705 ± 0.0737), DB-TT 0.8143 ± 0.1582) and WB-TT (0.8216 ± 0.0951) ($P > 0.05$). Likewise, expression of *HSP70-2* in CJ-OV (0.6996 ± 0.0718) and WB-OV (0.5891±0.0356) was not different ($P > 0.05$). The expression level of this gene in males at 10-16 DAM (0.8779 ± 0.1260) was higher than that at 6-9 DAM (0.6528 ± 0.0595) ($P < 0.05$) (Fig.3.86-3.87). In addition, the expression levels of this gene in WB-TT and molting groups were not different ($P > 0.05$).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

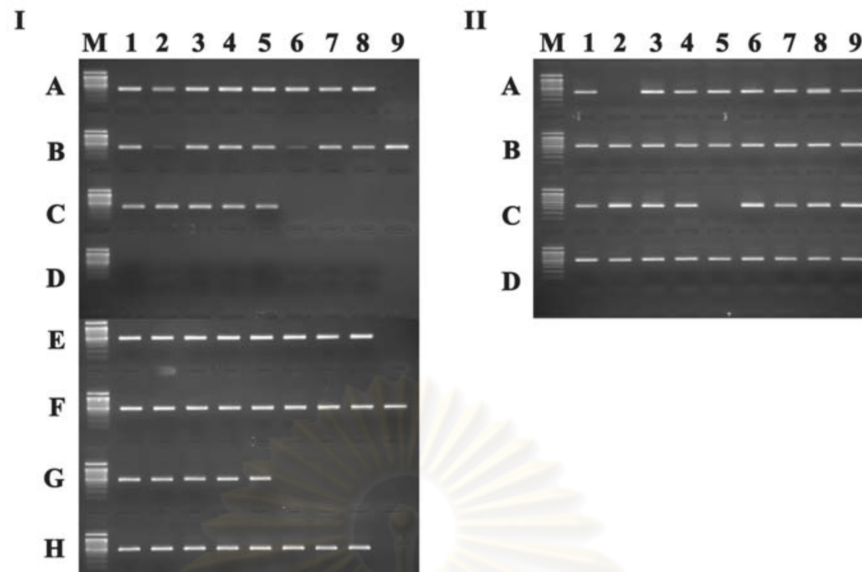


Figure 3.76 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *PMTST1* (panel I: A-D and panel II: A and C) and *EF-1α* (panel I: E-H and panel II: B and D) of testis from wild broodstock (WB-TT, I-A and I-E, lanes 1-8), domesticated broodstock (DB-TT, I-B and I-F, lanes 1-9), cultured juveniles (CJ-TT, I-C and I-G, lanes 1-5), and ovaries from wild broodstock (WB-OV, I-D and I-H; lanes 1-4), and cultured juveniles (CJ-OV, I-D and I-H; lanes 5-8) and testes from broodstock after molting for Day 1-5 (II-A and II-C; lanes 1-6), Day 6-9 (II-A and II-C; lanes 1-6 and II-B and II-D; lane 1), and Day 10-16 (II-B and II-D; lanes 2-9). Lanes M = 100 bp DNA ladder.

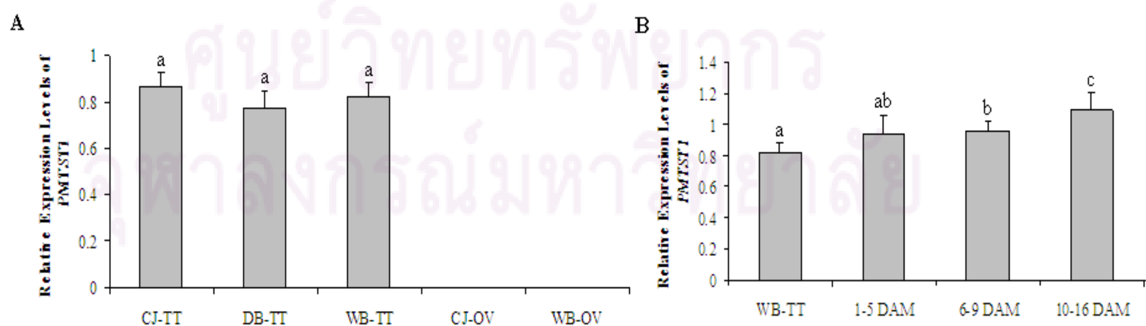


Figure 3.77 Histograms showing the expression levels of *PMTST1* in gonads of *P. monodon*. The same letters indicate that the expression levels were not significantly different (P>0.05).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

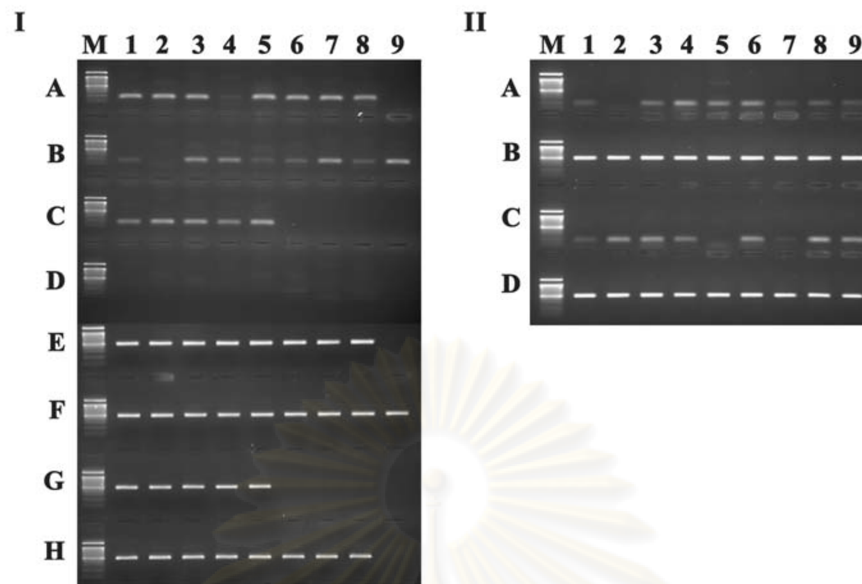


Figure 3.78 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *MIPP* (panel I: A-D and panel II: A and C) and *EF-1α* (panel I: E-H and panel II: B and D) of testis from wild broodstock (WB-TT, I-A and I-E, lanes 1-8), domesticated broodstock (DB-TT, I-B and I-F, lanes 1-9), cultured juveniles (CJ-TT, I-C and I-G, lanes 1-5), and ovaries from wild broodstock (WB-OV, I-D and I-H; lanes 1-4), and cultured juveniles (CJ-OV, I-D and I-H; lanes 5-8) and testes from broodstock after molting for Day 1-5 (II-A and II-C; lanes 1-6), Day 6-9 (II-A and II-C; lanes 1-6 and II-B and II-D; lane 1), and Day 10-16 (II-B and II-D; lanes 2-9). Lanes M = 100 bp DNA ladder.

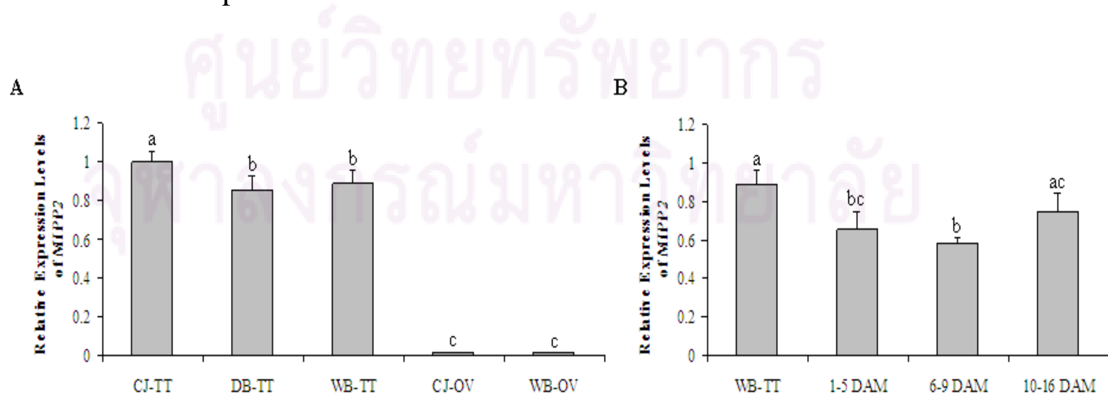


Figure 3.79 Histograms showing the expression levels of *MIPP* in gonads of *P. monodon*. The same letters indicate that the expression levels were not significantly different (*P*>*0.05*)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

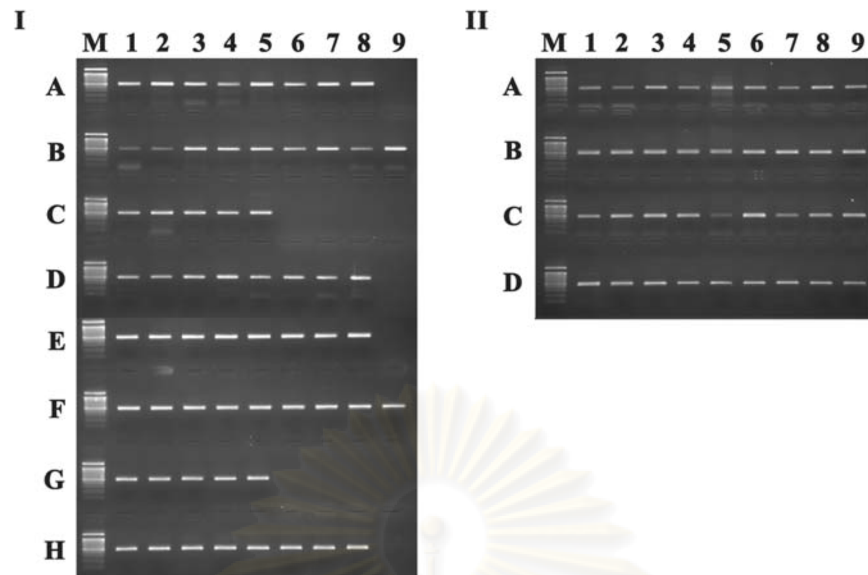


Figure 3.80 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *Tra-2* (panel I: A-D and panel II: A and C) and *EF-1α* (panel I: E-H and panel II: B and D) of testis from wild broodstock (WB-TT, I-A and I-E, lanes 1-8), domesticated broodstock (DB-TT, I-B and I-F, lanes 1-9), cultured juveniles (CJ-TT, I-C and I-G, lanes 1-5), and ovaries from wild broodstock (WB-OV, I-D and I-H; lanes 1-4), and cultured juveniles (CJ-OV, I-D and I-H; lanes 5-8) and testes from broodstock after molting for Day 1-5 (II-A and II-C; lanes 1-6), Day 6-9 (II-A and II-C; lanes 1-6 and II-B and II-D; lane 1), and Day 10-16 (II-B and II-D; lanes 2-9). Lanes M = 100 bp DNA ladder.

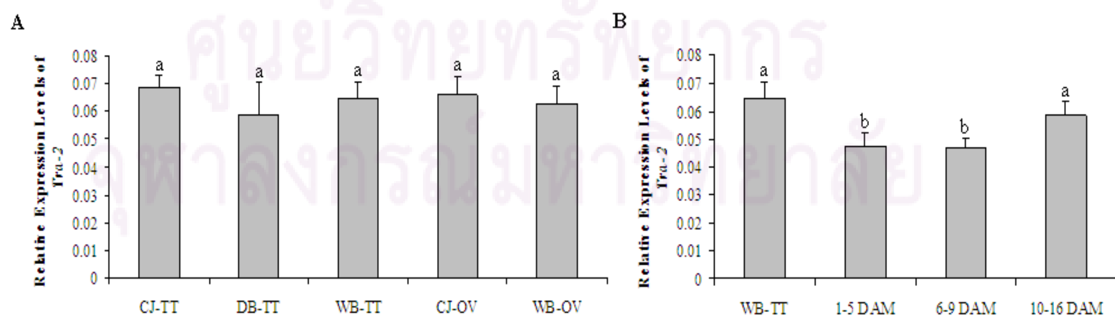


Figure 3.81 Histograms showing the expression levels of *Tra-2* in gonads of *P. monodon*. The same letters indicate that the expression levels were not significantly different (*P>0.05*).

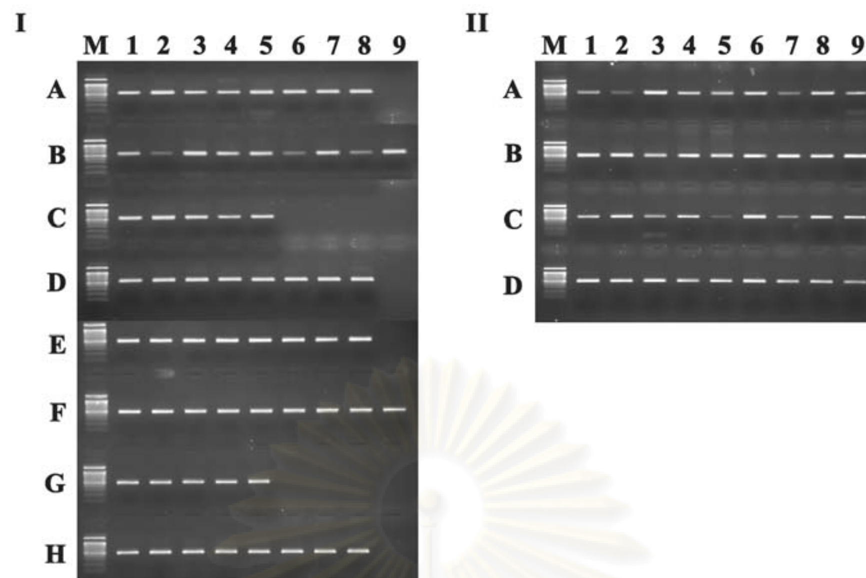


Figure 3.82 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *prohibitin 2* (panel I: A-D and panel II: A and C) and *EF-1α* (panel I: E-H and panel II: B and D) of testis from wild broodstock (WB-TT, I-A and I-E, lanes 1-8), domesticated broodstock (DB-TT, I-B and I-F, lanes 1-9), cultured juveniles (CJ-TT, I-C and I-G, lanes 1-5), and ovaries from wild broodstock (WB-OV, I-D and I-H; lanes 1-4), and cultured juveniles (CJ-OV, I-D and I-H; lanes 5-8) and testes from broodstock after molting for Day 1-5 (II-A and II-C; lanes 1-6), Day 6-9 (II-A and II-C; lanes 1-6 and II-B and II-D; lane 1), and Day 10-16 (II-B and II-D; lanes 2-9). Lanes M = 100 bp DNA ladder.

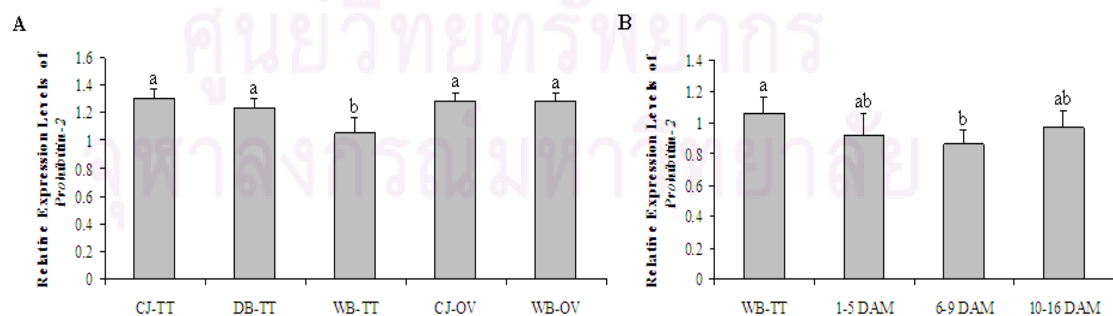


Figure 3.83 Histograms showing the expression levels of *prohibitin 2* in gonads of *P. monodon*. The same letters indicate that the expression levels were not significantly different ($P > 0.05$).

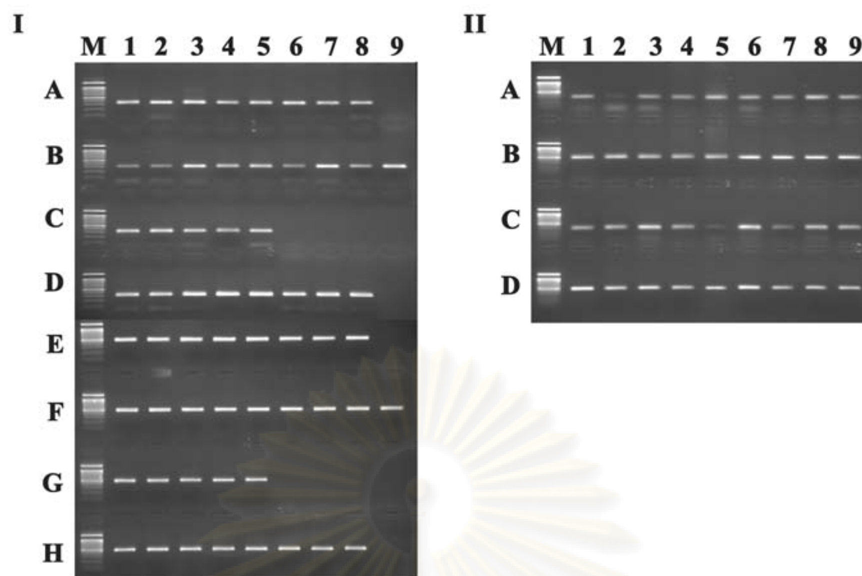


Figure 3.84 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *Trap240* (panel I: A-D and panel II: A and C) and *EF-1α* (panel I: E-H and panel II: B and D) of testis from wild broodstock (WB-TT, I-A and I-E, lanes 1-8), domesticated broodstock (DB-TT, I-B and I-F, lanes 1-9), cultured juveniles (CJ-TT, I-C and I-G, lanes 1-5), and ovaries from wild broodstock (WB-OV, I-D and I-H; lanes 1-4), and cultured juveniles (CJ-OV, I-D and I-H; lanes 5-8) and testes from broodstock after molting for Day 1-5 (II-A and II-C; lanes 1-6), Day 6-9 (II-A and II-C; lanes 1-6 and II-B and II-D; lane 1), and Day 10-16 (II-B and II-D; lanes 2-9). Lanes M = 100 bp DNA ladder.

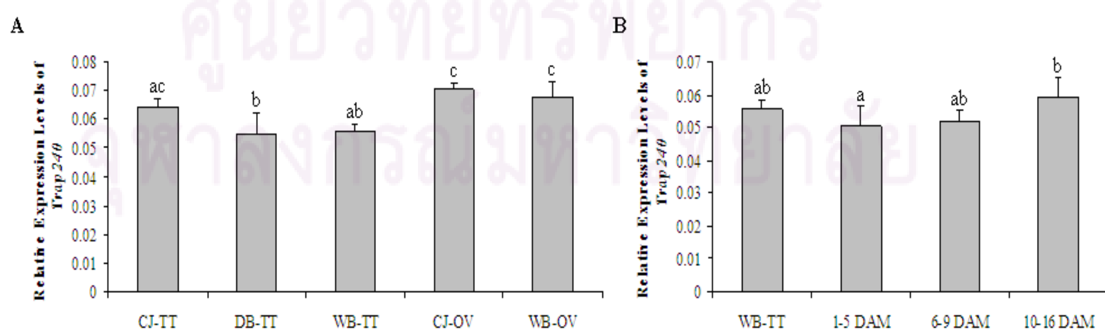


Figure 3.85 Histograms showing the expression levels of *Trap240* in gonads of *P. monodon*. The same letters indicate that the expression levels were not significantly different ($P > 0.05$).

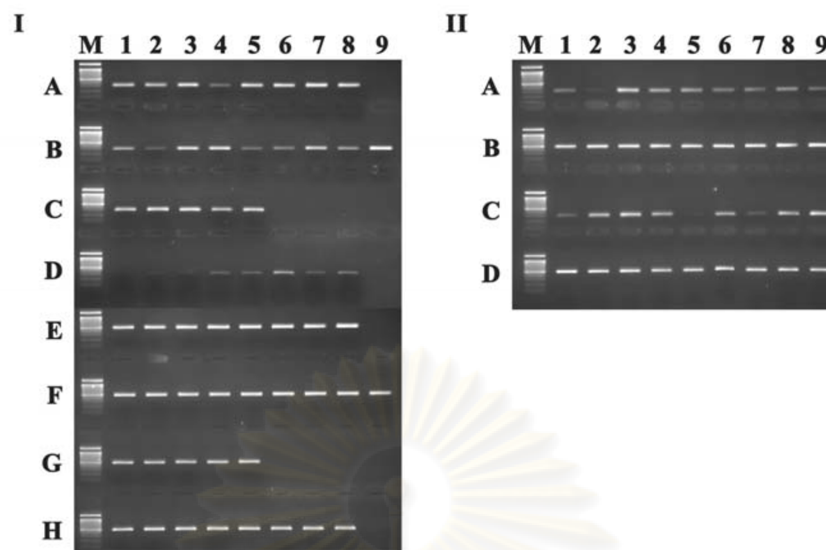


Figure 3.86 A 1.6% ethidium bromide-strained agarose gel showing the expression levels of *HSP70-2* (panel I: A-D and panel II: A and C) and *EF-1α* (panel I: E-H and panel II: B and D) of testis from wild broodstock (WB-TT, I-A and I-E, lanes 1-8), domesticated broodstock (DB-TT, I-B and I-F, lanes 1-9), cultured juveniles (CJ-TT, I-C and I-G, lanes 1-5), and ovaries from wild broodstock (WB-OV, I-D and I-H; lanes 1-4), and cultured juveniles (CJ-OV, I-D and I-H; lanes 5-8) and testes from broodstock after molting for Day 1-5 (II-A and II-C; lanes 1-6), Day 6-9 (II-A and II-C; lanes 1-6 and II-B and II-D; lane 1), and Day 10-16 (II-B and II-D; lanes 2-9). Lanes M = 100 bp DNA ladder.

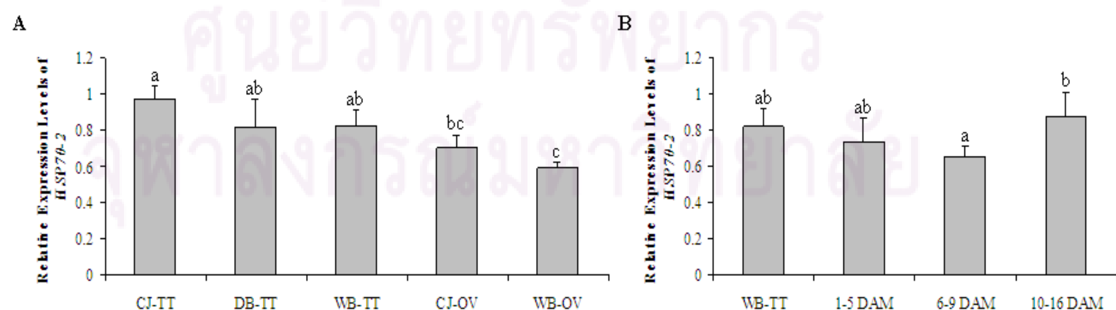


Figure 3.87 Histograms showing the expression levels of *HSP70-2* in gonads of *P. monodon*. The same letters indicate that the expression levels were not significantly different (*P>0.05*).

3.7 Quantitative analysis of interesting genes in testes by real-time PCR

3.7.1 Expression level of *Dmc1* and *progestin receptor membrane component 1 (PGRMC1)* in testes of juvenile *P. monodon* in response to dopamine administration

The expression levels of *Dmc1* and *PGRMC1* in juvenile *P. monodon* in response to dopamine (DA) administration (10^{-6} mol/shrimp) at different time points (3, 6, 12, and 24 hours) were examined using quantitative real-time PCR analysis.

Standard curves of *Dmc1*, *PGRMC1*, and *EF-1α* of were constructed (Fig. 3.88-3.89). For quantitative analysis of specimens, 50 ng of first-strand cDNA was used as the template. Results from real-time PCR revealed that expression levels of *Dmc1* in testes were not significantly altered after dopamine treatment ($P > 0.05$) whereas *PGMRC1* was up-regulated at 3 hr post treatment ($P < 0.05$) (Fig. 3.90). These preliminary results suggested that dopamine might not inhibit spermatogenesis in *P. monodon*. However, further confirmative studies need to be conducted in *P. monodon* broodstock at both mRNA and protein levels. Raw data on relative expression level of these genes are shown in Appendix C.

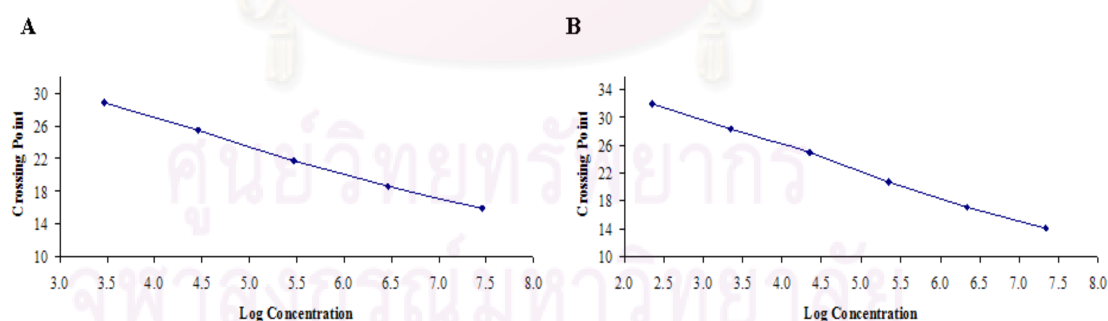


Figure 3.88 Standard curve of *Dmc1* (A; r^2 for the standard curve = 0.996820) and *EF-1α* (B; r^2 for the standard curve = 0.99327) in testes of juvenile *P. monodon*. *EF-1α* was used as an internal control for *Dmc1*.

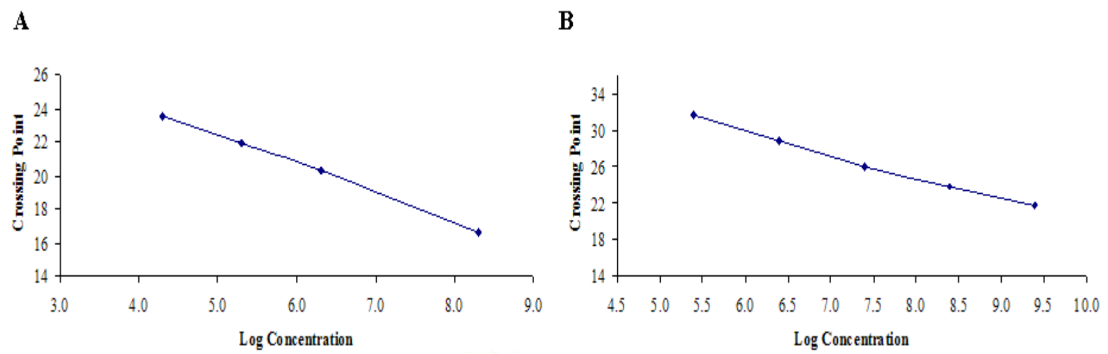


Figure 3.89 Standard curve of *PGRMC1* (A; r^2 for the standard curve = 0.99820) and *EF-1α* (B; r^2 for the standard curve = 0.99520) in testes of juvenile *P. monodon*. *EF-1α* was used as an internal control for *PGRMC1*.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

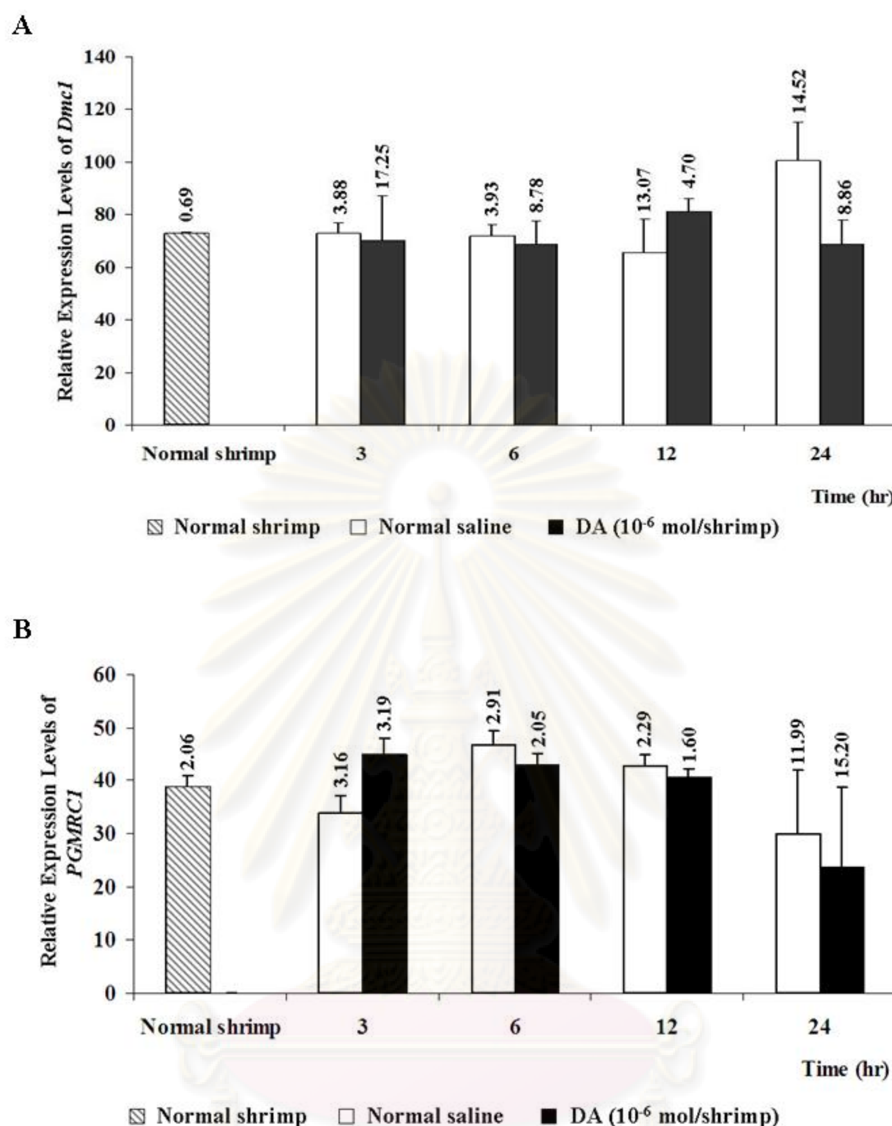


Figure 3.90 Real-time PCR analysis illustrating relative expression levels of *Dmc1* in testes of juvenile *P. monodon* after injected with normal saline or dopamine (10^{-6} mol/shrimp) for 3, 6, 12 and 24 hr. The normal shrimp was also included as the control. An asterisk indicates significant up-regulation of the relative expression level of *PGMRC1* in dopamine-treated juvenile *P. monodon*. Numbers above the histograms reveal standard deviation of the treatment.

3.7.2 Expression levels of *SUMO-1, cyclophilinA (CYA), Dmc1*, *saposin* and *spermatogonial stemcell renewal factor* at different stage of testes and ovaries of *P. monodon*

The expression levels of transcripts related to testicular development including *SUMO-1, cyclophilin A (CYA), Dmc1*, *saposin* and *spermatogonial stemcell renewal factor* in testes of wild *P. monodon* broodstock after molting, wild broodstock (WB-TT), domesticated broodstock (DB-TT), cultured juvenile (CJ-TT) and ovaries from wild broodstock (WB-OV) and cultured juveniles (CJ-OV) were examined using real-time PCR analysis.

The standard curve of each gene was constructed (Fig. 3.91-3.93) using 50 ng of the first strand cDNA template for *SUMO-1*, *CYA*, *saposin, spermatogonial stem cell renewal factor* and *EF-1α* and 75 ng for *Dmc1*.

Quantitative real-time PCR clearly illustrated that *CYA* was more abundantly expressed in ovaries than testes ($P < 0.05$, Fig. 3.94). The expression level of CJ-OV preferentially expressed to WB-OV ($P < 0.05$). In addition, the expression level of this gene in testes at 10-16 DAM was more abundant than that of 6-9 DAM ($P < 0.05$). The expression level of *CYA* in WB-TT was greater than that of 6-9 DAM ($P < 0.05$).

SUMO-1 was significantly expressed lower in domesticated broodstock than in wild and cultivated *P. monodon* males ($P < 0.05$, Fig. 3.94) but not in wild and cultured juvenile females ($P > 0.05$). The expression levels of *SUMO-1* in testes and ovaries were not significantly different ($P > 0.05$). In addition, the expression level of this gene in testes at 10-16 DAM was higher than that at 1-5 and 6-9 DAM ($P < 0.05$). The expression level of *SUMO-1* in WB-TT was higher than that at 1-5 and 6-9 DAM ($P < 0.05$).

Saposin was more abundantly expressed in testes than ovaries of *P. monodon* ($P < 0.05$, Fig. 3.95). The expression levels of this transcript were not significantly different in different stages of testes (WB-TT, DB- TT and CJ-TT) and ovaries (WB-TT and CJ-TT). Moreover, the expression levels of *saposin* in the molting groups

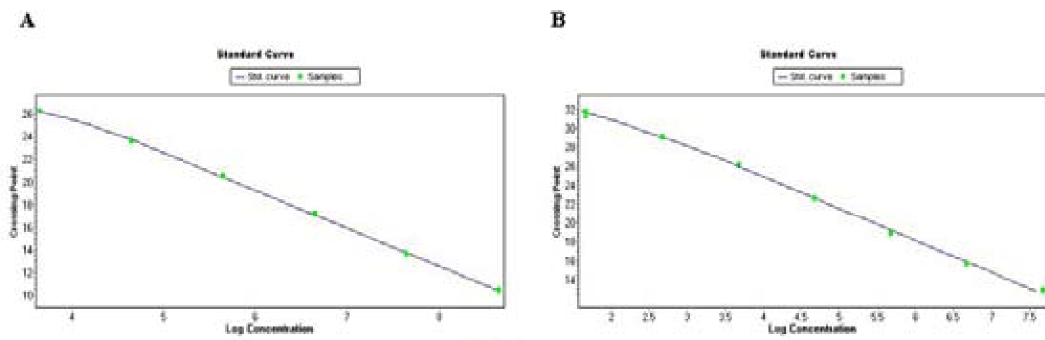


Figure 3.91 Standard curve of *SUMO-1* (efficiency of the amplification = 1.993) and *CYA* (efficiency of the amnplificaqtion = 1.980)

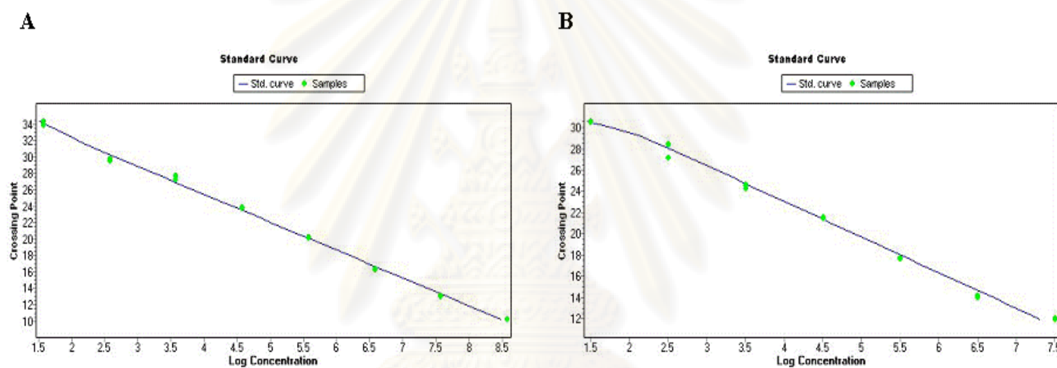


Figure 3.92 Standard curve of *Dmc1* (efficiency for the amplification = 1.971) and *saposin* (efficiency of the amplification = 1.984)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

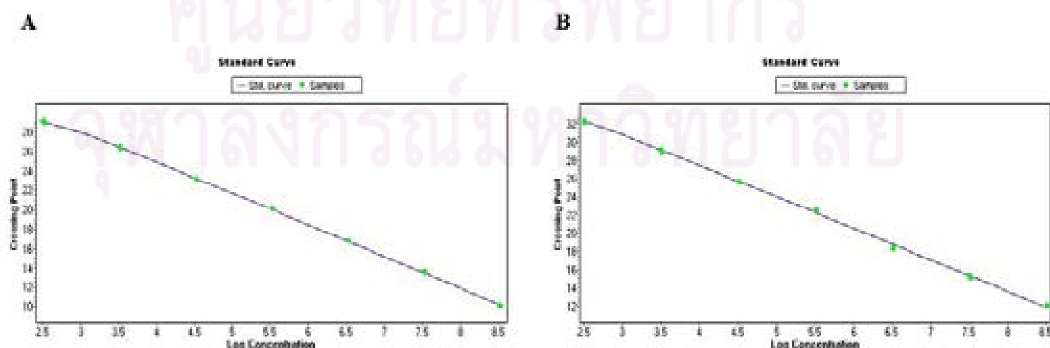


Figure 3.93 Standard curve of *spermatogonial stem-cell renewal factor* (efficiency of the amplification = 2.027). *EF-1α* (efficiency for the standard curve = 1.945)

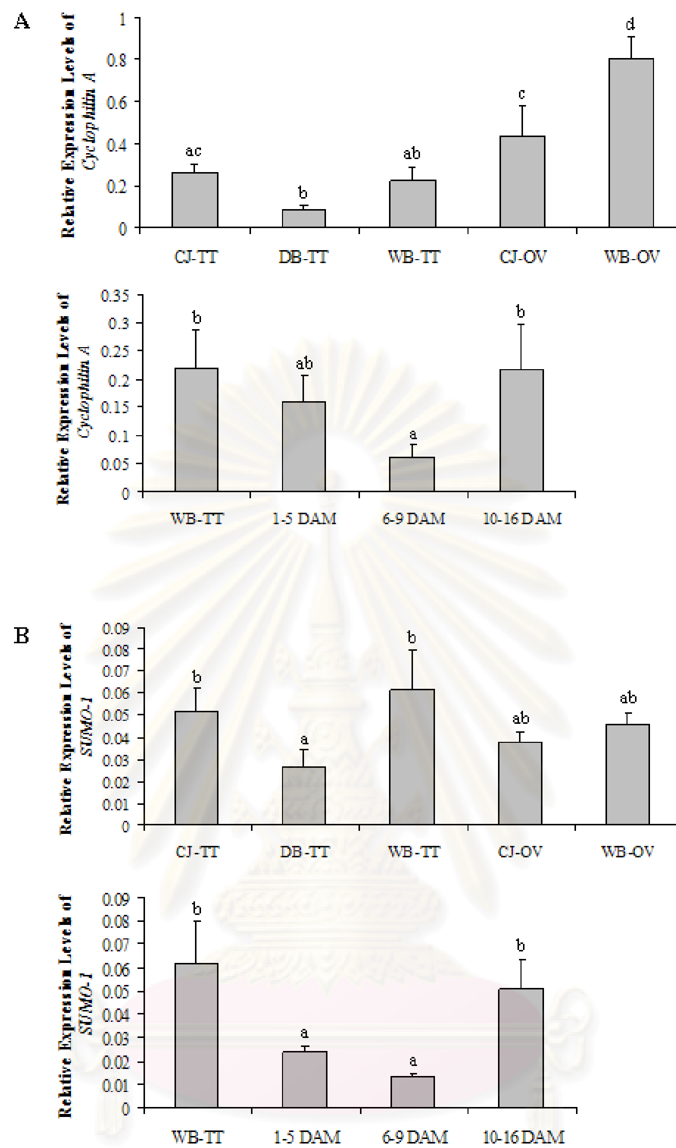


Figure 3.94 Real-time PCR analysis illustrating the relative expression levels of *CYA* (A) and *SUMO-1* (B) in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting. The same letters indicate that the expression levels were not significantly different (*P*>0.05).

CJ-TT = testes of cultured juvenile, WB-TT = testes of wild broodstock, DB-TT = testes of domesticated broodstock, CJ-OV = ovaries of cultured juvenile and WB-OV = ovaries of wild broodstock.

were not different. However, the *saposin* level of WB-TT was greater than that of 1-5 and 6-9 DAM ($P < 0.05$).

Relative expression levels (normalized by *EF-1α*) of *spermatogonial stem-cell renewal factor* was quite low compared to the control. Accordingly, the absolute expression levels (without normalization) were also illustrated and both parameters were considered together. The expression level of *spermatogonial stem-cell renewal factor* in domesticated broodstock was significantly lower than that in wild and cultivated *P. monodon* males, $P < 0.05$, Fig. 3.96) but not in cultured *P. monodon* juveniles ($P > 0.05$).

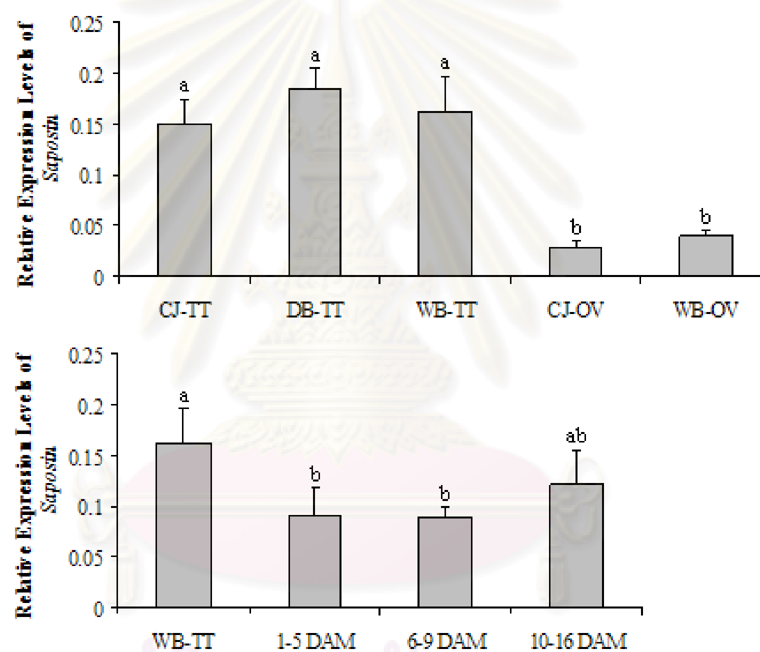


Figure 3.95 Real-time PCR analysis illustrating the relative expression levels of *saposin* in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting. The same letters indicate that the expression levels were not significantly different ($P > 0.05$).

CJ-TT = testes of cultured juvenile, WB-TT = testes of wild broodstock, DB-TT = testes of domesticated broodstock, CJ-OV = ovaries of cultured juvenile and WB-OV = ovaries of wild broodstock.

The relative expression level of *spermatogonial stem-cell renewal factor* was preferentially expressed in testes to ovaries of *P. monodon* ($P<0.05$). In addition, both relative and absolute expression levels of this gene in testes of 10-16 DAM was more abundant than those at 1-5 and 6-9 DAM ($P < 0.05$) and the expression level of this gene in WB-TT was significantly greater than that of the molting groups ($P < 0.05$). When absolute expression levels of *spermatogonial stem-cell renewal factor* were considered the highest level was found in WB-OV resulting in a preferential expression pattern of this gene in ovaries ($P < 0.05$).which is contradictory to results inferred from the relative expression levels.

Dmc1 was less abundantly expressed in testes of domesticated broodstock than in wild and cultivated *P. monodon* males ($P < 0.05$, Fig. 3.97) but not significantly different from ovaries of both cultured juveniles and wild broodstock ($P > 0.05$). The expression level of this gene was more preferential in testes to ovaries ($P < 0.05$). In addition, both relative and absolute expression level of this gene in testes of 10-16 DAM was greater than those at 1-5 and 6-9 DAM ($P < 0.05$). However, the levels of *Dmc1* were not different when absolute expression levels rather than relative expression levels were considered.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

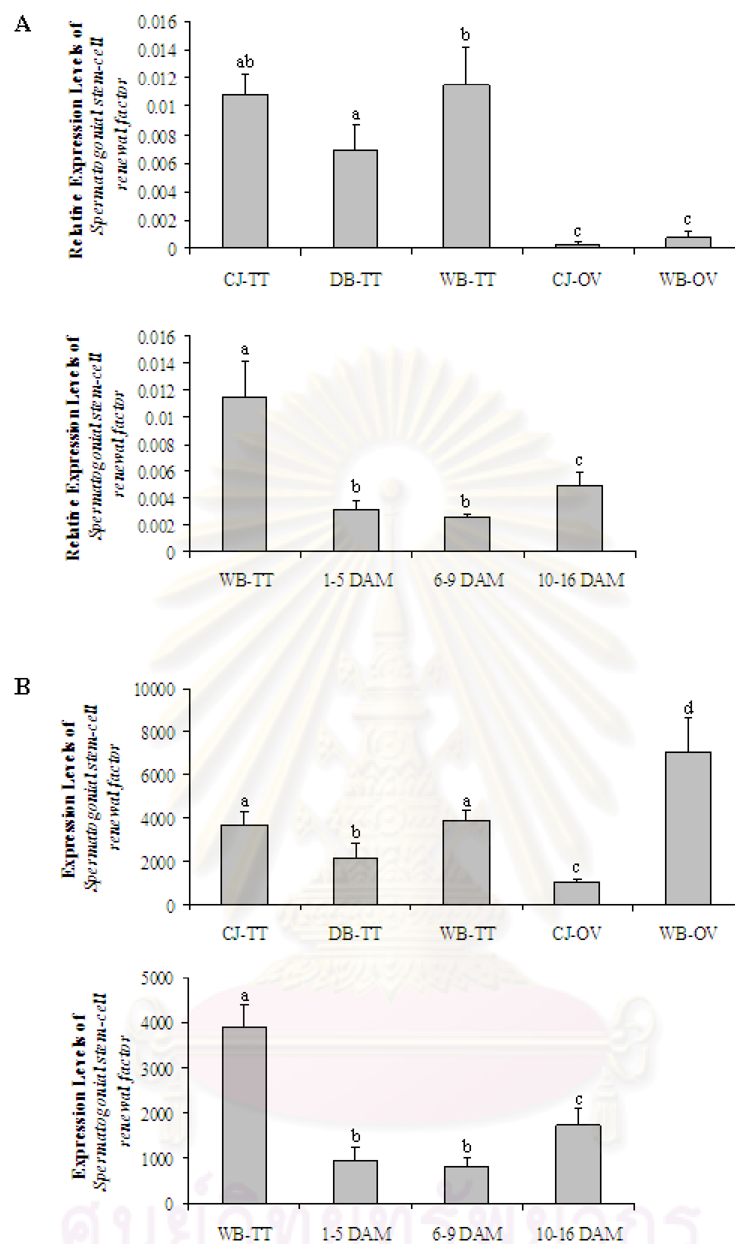


Figure 3.96 Real-time PCR analysis illustrating the relative (A) and absolute (B) expression levels of *spermatogonial stem cell renewal factor* in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting. The same letters indicate that the expression levels were not significantly different ($P > 0.05$).

CJ-TT = testes of cultured juvenile, WB-TT = testes of wild broodstock, DB-TT = testes of domesticated broodstock, CJ-OV = ovaries of cultured juvenile and WB-OV = ovaries of wild broodstock.

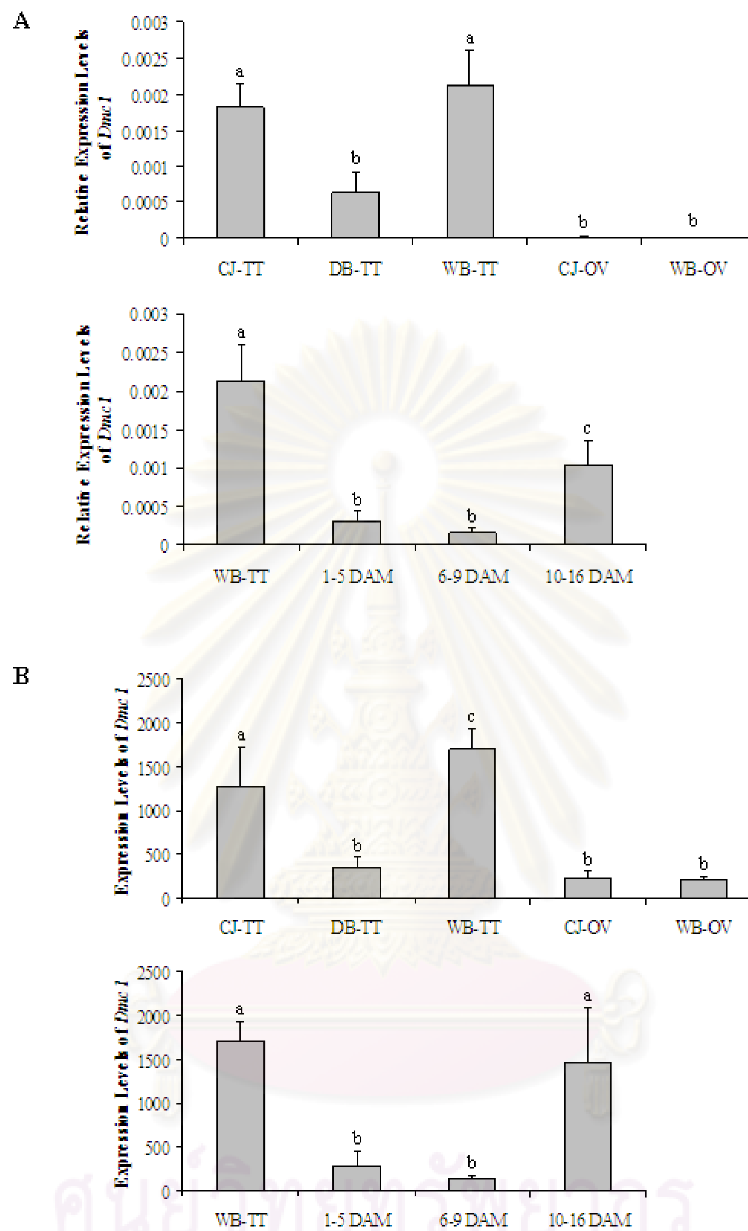


Figure 3.97 Real-time PCR analysis illustrating the relative (A) and absolute (B) expression levels of *Dmc1* in testes and ovaries of *P. monodon* and in testes of *P. monodon* broodstock after molting. The same letters indicate that the expression levels were not significantly different (P>0.05).

CJ-TT = testes of cultured juvenile, WB-TT = testes of wild broodstock, DB-TT = testes of domesticated broodstock, CJ-OV = ovaries of cultured juvenile and WB-OV = ovaries of wild broodstock.

3.8 Large scale examining expression patterns of genes in testes of *P. monodon* by microarray analysis

Microarray slides fabricated with transcripts from *V. harveyi* infected hemocyte and non-infected hemocyte cDNA libraries of *P. monodon* and WSSV infected hemocyte and non-infected hemocyte cDNA libraries of *M. japonicus* were used for microarray analysis. A total of spotted cDNA was 2,036 ESTs composing of 1282 and 754 ESTs from cDNA libraries of *P. monodon* and *M. japonicus*, respectively. The former consisted of 408 EST representing known genes and 874 ESTs representing unknown genes. The latter contained 254 and 469 ESTs representing known and unknown genes, respectively.

3.8.1 Gene differentially expressed in testes of juvenile and broodstock *P. monodon*

Testis cDNA of juvenile and broodstock *P. monodon* and *vice versa* were labeled with Cy3 and Cy5 and used as the probes for microarray analysis. Only transcripts exhibiting ≥ 2 folds or ≤ 0.5 fold between the target and the control specimens were considered as up- or down-regulation.

The number of up- or down-regulated genes in testes of broodstock (cDNAs in testes of juvenile shrimp were used as the control, Fig. 3.98) and in testes of juvenile shrimp (cDNAs in testes of broodstock were used as control, Fig. 3.99) are showed in Table 3.16.

In *P. monodon* broodstock (WB-TT), 350 of 1891 positively detected transcripts (18.5%) revealed differential expression patterns. Of these, 138 (39.4% of differentially expressed transcripts) and 212 (60.6%) transcripts were up- and down-regulated, respectively.

In juvenile *P. monodon* (CJ-TT), 105 of 1792 positively detected transcripts (5.9%) revealed differential expression patterns. Of these, 54 (51.4% of differentially expressed transcripts) and 51 (48.6%) transcripts were up- and down-regulated, respectively.

Table 3.16 A summary on the number of up- or down-regulated genes in testes of broodstock and juvenile *P. monodon* analyzed by microarrays

	Broodstock	**Juvenile**
Total number of signally detected genes	**1891**	**1792**
Up-regulated genes	**138**	**54**
-Up-regulated known genes	44	34
-Up-regulated unknown genes	94	20
Down-regulated genes	**212**	**51**
-Down-regulated known genes	63	9
-Down-regulated unknown genes	149	42
Total number of up-/down-regulated genes	**350**	**105**

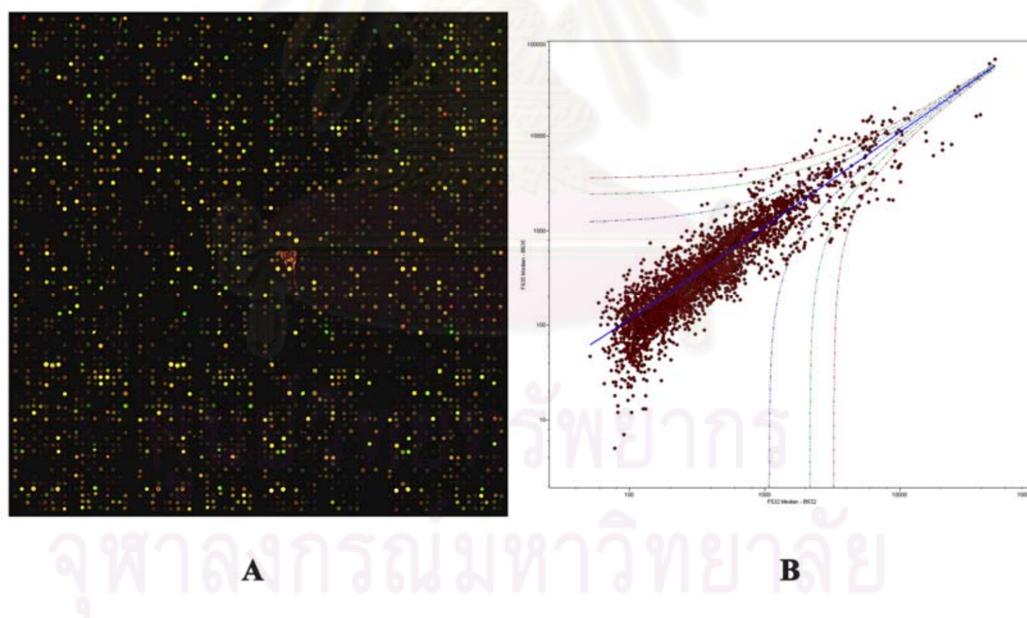


Figure 3.98 Microarray analysis using cDNA of testes of juveniles and broodstock of *P. monodon* labeled with Cy3 and Cy5 as the probes, respectively.

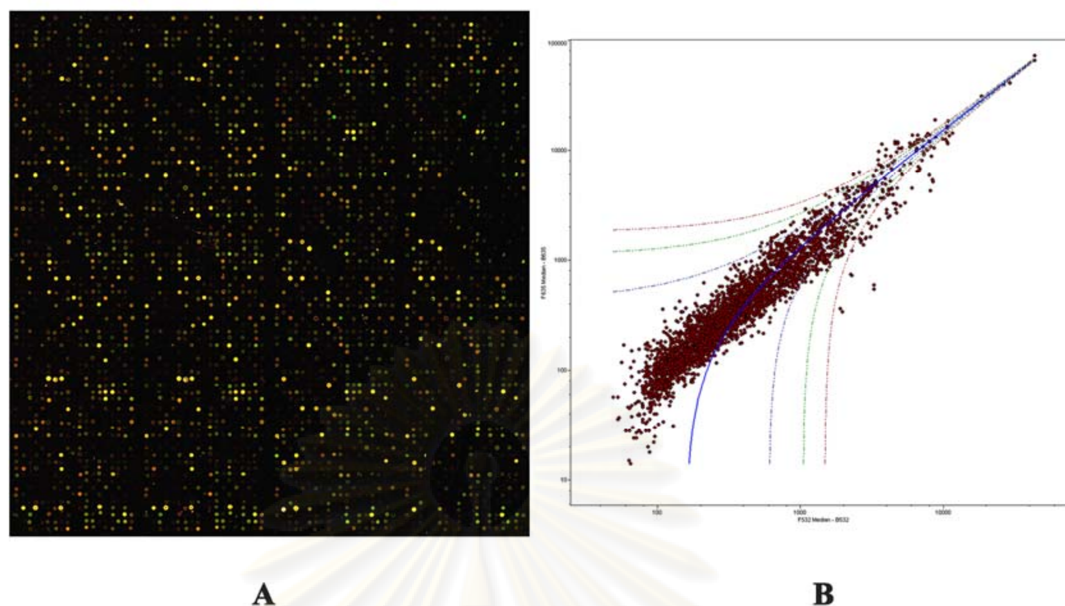


Figure 3.99 Microarray analysis using cDNA of broodstock and juvenile testes of *P. monodon* labeled with Cy3 and Cy5 as the probes, respectively.

Table 3.17 Examples of up- or down-regulated genes in testes of broodstock in comparison with those in testes of juvenile *P. monodon*

No.	Name	ID	Species of the original transcript	Tested samples	Ratio of Medians (635/532)	Up- or down-regulation
Up-regulated in WB-TT but down-regulated in CJ-TT						
I-A	unknown	LPV0054	*P. monodon*	CJ-TT	0.19	Down
I-B				WB-TT	5.427	Up
II-A	unknown	IF174	*P. monodon*	CJ-TT	0.182	Down
II-B				WB-TT	6.124	Up
III-A	unknown	IF361	*P. monodon*	CJ-TT	0.492	Down
III-B				WB-TT	2.276	Up
Up-regulated in CJ-TT but down-regulated in WB-TT						
IV-A	unknown	HpaN0480	*P. monodon*	CJ-TT	2.482	Up
IV-B				WB-TT	0.446	Down
V-A	unknown	LPV0175	*P. monodon*	CJ-TT	2.01	Up
V-B				WB-TT	0.383	Down

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Three unknown genes were up-regulated in testes of *P. monodon* broodstock but down-regulated in those of *P. monodon* juvenile. There were LPV0054, IF174, and IF361 from cDNA libraries of *P. monodon* (Table 3.17 and Fig.3.100). In contrast, two unknown genes (HpaN0480 and LPV0175) were down-regulated in *P. monodon* broodstock but up-regulated in juvenile *P. monodon* (Table 3.17 and Fig.3.100).

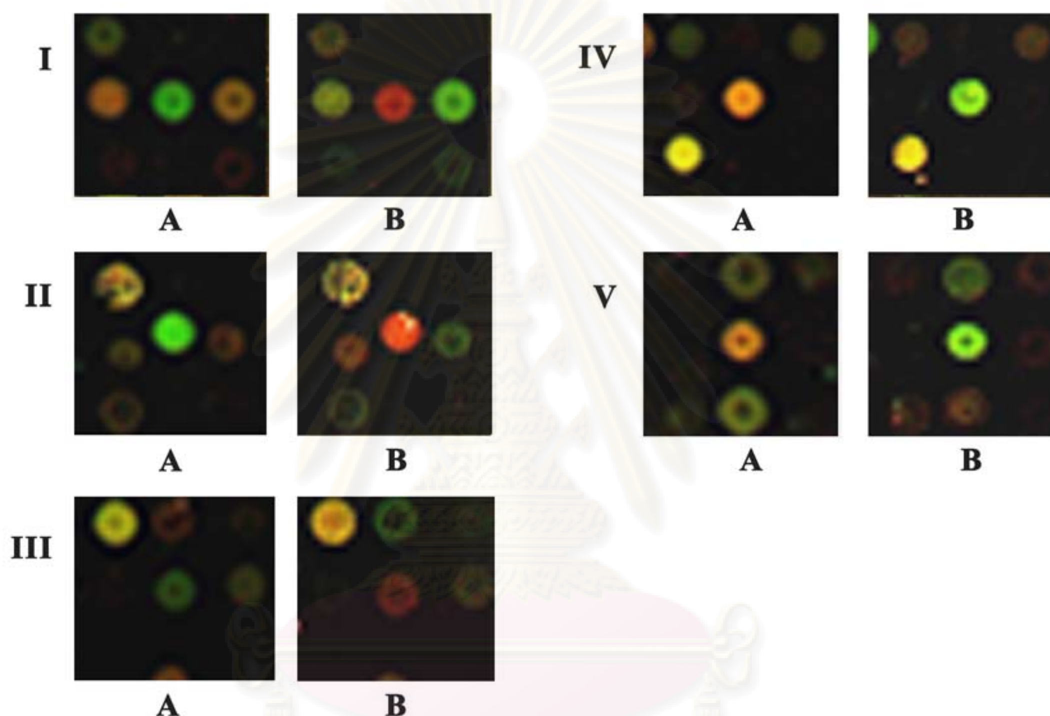


Figure 3.100 Microarray analysis showing up- and/or down-regulation of examined genes in testes of juvenile (A) and broodstock (B) *P. monodon*.

3.8.2 Gene expression of dopamine-treated testes at 3 and 6 hours post injection of juvenile *P. monodon*

Genes differentially expressed in testes of normal and dopamine-treated juvenile shrimps (10^{-6} mol/shrimp) at 3 and 6 hours post injection were examined. Testis cDNA of normal and dopamine-treated juvenile shrimps at 3 and 6 hours were used as probes by labeling with Cy3 and Cy5, respectively.

The numbers of up- or down-regulated genes of dopamine-treated testes at 3 hours and 6 hours post injection (testis cDNA of normal shrimp were used as the control) were shown in Table 3.18 and Fig. 3.101-3.102, respectively.

In comparison between dopamine-treated shrimp at 3 hr post injection and normal shrimp, 1924 ESTs gave the positive signal and 274 of these (14.2%) displayed differential expression patterns. A total of 186 (67.9% of differentially expressed transcripts) and 88 (32.1%) transcripts were up- and down-regulated, respectively.

Likewise, 1889 positively hybridized signals were observed when cDNAs of dopamine-treated shrimp at 6 hr post injection were compared with those of the normal shrimp. A total of 229 (12.1%) exhibited differential expression and 113 (49.3%) and 116 (50.7%) transcripts were up- and down-regulated, respectively.

Four unknown transcripts were up-regulated at 3hr but returned to the normal state at 6 hr post dopamine injection. They were PJH145r, PJA128r, HpaN0140, and LPV0054 (Table 3.19 and Fig.3.103). In contrast, an unknown N331 transcript did not revealed differential expression at 3 hr but was up-regulated at 6 hr post dopamine injection. In addition, two unknown genes (IF174 and PJA108f) were up-regulated at both 3 hr and 6 hr post dopamine injection.

Table 3.18 A summary on the number of up- or down-regulated genes of testes of dopamine-treated *P. monodon* at 3 hours and 6 hours post injection

	Hours post injection	
	3 hr	**6 hr**
Total number of signally detected genes	**1924**	**1889**
Up-regulated genes	**186**	**113**
- Up-regulated known genes	65	49
- Up-regulated unknown genes	121	64
Down-regulated genes	**88**	**116**
- Down-regulated known genes	16	23
- Down-regulated unknown genes	72	93
Total number of up- or down-regulated genes	**274**	**229**

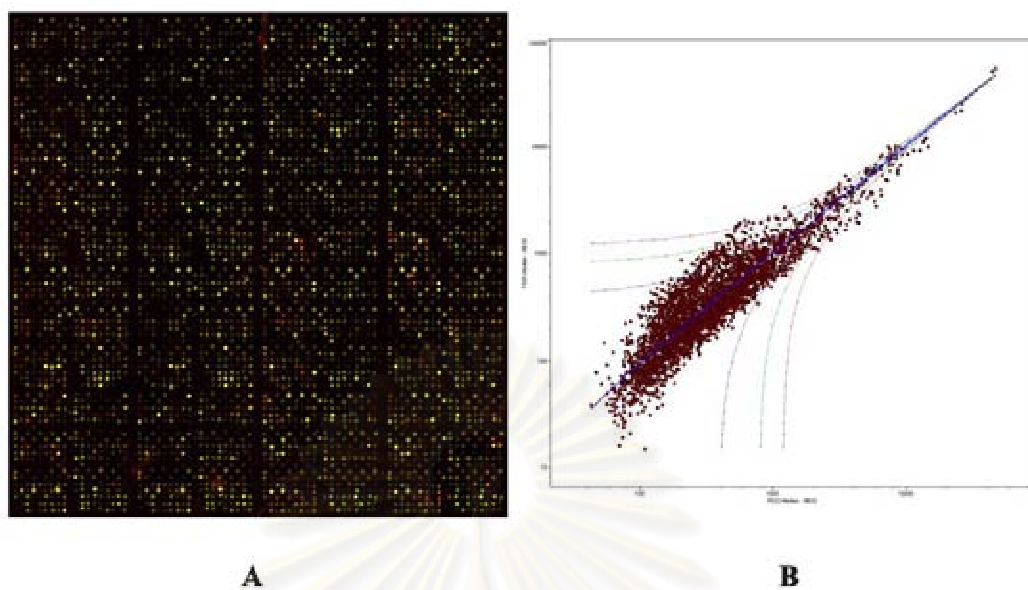


Figure 3.101 Microarray analysis using cDNA of normal testes and dopamine-treated testes of juvenile shrimps (10^{-6} mol/shrimp) at 3 hours post injection labeled with Cy3 and Cy5 as the probes, respectively.

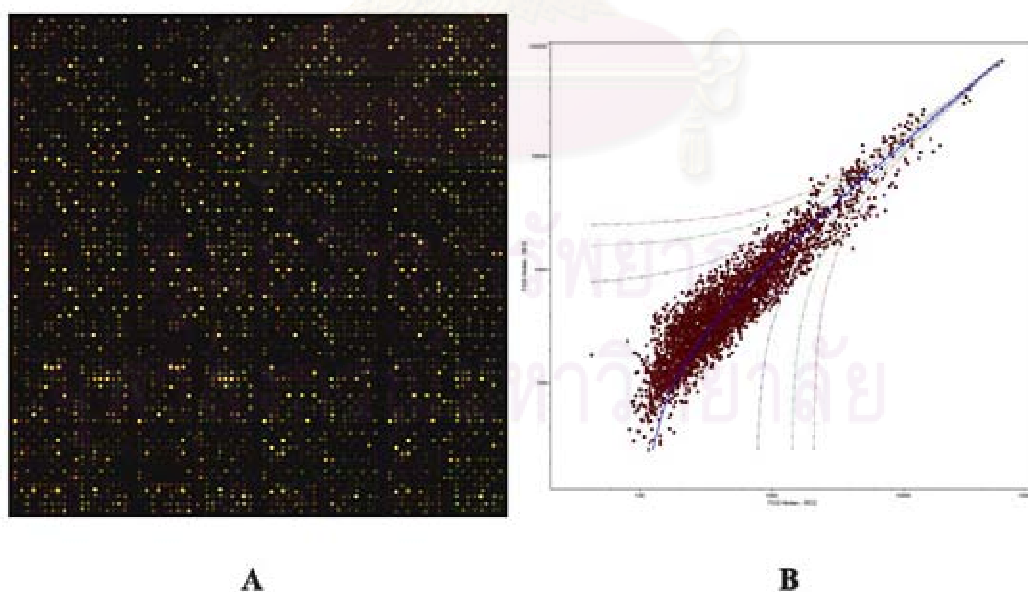


Figure 3.102 Microarray analysis using cDNA of normal testes and dopamine-treated testes of juvenile shrimps (10^{-6} mol/shrimp) at 6 hours post injection labeled with Cy3 and Cy5 as the probes, respectively.

Table 3.19 Examples of up- or down-regulated genes in testes of normal and dopamine-treated juveniles of *P. monodon*

No.	Name	ID	Species		Ratio of Medians (635/532)	Up- or down-reguration
Up- regulated at 3 hr but not at at 6 hr post dopamine injection						
I-A	unknown	PJH145r	*M. japonicus*	3 hr	3.165	Up
I-B				6 hr	1.238	Normal
I-A	unknown	PJA128r	*M. japonicus*	3 hr	3.368	Up
I-B				6 hr	1.362	Normal
II-A	unknown	HpaN0140	*P. monodon*	3 hr	3.251	Up
II-B				6 hr	1.743	Normal
III-A	unknown	LPV0054	*P. monodon*	3 hr	3.093	Up
III-B				6 hr	1.834	Normal
Up- regulated at both 3 hr and 6 hr post dopamine injection						
IV-A	unknown	IF174	*P. monodon*	3 hr	3.459	Up
IV-B				6 hr	2.400	Up
V-A	unknown	PJA108f	*M. japonicus*	3 hr	3.255	Up
V-B				6 hr	2.579	Up
Up-regulated at 6 hr but not at 3 hr post dopamine injection						
VI-A	unknown	N331	*P. monodon*	3 hr	1.692	Normal
VI-B				6 hr	3.316	Up

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

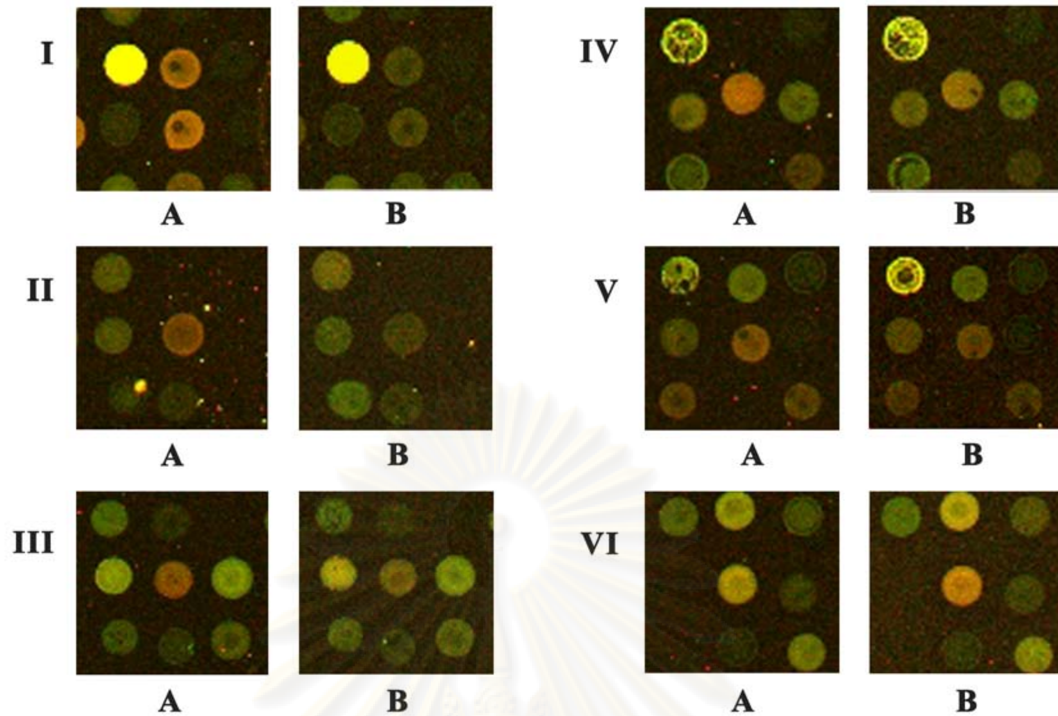


Figure 3.103 Microarray analysis showing up- and/or down-regulation of examined genes in testes of dopamine-treated juveniles of *P. monodon* at 3 hr (A) and 6 hr (B) post injection.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

3.9 *In vitro* expression of interesting genes using the bacterial expression system

3.9.1 Construction of recombinant plasmid in cloning and expression vector

Five recombinant plasmids carrying the full length cDNA of *spermatogonial stem cell renewal factor, Dmc1*, *PGRMC1*, *cyclophilin A*, and *SUMO-1* were successfully constructed for *in vitro* expression of the corresponding recombinant protein.

Spermatogonial stem cell renewal factor was initially found from the ovary cDNA library and the full length cDNA of this gene in ovaries was characterized by RACE-PCR. The full length cDNA of ovarian *spermatogonial stem cell renewal factor* was 2759 bp in length with an ORF of 1353 bp corresponding to 450 amino acids. This transcript significantly matched that of *Danio rerio* (E-value = 6e-109) (Sittikankeaw, 2006). The ORF of the testis form was successfully amplified using primers designed from the full length cDNA from ovaries. The ORF sequence of testis form significantly matched that of *Danio rerio* (E-value = 3e-103). Nucleotide sequence and deduced amino acid of both ORF were aligned (Fig. 3.104-3.105). Eleven polymorphic nucleotides causing 6 nonsynonymous mutations between these isoforms were observed.

A primer pair was designed to amplify cDNA representing the mature peptide of *spermatogonial stem cell renewal factor*, *Dmc1* and *PGRMC1*. The amplified full length cDNA was ligated, cloned into pGEM-T easy vector and transformed into *E. coli* JM109.

In constrast, recombinant plasmid containing the full length cDNA of *SUMO*-1 and *cyclophilin A* were already existent in testis cDNA library. Plasmid DNA of the positive clone was re-sequenced to confirm the orientation and nucleotide sequence of a particular recombinant clone. Recombinant plasmid in the cloning vector was used as the template for amplification of the fragment containing a mature peptide of *SUMO*-1 and *cyclophilin A*. The amplification product was digested with *Nde* I and *Eco* RI, eluted, and ligated into pET32a expression vector and transformed into *E. coli* BL21(DE3) codon+ RIPL.

```
ORF-TT      ATGAGCGCTGCACAGACCTCTCAAGGCAGCTGGAGGATTCCGGACCTTCTTGCCATGAAG 60
ORF-OV      ATGAGCGCTGCACAGACCTCTCAAGGCAGGTGGAGGATCCCGGACCTTCTGTCCATGAAG 60
            ***************************** ******** ***********  ********
ORF-TT      CGCGATGGCCTGGCCTATTCGGAGGACCAGATCGCCTTCTTGGTCCGGTCGGTCTCGGAT 120
ORF-OV      CGCGATGGCCTGGCCTATTCGGAGGACCAGATCGCCTTCTTGGTCCGGTCGGTCTCGGAT 120
            ************************************************************
ORF-TT      CGGTCCATGGACGACTGTCAGCTGGGGGCGCTCCTGATGGCCATCAAGCTGCAGGATATG 180
ORF-OV      CGGTCCATGGACGACTGTCAGCTGGGGGCGCTCCTGATGGCCATCAAGCTGCAGGATATG 180
            ************************************************************
ORF-TT      ACGGACGCAGAGACGATCGCCCTCACTAAGGGCATGAGGGACTCAGGAAGCGTGTTCTCG 240
ORF-OV      ACGGACGTAGAGACGATCGCCCTCACTAAGGGCATGAGGGACTCAGGAAGTGTGTTCTCG 240
            ******* ****************************************** *********
ORF-TT      TGGCCGAAGGACTGGCGCGTCGTGGACAAGCACAGCACGGGCGGCGTGGGTGACAAGGTG 300
ORF-OV      TGGCCGAAGGACTGGCGCGTCGTGGACAAGCACAGCACGGGCGGCGTGGGTGACAAGGTG 300
            ************************************************************
ORF-TT      TCCCTGGCGCTGGCCCCCGCCCTCGCCGCCTGCGGCTTCAAGGTGCCCATGATCAGCGGG 360
ORF-OV      TCCCTGGCGCTGGCCCCCGCCCTCGCCGCCTGCGGCTTCAAGGTGCCCATGATCAGCGGG 360
            ************************************************************
ORF-TT      CGCGGCCTCGAGCACACGGGGGGCACGCTCGACAAGCTGGAGAGCATCCCAGGCTTCAAG 420
ORF-OV      CGCGGCCTCGAGCACACGGGGGGCACGCTCGACAAGCTGGAGAGCATCCCAGGCTTCAAG 420
            ************************************************************
ORF-TT      GTGTCTCTGACGGAGGCCGAGATGAAGACGGCGCTGGAGGAGGTCGGCTGCTGTATCGTA 480
ORF-OV      GTGTCTCTGACGGAGGCCGAGATGAAGACGGCGCTGGAGGAGGTCGGCTGCTGTATCGTA 480
            ************************************************************
ORF-TT      GGCCAGACCGCCGACATCGTACCTGCTGACAGACGCATGTACGCCGCAAGAGACGTCACT 540
ORF-OV      GGCCAGACCGCCGACATCGTACCTGCTGACAGACGCATGTACGCCGCAAGAGACGTCGCT 540
            ********************************************************* **
ORF-TT      TCAACCGTCAAATCTGTGCCGCTCATCGTCTCGTCCATCATCAGCAAGAAGGCTGCGGAG 600
ORF-OV      TCAACCGTCAAATCTGTGCCGCTCATCGTCTCGTCCATCATCAGCAAGAAGGCTGCGGAA 600
            *********************************************************** 
ORF-TT      ACCGTGAGCGGGCTGGTGCTCGACGTCAAGTTCGGCGGAGGAGCCTTCATGAAGACCCAG 660
ORF-OV      ACCGTGAGCGGGCTGGTGCTCGACGTCAAGTTCGGCGGAGGAGCCTTCATGAAGACCCAG 660
            ************************************************************
ORF-TT      GAGGAGGCAGGGGCGCTGGCCAAGAAAATGGTGGATGTGGCCAACGGCGTGGGCATGGCC 720
ORF-OV      GAGGAGGCAGGGGCGCTGGCCAAGAAAATGGTGGATGTGGCCAACGGCGTGGGCATGGCC 720
            ************************************************************
ORF-TT      ACGACGGCCCTCCTGACCACGATGGATATCCCGCTCGGCAGGGCCATCGGCAACGCCCTC 780
ORF-OV      ACGACGGCCCTCCTGACCACGATGGATATCCCGCTCGGCAGGGCCATCGGCAACGCCCTC 780
            ************************************************************
ORF-TT      GAGGTGCGGGAGTCGCTGGAGTGTCTTCGGGGCAACGGACCGGAGGACCTTGAGGAGCTC 840
ORF-OV      GAGGTGCGGGAGTCGCTGGAGTGTCTTCGGGGCAACGGACCGGAGGACCTTGAGGAGCTC 840
            ************************************************************
ORF-TT      GTAACGCACCTGGGCGGAGAATTACTGCTGGGTGCGGGAGCGGCCTCTACGCTGGATGAA 900
ORF-OV      GTAACGCACCTGGGCGGAGAATTACTGCTGGGTGCGGGAGCGGCCTCTACGCTGGATGAA 900
            ************************************************************
ORF-TT      GCTCGCCAGAAGCTGGCCAAGGCTCTGAGGGATGGCAGTGCCAGAACGGCTTTCTGCAAT 960
ORF-OV      GCTCGCCAGAAGCTGGCCAAGGCTCTGAGGGATGGCAGTGCCAGAACGGCTTTCTGCAAT 960
            ************************************************************
ORF-TT      ATGATACAGAAGCAGGGTGTCACCAAGAGTGTAGCAGAGGCACTGTGCGGCAATGTTCCC 1020
ORF-OV      ATGATACAGAAGCAGGGTGTCACCAAGAGTGTAGCAGAGGCACTGTGCGGCAATGTTCCC 1020
            ************************************************************
ORF-TT      GACTACTCCCATCTACCTTCCTCGGCTCATGTCACTGCCGTCAAAGCTGCTTCCCCAGGA 1080
ORF-OV      GACTACTCCCATCTACCTTCCTCGGCTCATGTCACTGCCGTCAAAGCTGCTTCCTCAGGA 1080
            ****************************************************** *****
ORF-TT      GTGCTAGTTGGTATGGATGCCATGACTATGGCGAAGATCAGTTTAGAACTCGGGGCTGGC 1140
ORF-OV      GTGCTAGTTGGTATGGATGCCATGACTATGGCGAAGATCAGTTTAGAACTCGGGGCTGGC 1140
            ************************************************************
ORF-TT      AGGAACAAGGTCGGCGACCCGATCAACTACAGCGTGGGCATAATGCTCGTTAAGGTGGTC 1200
ORF-OV      AGGAACAAGGTCGGCGACCCGATCAACTACAGCGTGGGCATAATGCTCATTAAGGTGGTC 1200
            ************************************************ ***********
ORF-TT      GGCGAGAGCGTGAAGGAAGGCGAGACGTGGGCAGAGCTGCATCACGATTCCTCACTGCCA 1260
ORF-OV      GGCGAGAGCGTGAAGGAAGGCGAGACGTGGGCAGAGCTGCACCACGATTCCTCACTGCCA 1260
            ***************************************** ******************
ORF-TT      CCCACCCTCCTACAGAGGATGCAGGGAGCCGTCACCATCAAGGCGTCGGCGGAAGCATGC 1320
ORF-OV      CCCACCCTCCTACAGAGGATGCAGGGAGCCGTCACCATCAAGGCGTCGGCGGAAGCATGC 1320
            ************************************************************
ORF-TT      AAGCCCTCGCGCGTTGCCGCTCGCGTTGTCTAG 1353
ORF-OV      AAGCCCTCGCGCGTTGCCGCTCGCGTTGTCTAG 1353
            *********************************
```

Figure 3.104 Alignments between the full length cDNA of *spermatogonial stem-cell renewal factor* initially characterized from testes (ORF-TT) and ovaries (ORF-OV) of *P. monodon*

```
ORF-TT      MSAAQTSQGSWRIPDLLAMKRDGLAYSEDQIAFLVRSVSDRSMDDCQLGALLMAIKLQDM 60
ORF-OV      MSAAQTSQGRWRIPDLLSMKRDGLAYSEDQIAFLVRSVSDRSMDDCQLGALLMAIKLQDM 60
            ********* *******:******************************************

ORF-TT      TDAETIALTKGMRDSGSVFSWPKDWRVVDKHSTGGVGDKVSLALAPALAACGFKVPMISG 120
ORF-OV      TDVETIALTKGMRDSGSVFSWPKDWRVVDKHSTGGVGDKVSLALAPALAACGFKVPMISG 120
            **.*********************************************************

ORF-TT      RGLEHTGGTLDKLESIPGFKVSLTEAEMKTALEEVGCCIVGQTADIVPADRRMYAARDVT 180
ORF-OV      RGLEHTGGTLDKLESIPGFKVSLTEAEMKTALEEVGCCIVGQTADIVPADRRMYAARDVA 180
            ***********************************************************:

ORF-TT      STVKSVPLIVSSIISKKAAETVSGLVLDVKFGGGAFMKTQEEAGALAKKMVDVANGVGMA 240
ORF-OV      STVKSVPLIVSSIISKKAAETVSGLVLDVKFGGGAFMKTQEEAGALAKKMVDVANGVGMA 240
            ************************************************************

ORF-TT      TTALLTTMDIPLGRAIGNALEVRESLECLRGNGPEDLEELVTHLGGELLLGAGAASTLDE 300
ORF-OV      TTALLTTMDIPLGRAIGNALEVRESLECLRGNGPEDLEELVTHLGGELLLGAGAASTLDE 300
            ************************************************************

ORF-TT      ARQKLAKALRDGSARTAFCNMIQKQGVTKSVAEALCGNVPDYSHLPSSAHVTAVKAASPG 360
ORF-OV      ARQKLAKALRDGSARTAFCNMIQKQGVTKSVAEALCGNVPDYSHLPSSAHVTAVKAASSG 360
            **********************************************************.*

ORF-TT      VLVGMDAMTMAKISLELGAGRNKVGDPINYSVGIMLVKVVGESVKEGETWAELHHDSSLP 420
ORF-OV      VLVGMDAMTMAKISLELGAGRNKVGDPINYSVGIMLIKVVGESVKEGETWAELHHDSSLP 420
            ************************************:***********************

ORF-TT      PTLLQRMQGAVTIKASAEACKPSRVAARVV 450
ORF-OV      PTLLQRMQGAVTIKASAEACKPSRVAARVV 450
            ******************************
```

Figure 3.105 Alignments of deduced amino acid sequences of *spermatogonial stem-cell renewal factor* from testes (ORF-TT) and ovaries (ORF-OV) of *P. monodon*

3.9.2 *In vitro* expression of recombinant protein

Expression of 4 recombinant clones of *spermatogonial stem-cell renewal factor* (47.21 kDa), *Dmc1* (37.54 kDa), *SUMO-1* (10.58 kDa), and *cyclophilin A* (18.86 kDa) and 8 clones of PGMRC1 (20.98 kDa) after induced by IPTG at 37°C were examined. Four respective proteins were found at the expected sizes after induced by IPTG for 3 hr. The recombinant proteins were stably expressed at 6 hr post IPTG induction (Fig. 3.106-107). Although recombinant PGRMC1 was also induced by IPTG, the levels of expressed PGMRC1 was very low. The extension of culture period from 6 hr to overnight at 37°C did not resolve a problem from low expression of this recombinant protein (Fig. 3.108). As a result, in vitro expression of PGMRC1 was not carried out further.

One of four examined recombinant clones of each protein was selected and the expression of the corresponding recombinant protein was examined at 0, 1, 2, 3, 6, 12 and 24 hr after IPTG induction at 37°C. Recombinant spermatogonial stem-cell renewal factor, Dmc1, SUMO-1 and cyclophilin A proteins were overexpressed since 1 hr after induced by IPTG and (Fig 3.109-3.112). The expression levels of recombinant spermatogonial stem cell renewal factor were comparable between 1-12 hr after IPTG induction. However, the overnight culture resulted in a slightly lower level of this expressed protein.

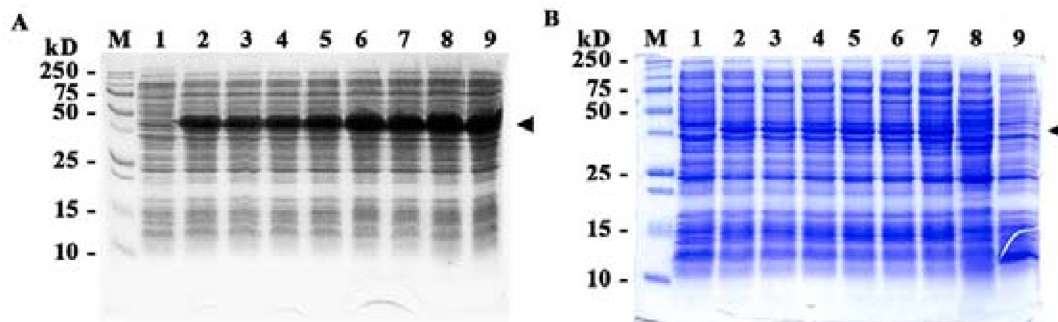


Figure 3.106 SDS-PAGE showing *in vitro* expression of four recombinant clones of *spermatogonial stem-cell renewal factor* (A) and *Dmc1* (B) of *P. monodon* at 0 hr (lane 1), 3 hr (lanes 2-5) and 6 hr (lanes 6-9) after induced by 1.0 mM IPTG.

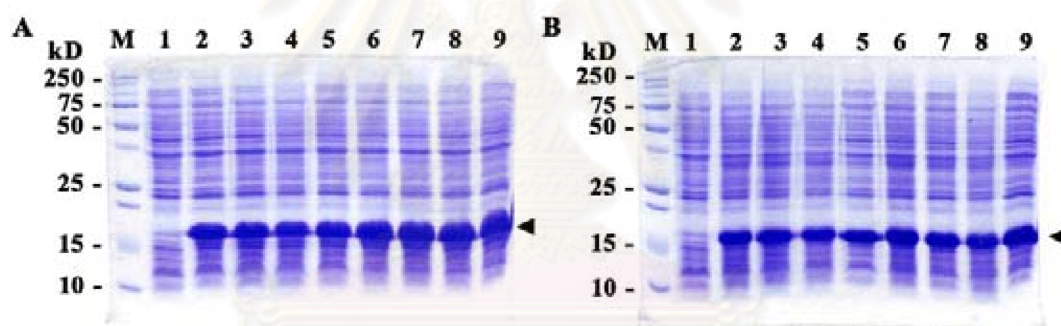


Figure 3.107 SDS-PAGE showing *in vitro* expression of four recombinant clones of *SUMO-1* (A) and *cyclophilin A* (B) of *P. monodon* at 0 hr (lane 1), 3 hr (lanes 2-5), and 6 hr (lanes 6-9) after induced by 1.0 mM IPTG.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

In contrast, lower expression levels of recombinant Dmc1 was found at 1 hr after IPTG induction. A greater level of expressed Dmc1 was observed when the culture period was extended between 2-12 hr after induction. The overnight culture of recombinant clones yielded the highest amount of recombinant Dmc1.

The expected protein band and that exhibiting a lower molecular weight were observed when total proteins from the recombinant clone of SUMO-1 were hybridized with Anti-His tag. Lower expression levels of recombinant SUMO-1 was found at 1 hr after IPTG induction. A greater level of expressed this recombinant protein was observed when the culture period was extended between 2 hr to overnight after IPTG induction. Nevertheless, the intensity of the interfering band was approximately identical to that of the expected target band throughout all culture period. The protein sequence of the expected SUMO-1 protein was further analyzed by LC-MS/MS.

Likewise, lower expression levels of recombinant cyclophilin A was found at 1 hr after IPTG induction. After 2 hr of induction, a greater level of expressed cyclophilin A was observed. This recombinant protein was stably expressed throughout the cultured period.

In addition, aliquots of the IPTG induced-culture (OD = 1) were collected. The soluble and insoluble protein fractions of each gene were analyzed by SDS-PAGE and western blot analysis. Recombinant spermatogonial stem cell renewal

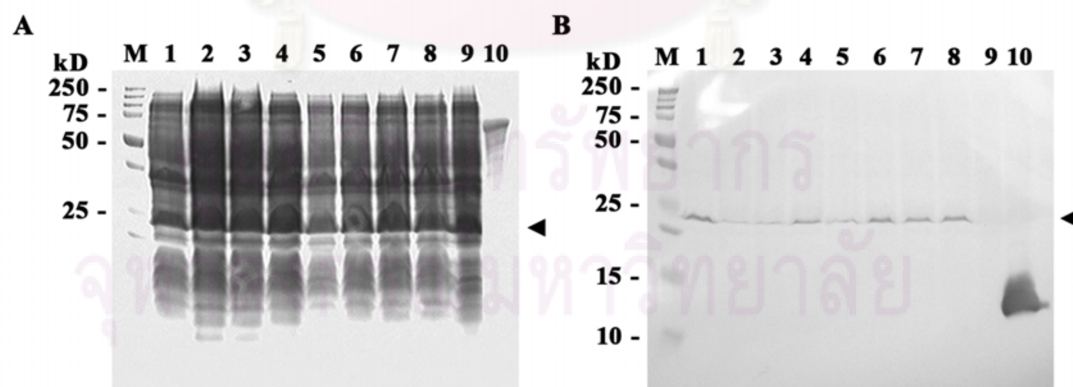


Figure 3.108 *In vitro* expression of eight recombinant clones of *PGRMC1* of *P. monodon* after induced by 1 mM IPTG overnight (lanes 1-8) using SDS-PAGE (A) and western blot analysis (B). *E. coli* BL21C+(DE3)RIPL cells (lane 8), and pET32a vector in *E. coli* BL21C+(DE3)RIPL cell (lane 9) were included as the control.

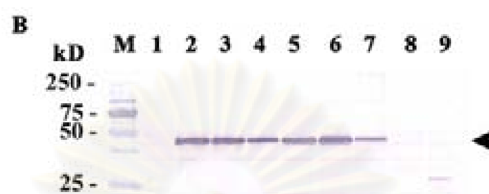
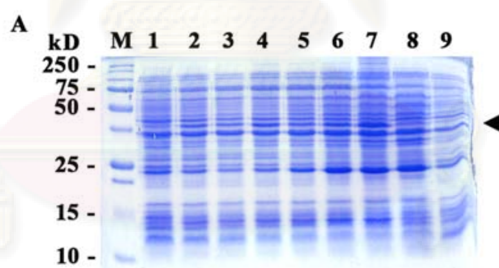
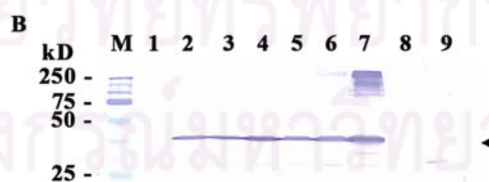


Figure 3.109 *In vitro* expression of recombinant spermatogonial stem cell renewal factor of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG (lanes 1-7) using SDS-PAGE (A) and western blot analysis (B). *E. coli* BL21C+(DE3)RIPL cells (lane 8), and pET32a vector in *E. coli* BL21C+(DE3)RIPL cells (lane 9) were included as the control.



ศูนย์วิทยทรัพยากร



จุฬาลงกรณ์มหาวิทยาลัย

Figure 3.110 *In vitro* expression of recombinant Dmc1 of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG (lanes 1-7) using SDS-PAGE (A) and western blot analysis (B). *E. coli* BL21C+(DE3)RIPL cells (lane 8), and pET32a vector in *E. coli* BL21C+(DE3)RIPL cells (lane 9) were included as the control.

factor and Dmc1 cultured at 37°C were expressed as the insoluble protein (Fig. 3.113 and 3.115), whereas recombinant cyclophilin A and SUMO-1 cultured at 37°C were found in both soluble and insoluble fraction at 6 hr after IPTG induction (Fig. 3.117 and 3.119).

The cultured temperature for recombinant clones of spermatogonial stem cell renewal factor, Dmc1, and cyclophilin A was then decreased from 37°C for 3 or 6 hr to 15°C overnight after the culture was induced by IPTG. Results showed that larger amounts of recombinant spermatogonial stem-cell renewal factor and cyclophilin A were expressed in the soluble fraction (Fig 3.114 and 3.118) whereas recombinant Dmc1 still was expressed in the insoluble fraction (Fig. 3.116).

Therefore the cultured temperature for further purification of recombinant spermatogonial stem-cell renewal factor and cyclophilin A were carried out at 15°C, overnight whereas recombinant Dmc1 was cultured at 37°C for 6 hr.

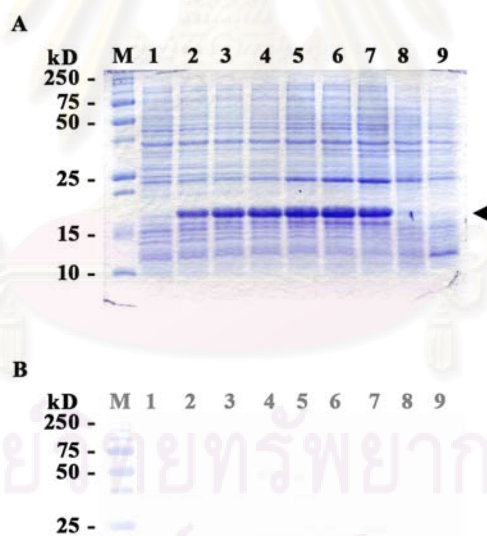


Figure 3.111 *In vitro* expression of recombinant SUMO-1 of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG (lanes 1-7) using SDS-PAGE (A) and western blot analysis (B). *E. coli* BL21C+(DE3)RIPL cells (lane 8), and pET32a vector in *E. coli* BL21C+(DE3)RIPL cells (lane 9) were included as the control.

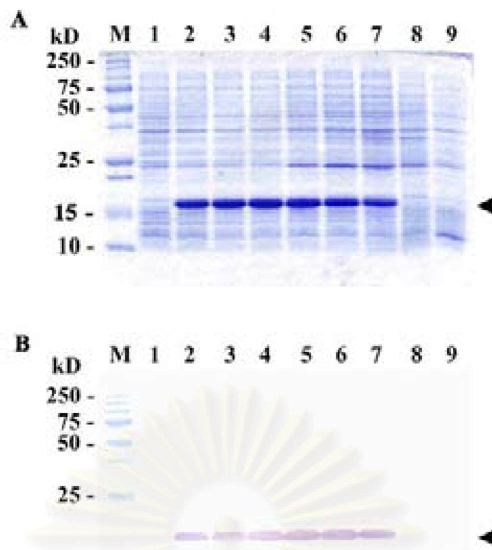


Figure 3.112 *In vitro* expression of recombinant cyclophilin A of *P. monodon* at 0, 1, 2, 3, 6, 12, and 24 hr after induced by 1 mM IPTG (lanes 1-7) using SDS-PAGE (A) and western blot analysis (B). *E. coli* BL21C+(DE3)RIPL cells (lane 8), and pET32a vector in *E. coli* BL21C+(DE3)RIPL cells (lane 9) were included as the control.

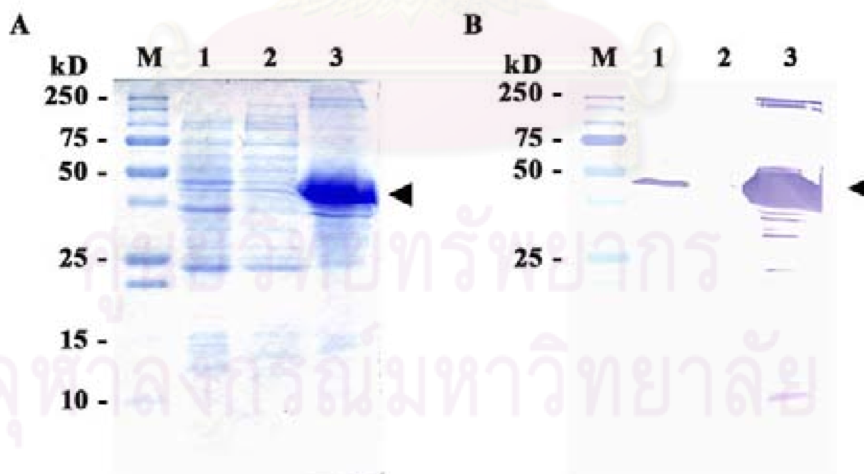


Figure 3.113 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant spermatogonial stem cell renewal factor after IPTG induction for 6 hr at 37°C as the insoluble protein. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 µg protein), and Lane 3: an insoluble protein fraction (25 µg protein).

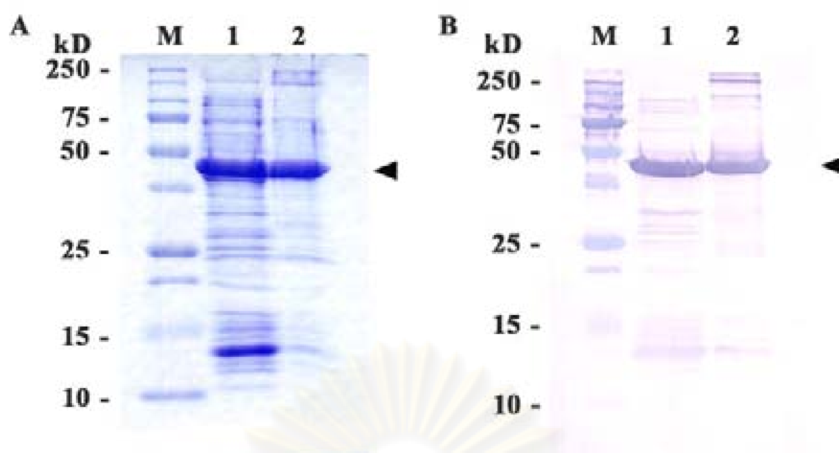


Figure 3.114 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant spermatogonial stem cell renewal factor after IPTG induction overnight at 15°C as both soluble and insoluble proteins. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 μg protein), and Lane 3: an insoluble protein fraction (25 μg protein).

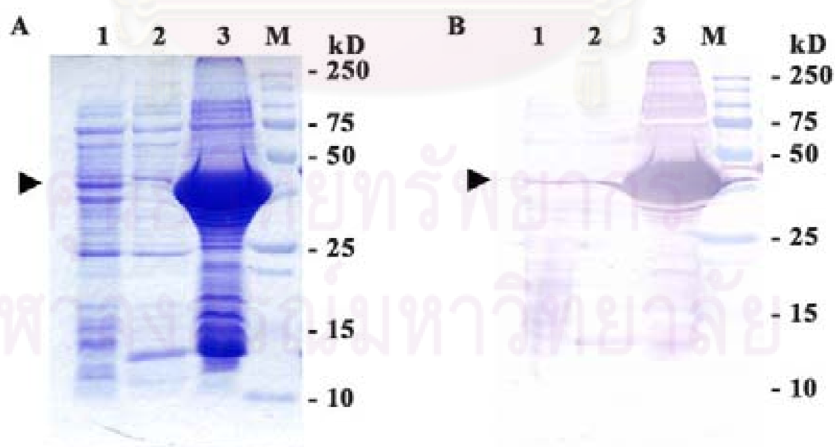


Figure 3.115 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant Dmc1 after IPTG induction for 6 hr at 37°C as the insoluble protein. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 μg protein), and Lane 3: an insoluble protein fraction (25 μg protein).

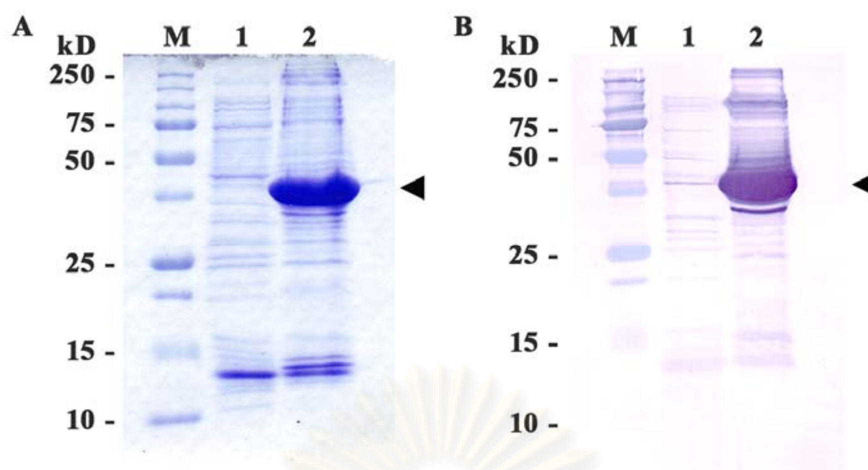
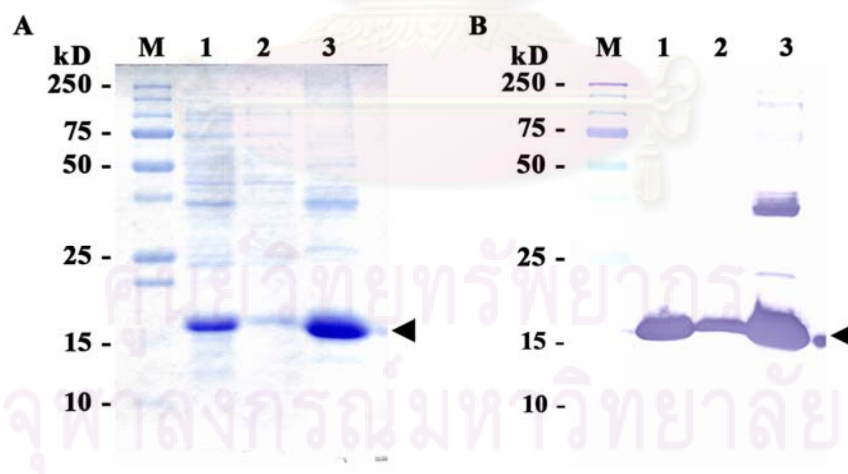


Figure 3.116 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant Dmc1 after IPTG overnight at 15°C as the insoluble protein. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 μg protein), and Lane 3: an insoluble protein fraction (25 μg protein).



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Figure 3.117 SDS-PAGE (A) and western blot analysis (B) showing more expression of recombinant cyclophilin A after IPTG induction for 3 hr at 37°C as the insoluble protein than the soluble protein. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 μg protein), and Lane 3: an insoluble protein fraction (25 μg protein).

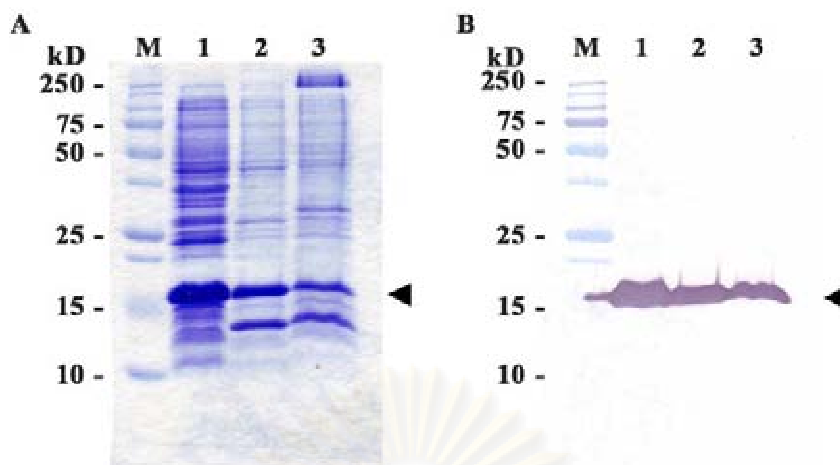


Figure 3.118 SDS-PAGE (A) and western blot analysis (B) showing expression of recombinant cyclophilin A after IPTG induction overnight at 15°C as both soluble and insoluble proteins. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 μg protein), and Lane 3: an insoluble protein fraction (25 μg protein).

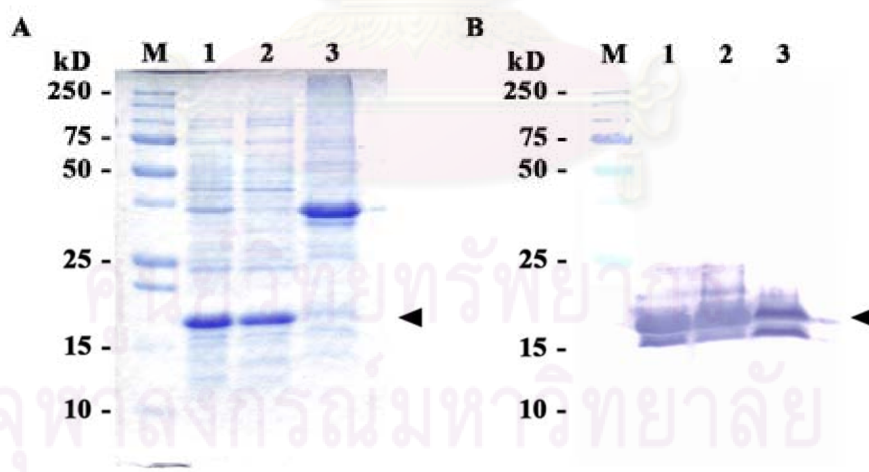


Figure 3.119 SDS-PAGE (A) and western blot analysis (B) showing equal expression of recombinant SUMO-1 after IPTG induction for 6 hr at 37°C as the soluble and insoluble proteins. Lane 1 = whole cells, Lane 2 = a soluble protein fraction (25 μg protein), and Lane 3: an insoluble protein fraction (25 μg protein).

3.9.3 Peptide sequencing of recombinant cyclophilin A and SUMO-1

Surprisingly, western blot analysis indicated 2 positively discrete protein bands of recombinant SUMO-1. Accordingly, The expected SUMO (upper band) and cyclophilin A were further analyzed by ESI-LC-MS/MS. Internal peptide sequence of recombinant SUMO-1 was K-I-K-V-V-G-Q-D-S-N-E-I-H-F-R-V which was significantly matched SUMO-1 like protein of *Artemia salina* (E value = 6.7E-09) and perfectly matched the deduced amino acid sequence obtained from EST. Likewise, that of cyclophilin A possessed R-I-V-M-E-L-R-A-D-V-V-P-K-T-A-E-N-F-R-**A** and K-H-T-G-P-G-T-L-S-M-A-N-A-G-P-N-T-N-G-S-Q-F-F-**L**-C-T-V-K-T which was significantly matched cyclophilin A of *Scophthalmus maximus* (1.8E-06) and nearly identical to R-I-V-M-E-L-R-A-D-V-V-P-K-T-A-E-N-F-R-**S** and K-H-T-G-P-G-T-L-S-M-A-N-A-G-P-N-T-N-G-S-Q-F-F-**I**-C-T-V-K-T deduced from the full length cDNA of *P. monodon* cyclophilin A. This confirmed that recombinant proteins obtained were actually SUMO-1 and cyclophilin A, respectively.

3.9.4 Purification of recombinant protein

3.9.4.1 The first trial

The soluble fractions of recombinant spermatogonial stem-cell renewal factor, SUMO-1 and cyclophilin A were purified as native proteins whereas the insoluble fraction of recombinant Dmc1 was purified under denaturing conditions.

Recombinant proteins were run through the column three times and washed twice. The first wash used 10 ml of the binding buffer (20 mM sodium phosphate and 500 mM NaCl) including 20 mM imidazole, pH 7.4 and the second was used 5-10 ml of the binding buffer including 80 mM imidazole, pH 7.4). After that recombinant proteins were eluted from column with 6 ml of the elution buffer (20 mM sodium phosphate, 500 mM NaCl and 500 mM imidazole, pH 7.4). Washed and eluted fractions were analyzed by SDS-PAGE and western blot (Fig 3.120-3.123). After purification, eluted proteins were keeping at -20 °C overnight.

When purified protein solutions of these recombinant proteins were thawed, white precipitated protein pellets were clearly seen in the eluted fraction 1-3. All

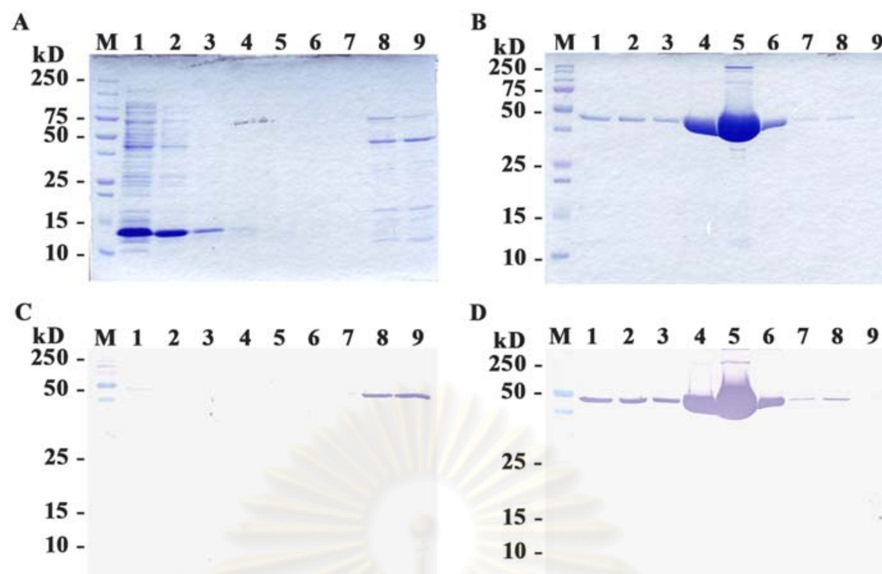


Figure 3.120 Purification of recombinant spermatogonial stem cell renewal factor of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-6 = the first wash fractions 1, 2, 3, 6, and 9 and lanes 7-9 = the second wash fractions 1, 2, and 3, respectively. B and D: lanes 1-3 = the second wash fractions 5, 7, and 9 and lanes 4-9 = eluted protein fractions 1-6, respectively.

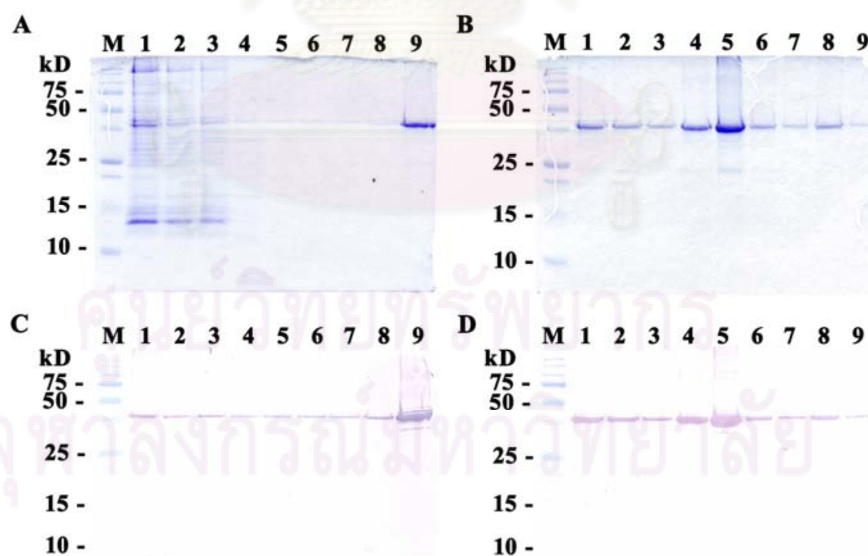


Figure 3.121 Purification of recombinant Dmc1 of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-7 = the first wash fractions 1, 2, 3, 5, 7 and 9 and lanes 8-9 = the second wash fractions 1 and 2, respectively. B and D: lanes 1-3 = the second wash fractions 3, 4, and 5 and lanes 4-9 = eluted protein fractions 1-6, respectively.

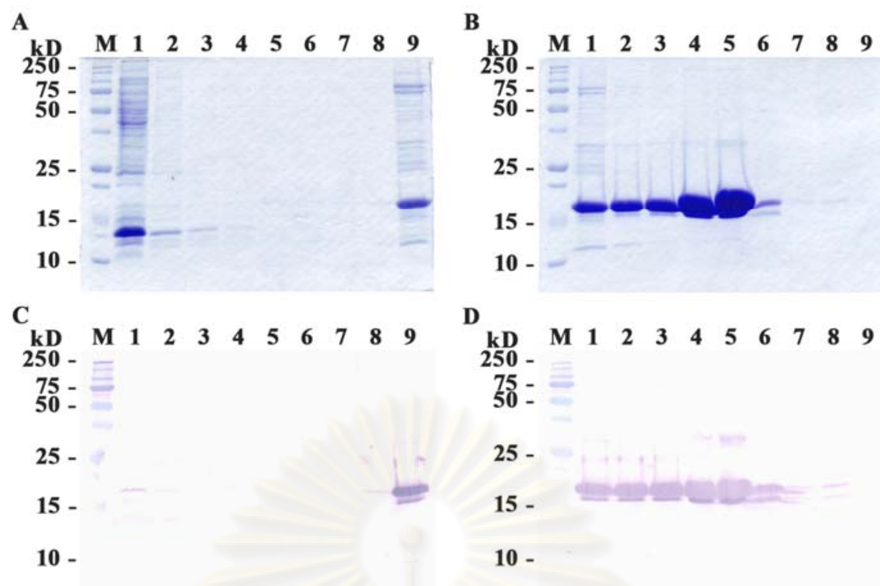
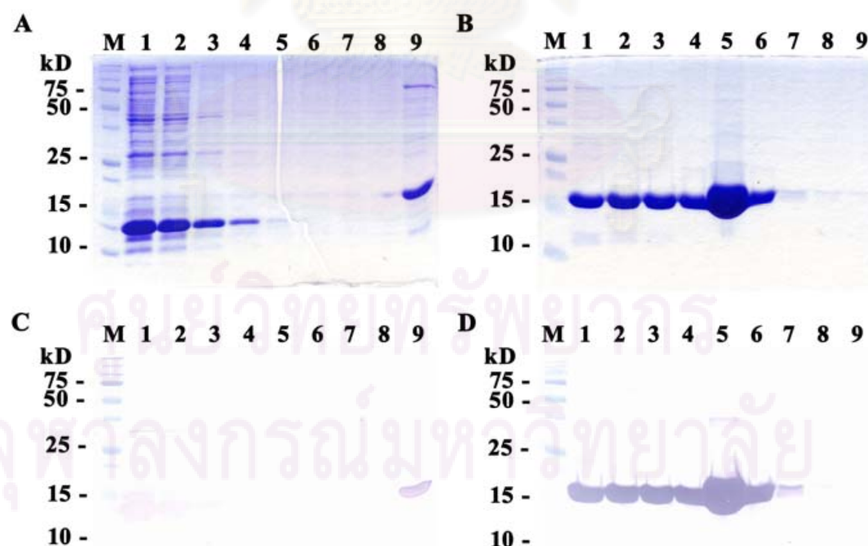


Figure 3.122 Purification of recombinant SUMO-1 of *P. monodon* (cultured at 37°C, for 6 hr). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-7 = the first wash fractions 1, 2, 3, 5, 7 and 9 and lanes 8-9 = the second wash fractions 1 and 2, respectively. B and D: lanes 1-3 = the second wash fractions 3, 5, and 7 and lanes 4-9 = eluted protein fractions 1-6, respectively.



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Figure 3.123 Purification of recombinant cyclophilin of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-7 = the first wash fractions 1, 2, 3, 5, 7 and 9 and lanes 8-9 = the second wash fractions 1 and 2, respectively. B and D: lanes 1-3 = the second wash fractions 3, 4, and 5 and lanes 4-9 = eluted protein fractions 1-6, respectively.

eluted fractions of each recombinant protein were then pooled and dialyzed with the PBS buffer to eliminate imidazole in the elution buffer. The dialyzed proteins were concentrated *in vacuo*. Non-purified and purified recombinant proteins were sent to Faculty of Medical Technology for production of polyclonal antibodies. Notably, the purified proteins from the first trial were not suitable for production of polyclonal antibody.

9.4.1.2 The second trial

The first trial for purification of recombinant proteins was unsuccessful. Accordingly, the second trial was performed. Recombinant SUMO-1 was excluded from the experiment because the upper target protein bands co-purified with the non-recombinant protein of *E. coli* after purification.

Recombinant spermatogonial stem cell renewal factor cyclophilin A and Dmc1 proteins were purified using the same methods used for the first trial excepted an additional washing step was included (5 ml of the binding buffer containing 50 mM imidazole, pH 7.4) (Fig 3.124-3.126). After purification, eluted recombinant proteins were kept at 4°C overnight. The precipitated proteins (white pellet) were slightly found in the eluted protein fraction 3. All except this fraction of recombinant spermatogonial stem-cell renewal factor and cyclophilin A were pooled and dialyzed with the PBS buffer. White precipitated proteins were still found in recombinant cyclophilin A. Purified proteins were concentrated by ultrafiltration.

All eluted fractions of recombinant Dmc1 was pooled, concentrated with ultrafiltration and size-fractionated by SDS-PAGE. The protein band was eluted with the protein elution buffer (25 mM Tris, 192 mM Glycine, and 0.1% SDS) using an Electro-Eluter (model 422, Bio-Rad). Non-purified and purified recombinant proteins were sent to Faculty of Medical Technology for further immunization of rabbit and determination of the titers of produced antibodies.

For production of the polyclonal antibody, degraded spermatogonial stem-cell renewal factor, cyclophilin A and Dmc1 from the first trial were boosted four times in separate rabbits. Serum of the immunized rabbit was collected and the titer was estimated by ELISA using the degraded protein as the antigen. However, ELISA

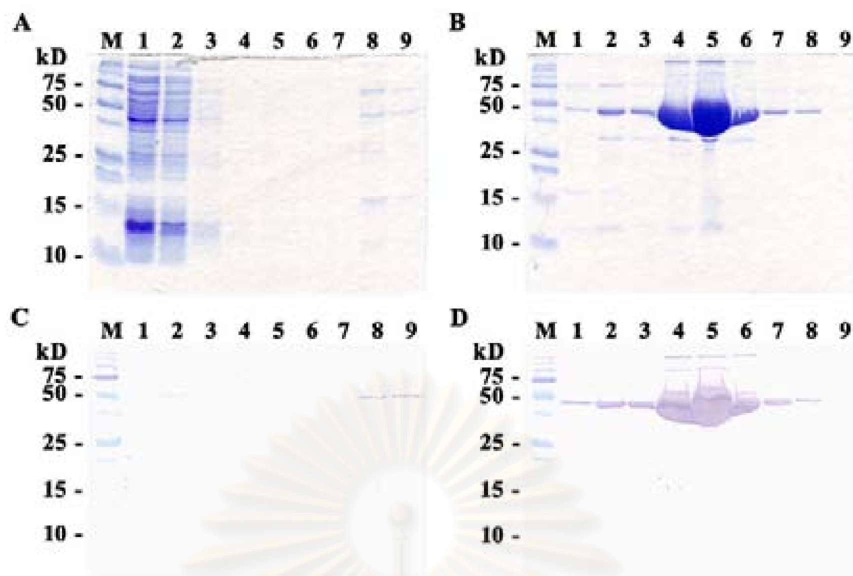


Figure 3.124 Purification of recombinant spermatogonial stem cell renewal factor of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-6 = the first wash fractions 1, 3, 5, 7 and 9 and lanes 7-9 = the second wash fractions 1, 3 and 5, respectively. B and D: lanes 1-3 = the third wash fractions 1, 3, and 5 and lanes 4-9 = eluted protein fractions 1-6, respectively.

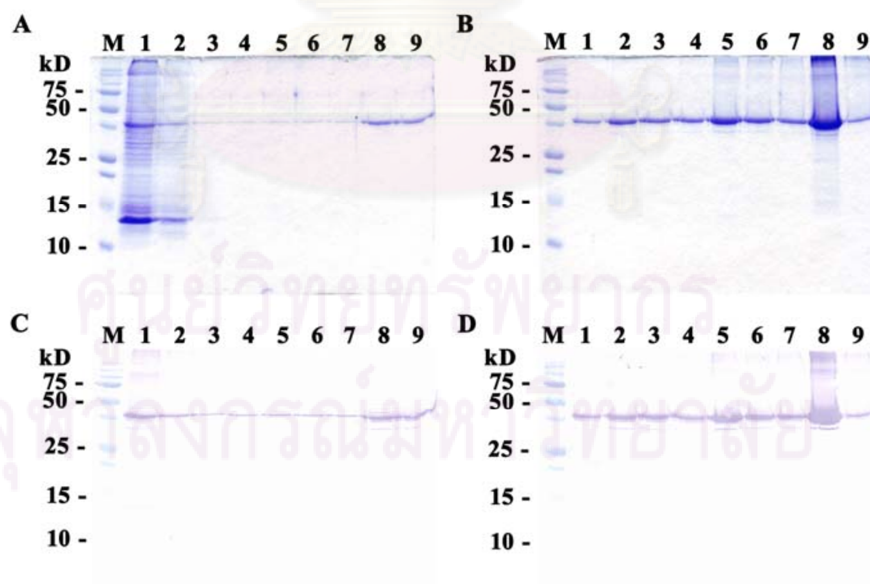


Figure 3.125 Purification of recombinant Dmc1 of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-6 = the first wash fractions 1, 3, 5, 7 and 9 and lanes 7-9 = the second wash fractions 1, 3 and 5, respectively. B and D: lanes 1-3 = the third wash fractions 1, 3, and 5 and lanes 4-9 = eluted protein fractions 1-6, respectively.

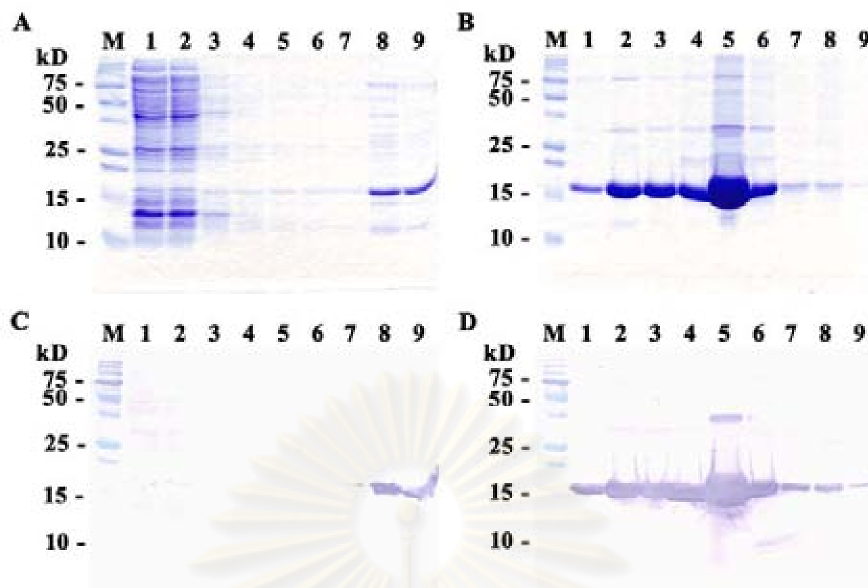


Figure 3.126 Purification of recombinant cyclophilin A of *P. monodon* (cultured at 15°C, overnight). Recombinant proteins were examined by using SDS-PAGE (A-B) and western blot analysis (C-D).

A and C: lane 1 = the soluble fraction after pass through the column, lanes 2-6 = the first wash fractions 1, 3, 5, 7 and 9 and lanes 7-9 = the second wash fractions 1, 3 and 5, respectively. B and D: lanes 1-3 = the third wash fractions 1, 3, and 5 and lanes 4-9 = eluted protein fractions 1-6, respectively.

did not give the positive results possibly owing to problems from precipitated proteins (antigen). Afterwards, the purified recombinant protein solutions from the second trial were used to boost the same rabbit. Titers of rabbit serum after the 4th and 5th immunization were examined using antigen from recombinant protein solutions (2nd trial).

Results from ELISA showed that polyclonal antibodies of all recombinant proteins were successfully produced. The positive titers were observed since the 4th immunization and the titers of polyclonal antibodies were greater after the 5th immunization. Nevertheless, the titer of polyclonal antibody for recombinant cyclophilin A could not be detected. This should be resulted from degradation of the antigen (purified cyclophilin A) (Table 3.20-3.21). Although polyclonal antibody of recombinant spermatogonial stem-cell renewal factor, cyclophilin A, and Dmc1 were successfully produced, the titers of these polyclonal antibodies were quite low.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.20 Summary of conditions used for purification of recombinant proteins in this study

	Recombinant protein			
	SUMO*	**Spermatogonial stem cell renewal factor****	**Dmc1****	**Cyclophilin A****
Size of protein (kDa)	10.58	47.21	37.54	18.86
Kind of protein	Soluble protein	Soluble protein	Insoluble protein	Soluble protein
Culture cells	1 Liter (Purify 25%)	500 ml (Purify 80%)	500 ml (Purify 50%)	500 ml (Purify 80%)
Conditions (temp and time after induction)	37 °C, 6 hr	15 °C, overnight	37 °C, 6 hr	15 °C, overnight
Weight of wet cells (g/L)	5.03	0.95	2.74	0.97
Weight of inclusion body fraction (g/L)	0.58	0.27	0.48	0.18
Protein concentration (mg protein/L of cultured cells)	51.86	89.01	55.05	129.83
Protein concentration of each elution (E1-E6)	E1 = 4.728 mg/ml, E2 = 3.945 mg/ml, E3 = 1.798 mg/ml, E4 = 0.487 mg/ml, E5 = 1.314 mg/ml, and E6 = 0.694 mg/ml	E1 = 9.78 mg/ml, E2 = 22.62 mg/ml, E3 = 3.88 mg/ml, E4 = 0.45 mg/ml, E5 = 0.28 mg/ml, and E6 = 0.08 mg/ml	E1 = 1.03 mg/ml, E2 = 2.05 mg/ml, E3 = 1.61 mg/ml, E4 = 1.33 mg/ml, E5 = 6.26 mg/ml, and E6 = 1.48 mg/ml	E1 = 5.56 mg/ml, E2 = 42.54 mg/ml, E3 = 5.06 mg/ml, E4 = 0.39 mg/ml, E5 = 0.41 mg/ml, and E6 = 0.13 mg/ml

*First purification trial

**Second purification trial

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.21 ELISA results illustrating the titers of polyclonal antibodies after rabbits were boosted by recombinant spermatogonial stem cell renewal factor, Dmc1 and cyclophilin A after rabbits were boosted 4 times

	Polyclonal antibody					
	Spermatogonial stem-cell renewal factor		Dmc1		Cyclophilin A	
Dilution of serum	**Pre-immunized serum**	**Immunized serum**	**Pre-immunized serum**	**Immunized serum**	**Pre-immunized serum**	**Immunized serum**
Antigen dilution: before the second purification						
1:500	0.785	1.222	0.922	2.247	1.150	2.868
1:2000	0.221	0.524	0.364	1.281	0.449	1.776
1:8000	0.049	0.162	0.072	0.475	0.116	0.686
1:32000	0.002	0.032	0.005	0.132	0.029	0.181
Antigen dilution: after the second purification						
1:500	0.273	0.688	0.018	0.387	0.058	0.795
1:2000	0.104	0.271	0.004	0.119	0.004	0.242
1:8000	0.027	0.079	0.001	0.040	-0.007	0.080
1:32000	0.008	0.023	-0.004	0.010	-0.013	0.018

Pre-immunized serum = serum from normal rabbit

Immunized serum = serum from rabbit injected with the recombinant protein 4 times

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table 3.22 ELISA results illustrating the titers of polyclonal antibodies after rabbits were boosted by recombinant spermatogonial stem cell renewal factor, Dmc1 and cyclophilin A after rabbits were boosted 5 times

	Polyclonal antibody					
	Spermatogonial stem-cell renewal factor		Dmc1		Cyclophilin A	
Dilution of serum	**Pre-immunized serum**	**Immunized serum**	**Pre-immunized serum**	**Immunized serum**	**Pre-immunized serum**	**Immunized serum**
Antigen dilution: before the second trial purification						
1:500	0.352	1.486	0.534	1.681	0.481	1.503
1:2000	0.072	1.057	0.214	1.009	0.194	0.869
1:8000	0.041	0.663	0.094	0.481	0.092	0.392
1:32000	0.042	0.249	0.057	0.181	0.062	0.149
Antigen dilution: after the second trial purification						
1:500	0.217	1.873	0.075	0.936	0.072	0.151
1:2000	0.074	1.394	0.064	0.438	0.061	0.084
1:8000	0.059	0.866	0.061	0.194	0.054	0.066
1:32000	0.061	0.363	0.057	0.095	0.051	0.063

Pre-immunized serum = serum from normal rabbit

Immunized serum = serum from rabbit injected with the recombinant protein 5 times

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER IV

DISCUSSION

Spermatogenesis consists of a series of complex cellular events, in which different genes express to ensure the proper development of spermatozoa (Abe, 1987 and 2004).). An initial step toward understanding molecular mechanisms of testicular and sperm development as well as sex differentiation cascades in *P. monodon* is the identification and characterization of sex-related genes expressed in testes of an economically important species like *P. monodon*. The basic information would provide an insight into molecular aspects governing reproductive processes for future functional studies in this species. In addition, the fundamental controls of growth in penaeid shrimp are largely unstudied. Genes encoding vertebrate-like growth factors and cell cycle regulating proteins should also be characterized. Integration of the knowledge from various molecular disciplines would provide markers that can be used to assist selective breeding programs of *P. monodon*.

Isolation and characterization of functionally important gene homologues in the conventional heart cDNA library of *P. monodon*

To characterize genes encoding vertebrate-like growth factors and cell cycle regulating proteins, a conventional heart cDNA library of *P. monodon* was established from mRNA of heart (4-month-old) of *P. monodon*. A total of 412 recombinant clones of the heart cDNA library having inserts greater than 500 bp in size were unidirectionally sequenced from the 5′ terminus. Sequences were blasted against data in the GenBank using Blast*N* and Blast*X* and 54.5% of the heart EST significantly matched with known sequences in the GenBank ($P < 10^{-4}$) whereas the remaining sequences were newly genes identified regarded as unknown transcripts (45.4 %).

A large number of unknown transcripts were found from the library. Among known transcripts in the heart library, those categorized to unknown transcripts (45.4%) were the most abundantly expressed group followed by those classified into a mitochondrial protein group (26.5%), an unidentified (hypothetical cDNA/DNA) group (17.0%). The remaining transcripts including those classified as members of miscellaneous functions (2.6%), ribosomal proteins (2.2%), gene expression

regulation and protein synthesis (2.2%), metabolism (1.9%), defense and homeostasis (1.2%), signal and communication (0.5%), cell division and DNA synthesis (0.5%), internal/external structure and mobility (0.25%) were rare abundant transcripts in this library.

It is not surprised to identify a large number of mitochondrial protein genes in the heart cDNA library because the high proportion of mitochondrial DNA/genomic DNA is usually found in active tissues (Imjongjairak et al., 2004). As a result, abundantly expression levels of mitochondrial protein transcripts should be typical for the heart tissue. This reflects a large proportion of EST representing mitochoindrial proteins in this library.

Several functional important transcripts including *troponin-T*, *high mobility protein 20A*, *neuroparin A*, *lipoamide dehydrogenase*) were identified. These transcripts which have previously reported to be related with growth of other organisms will be further characterized. In addition, the full length cDNA of a homologue of *myosin 1 light chain* (*MLC1*, ORF of 465 bp encoding a protein of 154 amino acid) which is significantly matched that of *Apis mellifera* (E-value = 7e-49) was also identified in this libraries.

Neurotransmitters and neuropeptides (e.g *MCL1* and *neuroparsin*) that are found in heart are important for cardiac regulation, brain development, growth and reproduction in several species. The roles of neuropeptides for growth and reproductive regulation have been reported in gastropods (i.e. abalone) for examples; mutation of the MCL1 protein causes brain disorders (Schmitt, et al, 2003). Neuroparsin A and B involved with brain development and reproductive physiology in gonads (Janssen, et al, 2001).

Transcripts involved with metabolisms were also found in the conventional heart cDNA library. Therefore, this library should be normalized for decreasing mitochondrial transcripts before ESTs are further sequenced of a larger number of recombinant clones.

Isolation and characterization of functionally important gene homologues in the conventional testis cDNA library of *P. monodon*

Due to difficulties in sexual maturation of captive *P. monodon*, molecular mechanisms of this process have long been of interest by aquaculture industries (Preechaphol *et al.*, 2007). Eight hundred and ninety-six recombinant clones from a testis cDNA library (the primary titer of 5.66 x 10^6 pfu/ml) were sequenced and analyzed. A total of 699 (67.4%) of sequenced clones significant match previously deposited genes in the GenBank whereas 290 sequences (32.6%) were regarded as unknown transcripts. Accordingly, unknown transcripts and functionally unidentified (hypothetical) protein-coding genes (114, 12.8%) predominated in this library.

The percentage of unknown transcripts found in the present study was lower than that previously reported in the normal (39.1%) and SSH (65.0%) cDNA libraries of testes of *H. asinina* (Amparyup et al., 2004) and a testis SSH library of *P. monodon* (96.7%, Leelatanawit et al., 2004). A large number of unknown transcripts found in testes of *P. monodon* indicated that further characterization (e.g. by RACE-PCR) is required to cover a larger part of their coding region.

The relationship between the number of clones sequenced and the accumulative numbers of unique transcripts indicated that the discovery rate of new transcripts still does not reach saturation after 896 recombinant clones were sequenced. Therefore, additional unique transcripts can still be identified by sequencing a larger number of recombinant clones.

In contrast to redundant transcripts identified in ovaries of *P. monodon* (7.5% of *TSP* and 8.3% of *peritrophin*; Preechaphol et al., 2007) and testes of *H. asinina* (7.6% of *sperm lysin*; Amparyup et al., 2004), highly redundant ESTs were not observed in the present library. This further suggested that the established library was reasonably diverse.

The full length cDNA of several functionally important genes (e.g. *TCTP*, *dynactin subunit 5*, *small nuclear ribonucleoprotein polypeptide G*, *ubiquitin conjugating enzyme E2* and *mitotic checkpoint BUB3*) were identified implying relative large insert sizes of this cDNA library. All except *TCTP* (AY186580),

ubiquitin conjugating enzyme E2 (DV738178 and CO777380) and *cyclophilin A* (EU164775 and DT624276) were reported for the first time in *P. monodon*.

Dynein is functionally related to the transport of various cytoplasmic organelles (Aniento et al., 1993). In *Drosophila*, cytoplasmic dynein and dynactin complex is required for the spermatid growth but not for axoneme assembly (Ghosh-Roy et al., 2004).

In addition, the full length cDNA of *cdc2*, *SUMO-1* and *cyclophilin A* were also found. During the meiotic development examined so far, the G2/M phase transition is controlled by the maturation promoting factor (MPF), a complex of cdc2 (cdk1) and cyclin B1. In rat, inhibition of cdk1 and cdk2 affected the spontaneous processing of the first and second meiotic division of male germ cells (Godet et al., 2004).

SUMO-1 plays an important role in diverse reproductive functions such as spermatogenesis and modulation of steroid receptor activity. A SUMO-binding motif has been identified in the androgen receptor (AR), progesterone receptor (PR) and glucocorticoid receptor (GR) suggesting distinct roles for sumoylation in steroid receptor activity for growth and reproduction (Koshiyama, et al., 2006). In the sumoylation pathway, SUMO is transferred to substrate lysine residues through the thioester cascade of ubiquitin activating enzyme E1 and ubiquitin conjugating enzyme E2 (UBE2), and SUMO ligase E3 functions as an adaptor between E2 and each substrate (Takahashi. and Kikuchi, 2005).

In *Marsupenaeus japonicus*, *UBE2* was expressed at a higher level in testes than in ovaries. The expression at the stage I (GSI = 0.33 ± 0.004, N = 5) was significantly lower than that of the stage II (GSI = 0.45 ± 0.12, N = 5) but comparable to that of the stage III (GSI = 0.57 ± 0.006, N = 5) of testes. *UBE2* in ovaries was up-regulated since the stage III of ovaries. This suggested that *UBE2* has an important role in spermatogenesis and oogenesis of *M. japonicus* (Shen et al., 2008).

Cyclophilins are small proteins that bind Cyclosporin A (CsA) and catalyze protein folding (Lang et al., 1987). Cyclophilins are characterized by a conserved CBD that is required for both CsA-binding and protein-folding activities (Page et al., 1996). Recently, a diverse cyclophilin, mog 1 was isolated and functionally

characterized. Binding of mog 1 to MEP-1 is essential for germline sex determination in *Caenorhabditis elegans* (Belfiore *et al*., 2004).

Isolation and characterization of functionally important gene homologues in the SSH libraries of testes of *P. monodon*

A total of 367 ESTs of testes (178 and 189 clones from the forward and reverse SSH libraries, respectively) of *P. monodon* were unidirectionally sequenced. The positive recombinant clones of both libraries was 95.1 and 96.4% where 82.4 and 86.2 of which having insert sizes > 250 bp in length. The percentage of ESTs significantly matched known genes in respective libraries was 37.1 (66 ESTs) and 53.5% (100 ESTs).

Peritrophin and *thrombospondin* which are abundantly expressed in ovaries of *P. monodon* were not found in the subtractive testis cDNA libraries. Relatively large numbers of genes encoding ribosomal proteins were found in both libraries (16 and 11 accounting for 8.9 and 5.6%, respectively).

Unlike a SSH library of *H. asinina* where ESTs representing sperm lysin (7.6%) were abundantly expressed, highly redundant ESTs were not observed and a relatively large number of known gene homologues were found in these libraries. This suggested that the established SSH libraries were reasonably diverse.

Seven known transcripts: *allergen Pen m 2*, *COI*, *EF-1α*, *GTP-binding protein, 26S proteasome non-ATPase subunit 12, receptor for activated protein kinase C* (*RACK*) and *myelodysplasia/myeloid leukemia factor* were found with low frequencies in both libraries suggesting that the cDNA subtraction was successful.

Unknown transcripts predominated in both libraries (112 and 87 clones accounting for 62.9 and 46.5%, E-value > 1e-04). The percentage of unknown transcripts in SSH libraries was much grater than that of the conventional library established from testis of *P. monodon* broodstock (290/889 clones, 32.6%) but lower than those found in the SSH library of testes of the tropical abalone, *Haliotis asinina* (125/160, 65%).

Full length cDNAs identified by EST analysis and RACE-PCR

Eleven full length cDNAs of functionally important genes were discovered from EST analysis of testis (10 transcripts) and heart (1 transcript) cDNA libraries. They were *SUMO-1*, *cyclophilin A*, *TCTP, small nuclear ribonucleoprotein polypeptide G*, *dynactin subunit 5*, *ubiquitin conjugating enzyme 2*, *mitotic checkpoint, Cdc2*, *thioredoxin* , *multiprotein bridging factor 1* and *MLC1*.

In addition, the full length cDNA of 16 functionally important genes for example, *transformer-2*, *meiotic recombination protein DMC1/LIM15 homolog isoform 1*, *progestin receptor membrane component 1*, *prohibitin-2 (a repressor of estrogen receptor activity)*, *multiple inositol polyphosphate phosphatase 2*, *prohibitin-2 (a repressor of estrogen receptor activity)*, and *growth factor receptor-bound protein* were successfully characterized by RACE-PCR

Progestins are sex steroid hormones that play important roles in gametogenesis. In fish, progestin also plays an important role in spermiation and sperm maturation (Miura et al., 2006). Recently, effects of 17α, 20β-dihydroxy-4-pregnen-3-one (DHP) on the initiation of meiosis in the male Japanese eel were examined. DHP induced DNA replication in spermatogonia but prevented DNA replication during the period in which meiosis was initiated (Miura et al., 2006).

Two totally distinct classes of putative membrane-bound progestin receptors have been reported in vertebrates. They are progestin membrane receptor component (PGMRC; subtypes 1 and 2) and membrane progestin receptors (mPR; subtypes α, β, γ) which have never studied in any crustacean (Mourot et al., 2006). The full length cDNA of *PGMRC1*, initially identified in the reverse SSH libraries, was successfully identified by RACE-PCR and reported for the first time in crustaceans.

Sex determination mechanisms have long been of major interest from both developmental and evolutionary points of view (Delvin and Nagahama, 2002). In *Drosophila*, sex determination is under the control of the sex lethal (*Sxl*) gene (Burtis et al., 1991). Sxl participates in the female-specific splicing of its own pre-mRNA. The downstream target of Sxl is the *transformer* (*Tra*) which encodes a non-functional truncated Tra protein in males (Inoue et al., 1990). The female Tra protein induces female-specific splicing of the *doublesex* (*dsx*) pre-mRNA in cooperation

with the Tra-2 gene product promoting female sexual development (Burtis et al., 1991; Jursnich and Burtis, 1993; An and Wensink, 1995).

The diploid chromosome numbers of penaeid shrimp are 2*N* = 88- 92 where *P. monodon* possesses 2*N* = 88 (Benzie, 1998). Neither sex chromosomes nor environmental sex determination has been reported in penaeid shrimp. The partial sequence of *Sxl* transcripts were previously found in hemocytes EST libraries of *P. monodon* (Tassanakajon et al., 2006) and *Marsupenaeus japonicus* (T. Aoki, personal communication). In this study, RACE-PCR was applied and successfully identified the full length cDNA of sex-related *Tra-2* in *P. monodon*. The deduced *PMTra-2* contained the RRM domain which is found in a variety of RNA binding proteins, including heterogeneous nuclear ribonucleoproteins (hnRNPs), proteins implicated in regulation of alternative splicing, and protein components of small nuclear ribonucleoproteins (snRNPs) (Liu et al., 2008). The discovery of *PMTra-2* indirectly suggested that sex determination of this species may have influenced by the *Sxl-dsx-Tra* pathway.

Sex-specific/differential expression markers in *P. monodon*

Transcripts in this study were identified from cDNAs expressed in the conventional and SSH cDNA libraries. Thirty-six of 59 examined gene homologues (e.g. *SUMO-1, CYA, dynactin subunit 5 and Trap240*) revealed a trend of preferential expression in ovaries to testes of *P. monodon*. *PMTST1* was restrictively expressed in testes but not ovaries whereas *HSP70-2* exhibited a trend of preferential expression in testes of *P. monodon*. As a result, a SSH library between cDNA from testes and ovaries should be established to obtain additional transcripts that exhibit preferential expression levels in testes of *P. monodon*.

Inx2 and *PGMRC1* did not revealed differential expression between ovaries and testes of *P. monodon* ($P > 0.05$). In contrast, *Dmc1* was expressed differentially in gonads (testes > ovaries) of *P. monodon* broodstock ($P < 0.05$). More importantly, *PGMRC1* was more abundantly expressed in testes of juveniles than broodstock of *P. monodon* ($P < 0.05$). *Inx2* and *Dmc1* were preferentially expressed in ovaries of juveniles to broodstock of *P. monodon* ($P < 0.05$).

Recently, Khamnamtong et al., (2006) identified sex-specific (or differential) expression markers in ovaries and testes of *P. monodon* by RNA-arbitrary-primed-polymerase chain reaction (RAP-PCR). Expression patterns of 25 transcripts were tested against the first stranded cDNA of ovaries and testes of 3-month-old and broodstock-sized *P. monodon (N* = 5 and *N* = 7 - 10 for females and *N* = 4 and *N* = 5 - 7 for males, respectively). Five (*FI-4*, *FI-44*, *FIII-4*, *FIII-39* and *FIII-58*) and two (*M457-A01* and *MII-51*) RAP-PCR derived unknown transcripts revealed female- and male-specific expression patterns implying that these unknown genes may contribute gonadal development and/or sex differentiation of *P. monodon*.

In the giant freshwater prawn (*Macrobrachium rosenbergii*), a suppression subtractive hybridization (SSH) male reproductive tract library was constructed to identify male-specific genes that could be involved in male development. A novel *Mar-Mrr* (*M. rosenbergii male reproduction-related gene*, 683 bp in length with an ORF of 333 bp) and the *Kazal-type peptidase inhibitor* (*KPI*, 736 bp, ORF of 405 bp) transcripts were identified and these genes were only expressed in the male reproductive tract of *M. rosenbergii* (Cao et al., 2006 and 2007).

The transcripts restrictively (or preferentially) expressed in ovaries but not testes of *P. monodon* found in this study can be additionally used as the responsive indicators for reproductive maturation at the present stages but their involvement for ovarian and oocyte maturation of *P. monodon* and/or differentiation of sexes in *P. monodon* need to be further examined.

In addition, a large scale expression analysis of genes indifferent stages of testicular development and those under the normal and dopamine-treated conditions were also examined by microarray analysis. Many transcripts exhibiting differential expression patterns were found. Based on the fact that genes fabricated on the examined microarray slides were from the hemocyte cDNA libraries of *P. monodon* and *M. japonicus*., microarrays containing genes expressed in gonads of shrimp are more appropriate for screening of reproductively related genes in *P. monodon*.

More recently, the first reproduction cDNA microarray (*Repro*Array^{GST}) containing 4,992 features amplified from cDNAs of ovary (1,920) and testis (3,072) EST libraries of *P. monodon* was constructed and subjected to a high-throughput gene

expression analysis in four different stages of ovarian development (previtellogenic, vitellogenic, early cortical rod and late cortical rod stages). This allows rapid examination of expression profiles of a large number of reproduction-related genes in both ovaries and testes to provide an insight into molecular aspects governing reproductive processes of *P. monodon* (N. Karoonuthaisiri, personal communication).

Gene expression and tissue distribution analysis are important and provide the basic information to set up the priority for further analysis of functional genes. Based on the fact that a particular gene may express in several tissues and possesses a different function in different tissues, tissue distribution analysis was carried out

Tissue distribution analysis revealed that *PMTST1* was more abundantly expressed in testes, less abundantly expressed in intestine but was not expressed in other tissues. In contrast, *MIPP2* was only expressed in gonads of *P. monodon* broodstock. *CYA*, *SUMO-1*, *prohibitin2*, and *Trap240* were constitutively expressed in all examined tissues while *HSP70-2* was more abundantly expressed in testes than other tissues. *PMTra-2* was expressed at high levels in testes and ovaries but less abundantly expressed in gills, heart, hemocytes, lymphoid organs, intestine and stomach.

MIPP functionally dephosphorylates a number of inositol phosphates, including Ins(1,3,4,5)P_4, $InsP_5$ and $InsP_6$ which is recognized to have important cellular actions as receptor-mobilized precursor pools for intracellular signals (Craxton et al., 1997). HSP70-2 is a testis-specific member of the HSP70 protein family known to play a critical role in the completion of meiosis during male germ cell differentiation (Govin et al., 2006).

Tissue distribution analysis also indicated that *Inx2* and *Dmc1* were more abundantly expressed in testes than ovaries while *PGMRC1* which was constitutively expressed in all examined tissues showed comparable expression levels in testes and ovaries of *P. monodon* broodstock. Therefore, *Inx2* and *Dmc1* may play the important role in spermatogenesis but not oogenesis whereas *PGMRC1* is functionally important for both spermatogenesis and oogenesis of *P. monodon*.

Invertebrate gap-junction proteins, Inxs, were originally identified in *Drosophila* and *Caenorhabditis* (Phelan *et al*., 1998). In *Bombyx mori*, northern

blotting and *in situ* hybridization revealed that *Bm-Inx2* was expressed across all developmental stages and in various tissues with high expression observed in the nervous system during embryogenesis. In contrast, *Bm-Inx4* was transiently expressed at the germ-band formation stage of embryogenesis, and was specifically expressed in ovaries and testes during the larval and pupal stages (Hong *et al*., 2008).

Spermatogenesis is an essential process for production of haploid gametes. During meiosis, a single round of DNA replication is followed by two successive rounds of nuclear divisions (Abe, 1987). Dmc1 is involved in meiotic recombination occurred during the meiotic prophase (Ozaki *et al*., 2006). RNA interference (RNAi) against endogenous *Dmc1* defects spermatogenesis in mice indicating its important roles in spermatogenesis (Shoji *et al*., 2005). Recently, the full length *Dmc1* cDNA was cloned from the testis of the Japanese eel (*Anguilla japonica*). *Dmc1* mRNA of *A. japonica* was abundantly expressed in the testes and ovaries and lower expressed in the brain. *In situ* hybridization revealed that *A. japonica Dmc1* was localized only in the primary spermatocytes implying its important role during the initial stages of spermatogenesis (Kajiura-Kobayashi *et al*., 2005).

Molecular mechanisms and expression patterns of genes controlling different steps of sperm maturation and testicular development should be examined for better understanding the reproductive maturation of *P. monodon* in captivity. The ability to identify sex-specific and sex-differential expression markers of *P. monodon* opens a possibility to study the initial expression and localization of these gene products in undifferentiated gonads of *P. monodon* by *in situ* hybridization. Typically, the study of the effectors involving shrimp reproduction has been limited to the determination of one effector at a time (Ibara, et al., 2007). Therefore integrated interactions among genes in gonads of *P. monodon* need to be examined, for example, by microarray analysis.

***In vitro* expression of recombinant proteins**

Recombinant spermatogonial stem cell renewal factor, Dmc1, PGRMC1, cyclophilin A, and SUMO-1 protein were successfully expressed *in vitro* for further examination on expression of these genes at both the translational levels. Recombinant spermatogonial stem cell renewal factor cyclophilin A and Dmc1

proteins were purified. The polyclonal antibodies of these recombinant proteins were successfully produced. Localization of these proteins at different stages of testicular development can be subsequently carried out using immunohistochemistry.

Expression of functionally important gene homologues and the potential to be applied as biomarkers for reduced maturation of *P. monodon*

In the red swamp crayfish (*Procambarus clarkia*) dopamine inhibited testicular maturation dose-dependently whereas its antagonist, spiperone induced testicular maturation (Sarojini et al., 1995) but effects of dopamine on spermatogenesis of *P. monodon* have not been reported. Real-time quantitative PCR revealed that expression levels of *Dmc1* in testes were not significantly altered after dopamine treatment ($P > 0.05$) whereas *PM-PGMRC1* was up-regulated at 3 hr post treatment ($P < 0.05$). This suggested that dopamine may not inhibit spermatogenesis of *P. monodon*. However, the preliminary results in this study should be further confirmed in *P. monodon* broodstock at both the mRNA and protein levels.

Domesticated male *P. monodon* yielded lower fertilization rates of zygotes and low survival rates of larvae than did wild male *P. monodon* (B. Withyachumnarnkul personal communication). In this study, the GSI values of wild shrimp were greater than those of domesticated broodstock even though their body weights were comparable implying a possible reduction of the maturation potential in domesticated shrimp. It has been reported that overexpression of *prohibitin* (also called a repressor of estrogen receptor activity, REA) inhibits *estrogen receptor alpha* (*ERα*) transcriptional activity in MCF-7 breast cancer cells (He *et al*., 2008). Lower expression levels of *SUMO-1 Dmc1*, and *spermatogonial stem-cell renewal factor* and greater levels of *prohibitin2* in domesticated than wild broodstock ($P < 0.05$) suggested that transcriptional levels of these genes may be used to indicate possible reduced degrees of reproductive maturation of different full-sib families in the on-going domestication program of *P. monodon* in Thailand.

Recently, Parnes et al. (2006) indicated that sexually mature males of *Litopenaeus vannamei* are strictly associated with molt cycles and are hormonally regulated periodically. A decline in quality of spermatophore as the molt cycle progressed was also reported in *P. indicus* (Muthuraman, 1997). Expression levels of

SUMO-1, *CYA, Tra-2, HSP70-2*, *MIPP2*, *prohibitin2*, *Trap240*, *Dmc1*, *saposin* and *spermatogonial stem-cell renewal factor* in testes of *P. monodon* were reduced whereas *PMTST1* was up-regulated after molting. Accordingly, expression levels of these transcripts during the molt cycles of male *P. monodon* should be examined and applied to explore the possible association between their expression levels and testicular development of *P. monodon*.

In the present study, a large number of cDNA including sex-related transcripts in testes of *P. monodon* were identified. The expression profiles of genes specifically expressed or those preferentially expressed in testes of *P. monodon* illustrated in this study implied that these genes may have contributed testicular development in *P. monodon*. Functionally analysis of genes involving testicular and sperm development can be further carried out for better understanding of the reproductive maturation of male *P. monodon* in captivity.

CHAPTER V

CONCLUSIONS

1. Conventional cDNA libraries of testis and heart of *P. monodon* were constructed. A total of 896 and 413 ESTs were sequenced, respectively and 606 ESTs (67.6%) and 283 ESTs (68.5%) significantly matched sequences in the GenBank whereas 290 ESTs (32.4%) and 130 ESTs (31.5%) were unknown transcripts, respectively. Genes functionally involved in cell growth and development, sex determination cascades, and testicular development were found.

2. Forward and reverse SSH cDNA libraries between cDNA in testes of broodstock and juvenile *P. monodon* were established to promote the identification of differentially expressed genes in testes of *P. monodon* juveniles and broodstock. In total, 178 and 187 ESTs from respective libraries were sequenced and 67 ESTs (37.1%) and 104 ESTs (54.0%) of which were known genes whereas 112 and 87 ESTs accounting for 62.9 and 46.0% were unknown genes, respectively.

3. The full length cDNA of 16 functionally important gene homologues were successfully characterized by RACE-PCR. These included *PMTST1*, *multiple inositol polyphosphate phosphatase 2* (*MIPP2*), *prohibitin-2*, *cell division kinase 7 (cdk7)*, *flotillin 2*, *growth factor receptor-bound protein*, *innexin 1*, *innexin 2*, *Rac-GTPase activating protein 1*, *transformer 2 (Tra-2)*, *Dmc1*, *progestin membrane receptor component 1 (PGMRC1)*, *saposin*, *troponin T isoform 3*, *Ero1L CG1333-PB isoform B*, and *dihydrolipoamide dehydrogenas*e, respectively.

4. Expression patterns of 59 gene homologues were non-quantitatively examined using RT-PCR analysis. *PMTST1* was only expressed in testes but not ovaries. *MIPP* and *MIPP2 HSP70-2* exhibited a trend of preferential expression in testes of *P. monodon*. However, thirty-six transcripts showed a trend of greater expression levels in ovaries than testes.

5. Tissue distribution analysis of 15 gene homologues revealed that *PMTST1* was expressed in testes and intestine but was not expressed in other tissues. *MIPP* and *MIPP2* were only expressed in gonads of *P. monodon* broodstock. *Tra-2* was

expressed at high levels in testes and ovaries but less abundantly expressed in other tissues. Other transcripts did not reveal more abundantly expression in gonads compared to other tissues.

6. Expression levels of 12 gene homologues were examined using semiquantitative RT-PCR or quantitative real-time PCR. Testis-specific expression of *PMTST1* was confirmed by semiquantitative RT-PCR. *Dmc1*, *saposin*, *spermatogonial stem-cell renewal factor*, *MIPP2* and *HSP70-2* were preferentially expressed in testes to ovaries ($P < 0.05$) but *CYA* and *Trap240* were differentially expressed in the opposite direction. Expression levels of *SUMO-1*, *Tra-2* and *prohibitin2* in ovaries and testes of *P. monodon* were not significantly different ($P > 0.05$).

7. Significant reduction of *SUMO-1*, *Dmc1*, and *spermatogonial stem-cell renewal factor* and increment of *prohibitin2* transcripts in domesticated broodstock ($P < 0.05$) suggested that these reproductively related genes may be used as biomarkers to evaluate reduced degrees of reproductive maturation in domesticated *P. monodon*.

8. Relative expression levels of *PGMRC1* but not *Dmc1* in testes of juvenile *P. monodon* were induced upon DA administration (10^{-6} mol/shrimp) at 3 h post treatment ($P < 0.05$).

9. Recombinant proteins of *Dmc1*, *spermatogonial stem-cell renewal factor, SUMO-1*, and *CYA* were successfully expressed *in vitro*. Polyclonal antibodies against *Dmc1*, *spermatogonial stem-cell renewal factor* and *CYA* were successfully produced.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

REFERENCES

Abe, S-I. 1987. Differentiation of spermatogenic cells from vertebrates *in vitro*. Int. Rev. Cytol. 109: 159-209.

Abe, S.-I. 2004. Hormonal control of meiosis initiation in the testis from Japanese newt, *Cynops pyrrhogaster*. Zool. Sci. 21: 691–704.

Alfaro, J., Zúñiga, G., and Komen, J. 2004. Induction of ovarian maturation and spawning by combined treatment of serotonin and a dopamine antagonist, spiperone in *Litopenaeus stylirostris* and *Litopenaeus vannamei*. Aquaculture 236: 511-522.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Amparyup, P., Klinbunga, S., Tassanakajon, A., Hirono, I., Aoki, T., and Jarayabhand, P. 2004. Expressed sequence tag (EST) analysis of ovaries and testes from the tropical abalone (*Haliotis asinina*). Mar. Biotechnol. 6: S365-S370.

Avise, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York.

An, W., and Wensink, P. C. 1995a. Integrating sex- and tissue-specific regulation within a single *Drosophila* enhancer. Genes Dev. 9: 256-266.

An, W., and Wensink, P. C. 1995b. Three protein binding sites from an enhancer that regulates sex- and fat body-specific transcription of *Drosophila yolk protein* genes. EMBO J. 14: 1221-1230.

Anderson, I., and Brass, A. 1998. Searching DNA databases for similarities to DNA sequences: when is a match significant? Bioinformatics 14(4): 349-356.

Aniento, F., Emans, N., Griffiths, G., and Gruenberg, J. 1993. Cytoplasmic dynein-dependent vesicular transport from early to late endosomes. J. Cell. Biol. 123: 1373-1387.

Bailey-Brock, J. H., and Moss, S. M. 1992. Penaeid taxonomy, biology and zoogeography. In Fast, A. W. and Lester, L. J. (eds.), Marine Shrimp Culture: Principles and Practices, pp. 9-23. Netherlands: Elsevier Science Publishers.

Baker, R. H., and Sakai, R. K. 1976. Male determining factor on chromosome 3 in the mosquito, *Culex tritaeniorhynchus*. J. Hered. 67: 289-294.

Belfiore, M., Pugnale, P., Saudan, Z., and Puoti, A. 2004. Roles of the *C. elegans* cyclophilin-like protein MOG-6 in MEP-1 binding and germline fates. Development 131: 2935-2945.

Bell, T.A., and Lightner, D.V. 1988. A Handbook of Normal Penaeid Shrimp Histology. World Aquaculture Society, USA: Allen Press, Inc.

Benzie, J. A. H. 1998. Penaeid genetics and biotechnology. Aquaculture 164: 23-47.

Braydich-Stolle, L., Kostereva, N., Dym, M., and Hofmann, M. C. 2007. Role of Src family kinases and N-Myc in spermatogonial stem cell proliferation. Developmental Biology 304: 34-45.

Browdy, C. 1998. Recent developments in penaeid broodstock and seed production technologies : improving the outlook for superior captive stocks. Aquaculture 164 : 3-21.

Burtis, K. C., Coschigano, K. T., Baker, B. S., and Wensink, P. C. 1991. The doublesex proteins of *Drosophila melanogaster* bind directly to a sex-specific yolk protein gene enhancer. EMBO J. 10: 2577-2582.

Callard, G. V., 1991. Spermatogenesis. In: Pang, K.T., Schreibman, M.P. (Eds.) Vertebrate endrocrinology; fundamentals and biochemical implication, Vol 4 Part A. San Diego: Academic Press. pp. 303-341.

Cao, J. X., Yin, G. L., and Yang, W. L. 2006. Identification of a novel male reproduction-related gene and its regulated expression patterns in the prawn, *Macrobrachium rosenbergii*. Peptides. 27: 728-735.

Cao, J. X., Dai, J. Q., Dai, Z. M., Yin, G. L., and Yang, W. J. 2007. A male reproduction-related Kazal-type peptidase inhibitor gene in the prawn,

Macrobrachium rosenbergii: molecular characterization and expression patterns. Mar. Biotechnol. 9: 45-55.

Craxton, A., Caffrey, J.J., Burkhart, W., Safrany, S.T., and Shears, S.B. 1997. Molecular cloning and expression of a rat hepatic multiple inositol polyphosphate phosphatase. Biochem. J. 328: 75-81.

Chiu, HT., et al. 2006. Dopamine induces transient modulation of the physiological responses of white leg shrimp, *Litopenaeus vannamei*. Aquaculture 251: 558–566.

De Grandi, A., Calvari, V., Bertini, V., Bulfone, A., Peverali, G., Camerino, G., Borsani, G., and Guioli, S. 2000. The expression pattern of a mouse *doublesex*-related gene is consistant with a role in gonadal differentiation. Mech. Dev. 90: 323-326.

Devlin R. H, and Nagahama Y. 2002. Sex determination and sex differentiation in fish. Aquaculture 208: 191–364.

Diatchenko, L., Lau, Y. C., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., Lukyanov, S., Lukyanov, K., Gurskaya, N., Sverdlov, E. D., and Siebert, P. D. 1996. Suppression subtractive hybridization: A method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc. Natl. Acad. Sci. USA 93: 6025-6030.

Eddy, E. M. 1999. Role of heat shock protein HSP70-2 in spermatogenesis. Rev. Reprod. 4: 23-30.

Erdman, S. E., Chen, H. J., and Burtis, K. C. 1996. Functional and genetic characterization of the oligomerization and DNA binding properties of the *Drosophila* Doublesex proteins. Genetics 144: 1639-1652.

Felsenstein, J. 1993. Phylip (Phylogenetic Inference Package) Version 3.5c. Department of Genetics, University of Washington, Seattle, USA.

Ghosh-Roy, A., Kulkarni, M., Kumar, V., Shirolikar, S., and Ray, K. 2004. Cytoplasmic dynein–dynactin complex is required for spermatid growth but not axoneme assembly in *Drosophila*. Mol. Biol. Cell 15: 2470–2483.

Godet, M., Damestoy, A., Mouradian, S., Rudkin, B.B., and Durand, P. 2004. Key role for cyclin-dependent kinases in the first and second meiotic divisions of rat spermatocytes. Biol. Reprod. 70: 1147–1152.

Govin, J., Caron, C., Escoffier, E., Ferro, M., Kuhn, L., Rousseaux, S., Eddy, E.M., Garin, J., and Khochbin, S. 2006. Post-meiotic shifts in HSPA2/HSP70.2 chaperone activity during mouse spermatogenesis. J. Biol. Chem. 281: 37888-37892.

Grimes, S. R. 2004. Testis-specific transcriptional control. Gene 343: 11-22.

He, B., Feng, Q., Mukherjee, A., Lonard, D.M., DeMayo, F.J., Katzenellenbogen, B.S., Lydon, J.P., and O'Malley, B.W. 2008. A repressive role for prohibitin in estrogen signaling. Mol. Endocrinol. 22: 344-360.

Hong, S-M, Kang, S-W, Goo, T-W., Kim, N-S., Lee, J-H., Kim, K-A., and Nho, S-K. 2008. Two gap junction channel (innexin) genes of the *Bombyx mori* and their expression. J. Insect Physiol. 54: 180-191.

Huang, X., and Madan, A. 1999. CAP3: A DNA sequence assembly program. Genome Res. 9: 868-877.

Ibara, A. M., Racotta, I. S., Arcos, F. G., and Palacios, E. 2007. Progress on the genetics of reproductive performance in penaeid shrimp. Aquaculture 268: 23-43.

Imjongjirak, C., Ngewsra, L., Pramual, C., Insuan, S., Pala-or, K., Klinbunga, S., and Sittipraneed,S. 2004. Development of a simple and rapid method for typing population origins of the honey bee (*Apis cerana*) in Thailand. J. Apicul. Res. 43: 17-20.

Inoue, K., Hoshijima, K., Sakamoto, H., and Shimura, Y. 1990. Binding of the Drosophila sex-lethal gene product to the alternative splice site of transformer primary transcript. Nature 344: 461-463.

Janssen, T., Claeys, I., Simonet, G., Loof, A.D., Girardie, J., and Broeck, J.V. 2001. cDNA cloning and transcript distribution of two different neuroparsin precursors in the desert locust, *Schistocerca gregaria*. Insect Mol Biol 10: 183-189.

Jursnich, V. A., and Burtis, K. C. 1993. A positive role in differentiation for the male doublesex protein of *Drosophila*. Dev. Biol. 155: 235-249.

Kajiura-Kobayashi, H., Kobayashi, T., and Nagahama, Y. 2005. Cloning of cDNAs and the differential expression of A-type cyclins and Dmc1 during spermatogenesis in the Japanese eel, a teleost fish. Dev. Dynamics 232: 1115-1123.

Khamnamtong, B., Thumrungtanakit, S., Klinbunga, S., Hirono, T. Aoki, T., and Menasvetas, P., 2006. Identification of sex-specific expression markers of the giant tiger shrimp (*Penaeus monodon*). J. Biochem. Mol. Biol. 39: 37-45.

Kim, H., Moon, C., and Shin, T. 2008. Immunohistochemical study of flotillin-1 in the rat testis during postnatal development. Acta Histochemica 110(3): 224-231.

Klinbunga, S., Penman, D. J., McAndrew, B. J. and Tassanakajon, A. 1999. Mitochondrial DNA diversity in three populations of the giant tiger shrimp, *P. monodon*. Mar. Biotechnol 1(2): 113-121.

Kobayashi, H., Minshull, J., Ford, C., Golsteyn, R., Poon, R., and Hunt, T. 1991. On the synthesis and destruction of A- and B-type cyclins during oogenesis and meiotic maturation in *Xenopus laevis*. J Cell Biol. 114: 755–765.

Koshiyama, A., Hamada, F. N., Namekawa, S. H., Iwabata, K., Sugawara, H., Sakamoto, A., Ishizaki, T., and Sakaguchi, K. 2006. Sumoylation of a meiosis-specific RecA homolog, Lim15/Dmc1, via interaction with the

small ubiquitin-related modifier (SUMO)-conjugating enzyme Ubc9. FEBS J. 273: 4003-4012.

Kuo, C.M. Hsu, C.R., and Lin, C.Y. 1995. Hyperglycemic effects of dopamine in tiger shrimp, *Penaeus monodon*. Aquaculture 135: 161-172.

Lang, K., Schmid, F. X., and Fischer, G. 1987. Catalysis of protein folding by prolyl isomerase. Nature 329: 268-270.

Leelatanawit, R., Klinbunga, S., Puanglarp, N., Tassanakajon, A., Jarayabhand, P., Hirono, I., Aoki, T., and Menasveta, P. 2004. Isolation and Characterization of Differentially Expressed Genes in Ovaries and Testes of the Giant Tiger Shrimp (*Penaeus monodon*). Mar. Biotechnol. 6: S506-S510.

Li, F., Xiang, J., Zhang, X., Zhou, L., Zhang, C., and Wu, C. 2003. Gonad development characteristics and sex ratio in triploid Chinese shrimp (*Fenneropenaeus chinensis*). Mar. Biotechnol. 5: 528-535.

Li, Y., Byrne, K., Miggiano, E., Whan, V., Moore, S., Keys, S., Crocos, P., Preston, N., and Lehnert, S. 2003. Genetic mapping of the kuruma prawn *Penaeus japonicus* using AFLP markers. Aquaculture 219: 143-156.

Limsuwan, C. 2004. Diseases of Pacific white shrimp (*Litopenaeus vannamei*) cultured in Thailand. Proceeding of the JSPS-NRCT International Symposium Joint Seminar 2004: Management of Food Safety in Aquaculture and HACCP. pp. 36-41. Thailand: Kasetsart University.

Liu, Z. J., and Cordes, J. F. 2004. DNA marker technologies and their applications in aquaculture genetics. Aquaculture 238: 1–37.

Liu, G. Y., Zhao, S. M., Song, X. L., Wang, J., and Gao, S. Z. 2008. Isolation, sequence identification and tissue expression distribution of three novel porcine genes-UBC1, TRA2A and SH21A. J. Appl. Animal Res. 33: 33-38.

Lyons, R.E., and Li, Y. 2002. Crustacean genomics: implications for future in aquaculture. AgBiotechNet 4: 1-6.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Menasveta, P., Piyatiratitivorakul, S., Rungsupa, S., Moree, N., and Fast, A.W. 1993. Gonadal maturation and reproductive performance of giant tiger prawn (*Penaeus monodon* Fabricius) from the Andaman Sea and pond-reared in Thailand. Aquaculture 116: 191–198.

Meyer C., Schmid R., Scriba P. C., and Wehling M. 1996. Purification and partial sequencing of high-affinity progesterone-binding site(s) from porcine liver membranes. Eur. J. Biochem. 239: 726-731.

Miura, T., Higuchi, M., Ozaki, Y., Ohta, T., and Miura, C. 2006. Progestin is an essential factor for the initiation of the meiosis in spermatogenetic cells of the eel. Proc. Nati. Acad. Sci. USA 103: 7333–7338.

Miura, T., Ohta, T., Miura, C. I., and Yamauchi, K. 2007. Complementary deoxyribonucleic acid cloning of spermatogonial stem cell renewal factor. Endocrinology 144(12): 5504-5510.

Moniot, B., Berta, P., Scherer, G., Sudbeck, P., and Poulat, F. 2000. Male specific expression suggests role of *DMRT1* in human sex determination. Mech. Dev. 91: 323-325.

Morales, C. R., Zhao, Q., Lefrancois, S., and Ham, D. 2000. Role of prosaposin in the male reproductive system: effect of prosaposin inactivation on the testes, epididymis, prostate, and seminal vesicles. Archives of Andrology 44(3): 173-186.

Mourot, B., Nguyen, T., Fostier, A., and Bobe, J. 2006. Two unrelated putative membrane-bound progestin receptors, progesterone membrane receptor component 1 (PGMRC1) and membrane progestin receptor (mPR) beta, are expressed in the rainbow trout oocyte and exhibit similar ovarian expression patterns Reprod. Biol. Endocrinol. 4, 6 (doi:10.1186/1477-7827-4-6).

Mullis, K.B., and Faloona, F.A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalysed chain reaction. Methods Enzymol. 155: 335–350.

Muthuraman, A. L. 1997. Effect of ecdysial cycle on the semen and spermatophore production of the Indian white shrimp (*Penaeus indicus* H.M. Edwards). Fish. Technol. 34: 11-14.

Page, A. P., MacNiven, K., and Hengartner, M. O. 1996. Cloning and biochemical characterization of the cyclophilin homlogues from the freeliving nematode *Caenorhabditis elegans*. Biochem. J. 317: 179-185.

Parnes, S., Raviv, R., Shechter, A., and Sagi, A. 2006. Males also have their time of the month! Cyclic disposal of old spermatophores, timed by the molt cycle, in a marine shrimp. J. Exp. Biol. 209: 4974-4983.

Phelan, P., Bacon, J. P., Davies, J. A., Stebbings, L. A., Todman, M. G., Avery, L., Baines, R. A., Barnes, T. M., Ford, C., Hekimi, S., Lee, R., Shaw, J. E., Starich, T. A., Curtin, K. D., Sun, Y. A., and Wyman, R. J. 1998. Innexins: a family of invertebrate gap-junction proteins. Trends Genet. 14: 348–349.

Preston, N. P., Sellars, M. J., and Coman, F. E. 2004. Ploidy manipulation induces sterility in Kuruma Prawns. Global Aquaculture Advocate 7: 70-71.

Preechaphol, R., Leelatanawit, R., Sittikhankaew, K., Klinbunga, S., Khamnamtong, B., Puanglarp, N., and Menasveta, P. 2007. Expressed sequence tag analysis for isolation of sex-related transcripts in the giant tiger shrimp (*Penaeus monodon*). J. Biochem. Mol. Biol. 40: 501-510.

Raymond, C. S., Shamu, C. E., Shen, M. M., Seifert, K. J., Hirsch, B., Hodgkin, J., and Zarkower, D. 1998. Evidence for evolutionary conservation of sex-determining genes. Nature 391: 391-695.

Raymond, C. S., Kettlewel, J. R., Hirsch, B., Bardwell, V. J., and Zarkower, D. 1999. Expression of *DMRT1* in the genital ridge of mouse and chicken embryos suggests a role in vertebrate sexual development. Dev. Biol. 215: 208-220.

Rungnapa Leelatanawit, “Cloning and characterization of cDNA involved in sex determination of the Giant Tiger Shrimp *Penaeus monodon*,” Master’s thesis, Faculty of Science, Chulalongkorn University, 2003.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406-425.

Sambrook, J., and Russell, D.W. 2001. Molecular Cloning: A Laboratory Manual, 3rd edition. Cold Spring Harbor, NY : Cold Spring Habor Laboratory Press.

Sarojini, R., Nagabhushanam, R., and Fingerman, M. 1995. *In vivo* effects of dopamine and dopaminergic antagonists on testicular maturation in the red swamp crayfish, *Procambarus clarkii.* Biol. Bull. 189: 340–346.

Sassone-Corsi, P. 1998. CREM: a master-switch governing male germ cells differentiation and apoptosis. Semin. Cell Dev. Biol. 9: 475-482.

Sassone-Corsi, P. 2002. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. Science 296: 2176-2178.

Schmitt, A., Gofferje, V., Weber, M., Meyer, J., Mossner, R., and Lesch, K.P. 2003. The brain-specific protein MLC-1 implicated in megalencephalic leukoencephalopathy with subcortical cysts is expressed in glial cells in the murine brain. Glia 44(3): 283-295.

Shen, M. M., and Hodgkin, J. 1988. *mab-3*, a gene required for sex-specific yolk protein expression and a male-specific lineage in *C. elegans*. Cell 54: 1019-1031.

Shen, B., Zhang, Z., Wang, Y., Wang, G., Chen, Y., Lin, P., Wang, S., and Zou, Z. 2008. Differential expression of ubiquitin-conjugating enzyme E2r in the developing ovary and testis of penaeid shrimp *Marsupenaeus japonicus*. Mol. Biol. Rep. (in press).

Shoji M., Chuma, S., Yoshida, K., Morita, T., and Nakatsuji, N. 2005. RNA interference during spermatogenesis in mice. Dev. Biol. 282: 524-534.

Staelens, J., Rombaut, D., Vercauteren, I., Argue, B., Benzie, J., and Vuylsteke, M. 2008. High-density linkage maps and sex-linked markers for the black tiger shrimp (*Penaeus monodon*). Genetics 179(2): 917-925.

Takahashi, Y., and Kikuchi, Y. 2005. Yeast PIAS-type Ull1/Siz1 is composed of SUMO ligase and regulatory domains. J. Biol. Chem. 280: 35822–35828.

Tassanakajon, A., Klinbunga, S., Paunglarp, N., Rimphanitchayakit, V., Udomkit, A., Jitrapakdee, S., Sritunyalucksana, K., Phongdara, A., Pongsomboon, S., Supungul, P., Tang, S., Kuphanumart, K., Pichyangkura, R., and Lursinsap, C. 2006. *Penaeus monodon* gene discovery project: The generation of an EST collection and establishment of a database. <u>Gene</u> 384: 104-112.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight metrix choice. <u>Nucleic Acids Res.</u> 22: 4673-4680.

Wilson, K. et al. 2002. Genetic mapping of the black tiger shrimp *Penaeus monodon* with amplified fragment length polymorphism. <u>Aquaculture</u> 204: 297-309.

Withyachumnarnkul, B., Boonsaeng, V., Flegel, T.W., Panyim, S., and Wongteerasupaya C. 1998. Domestication and selective breeding of *Penaeus monodon* in Thailand. In Felgel, T. (Ed.), <u>Proceedings to the Special Session on Advances in Shrimp Biotechnology, The Fifth Asian Fisheries Forum: International Conference on Fisheries and Food Security Beyond the Year 2000</u>, pp. 73-77. 11-14 November 1998. Thailand: Chiangmai.

Zhu, D., Dix, D. J., and Eddy, E. M. 1997. HSP70-2 is required for CDC2 kinase activity in meiosis I of mouse spermatocytes. <u>Development</u> 124: 3007-3014.

Zirkin, B. R. 1993. Regulation of spermatogenesis in the adult mammals: gonadotropin and androgens. In: Desjardins, C., Ewing, L.L. (Eds.), Cell and Molecular Biology of the Testis. New York: Oxford University Press. pp. 166-188.

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

APPENDICES

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Appendix A

Table A1 Examples of transcripts from a testis cDNA library categorized as members of cell division/DNA synthesis, repair and replication (29 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0077-W	*14-3-3-like protein*	*Penaeus monodon*	0
TT-N-S01-1039-W			1e-112
TT-N-S01-0891-W	*AAA ATPase domain-containing protein*	*Strongylocentrotus purpuratus*	3e-13
TT-N-S01-0076-W	*Actin-binding protein anillin, contractile ring component anillin*	*Xenopus laevis*	3e-56
TT-N-S01-0241-W	*Anti-silencing factor 1 CG9383-PA*	*Drosophila melanogaster*	5e-61
TT-N-S01-0153-W	*BUB3 budding uninhibited by benzimidazoles 3 homolog (yeast)*	*Xenopus tropicalis*	2e-72
TT-N-S01-0190-W	*Cell division cycle 2*	*Danio rerio*	2e-31
TT-N-S01-0169-W	*Cell division protein kinase 7 (CDK-activating kinase) (CAK)*	*Mus musculus*	8e-85
TT-N-S01-0752-W	*checkpoint kinase 1 (Serine/threonine-protein kinase)*	*Strongylocentrotus purpuratus*	9e-76
TT-N-S01-0525-W	*CHK1 checkpoint homolog (Serine/threonine-protein kinase)*	*Xenopus tropicalis*	3e-06
TT-N-S01-0695-W	*Cyclin dependent kinase 2*	*Sphaerechinus granularis*	8e-82
TT-N-S01-1020-W	*DEAD (Asp-Glu-Ala-Asp) box polypeptide 31 isoform 2*	*Pan troglodytes*	4e-26
TT-N-S01-0187-W	*DNA polymerase beta*	*Xenopus laevis*	3e-59
TT-N-S01-0412-W	*DNA replication licensing factor MCM6 (Mis5 homolog)*	*Strongylocentrotus purpuratus*	5e-71
TT-N-S01-0042-W	*Methyl-CpG binding domain protein 4*	*Xenopus tropicalis*	5e-14
TT-N-S01-0371-W	*novel protein similar to vertebrate PCTAIRE protein kinase 2 (PCTK2)*	*Danio rerio*	2e-95
TT-N-S01-0198-W	*Nucleic acid-associated protein 36*	*Asterina pectinifera*	1e-08
TT-N-S01-0895-W	*Ornithine decarboxylase (ODC)*	*Apis mellifera*	6e-15
TT-N-S01-0129-W	*Oncoprotein nm23*	*Litopenaeus vannamei*	1e-66
TT-N-S01-0192-W	*Origin recognition complex, subunit 1-like*	*Xenopus tropicalis*	5e-04
TT-N-S01-0733-W	*peptidyl-prolyl cis-trans isomerase (Cyclophilin 1)*	*Bombyx mori*	6e-72
TT-N-S01-0556-W	*RAD1 homolog isoform 1*	*Rattus norvegicus*	5e-13
TT-N-S01-0220-W	*RAS protein*	*Bombyx mori*	1e-56
TT-N-S01-0203-W	*RAS-related GTP binding protein*	*Bombyx mori*	4e-62
TT-N-S01-0899-W	*Replication protein A1*	*Mus musculus*	5e-79
TT-N-S01-0912-W	*Replication protein A2*	*Mus musculus*	3e-16
TT-N-S01-0425-W	*SMC1 structural maintenance of chromosomes 1-like 1*	*Macaca mulatta*	2e-08
TT-N-S01-0216-W	*Split hand/foot malformation (ectrodactyly) type 1*	*Xenopus tropicalis*	5e-21
TT-N-S01-0657-W	*WD repeat domain 61 (Meiotic recombination REC14 protein homolog)(WDR61)*	*Bombyx mori*	3e-64

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A2 Examples of transcripts from a testis cDNA library categorized as members of defense and homeostasis (30 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0295-W	*70 kD heat shock protein*	*Mirocaris fortunata*	1e-92
TT-N-S01-0588-W			1e-122
TT-N-S01-0894-W			5e-68
TT-N-S01-0066-W	*Antimicrobial peptide*	*Fenneropenaeus chinensis*	3e-60
TT-N-S01-0284-W	*Antimicrobial peptide*	*Litopenaeus setiferus*	1e-43
TT-N-S01-0285-W			8e-44
TT-N-S01-0020-W	*Cyclophilin A*	*Chlamys farreri*	7e-74
TT-N-S01-0815-W	*Defender against apopototic cell death 1*	*Argopecten irradians*	2e-07
TT-N-S01-0882-W	*Ferritin light chain-like*	*Culicoides sonorensis*	6e-05
TT-N-S01-0351-W	*Heat shock cognate 70 protein*	*Trichoplusia ni*	4e-32
TT-N-S01-0846-W	*Heat shock protein 60*	*Liriomyza sativae*	1e-109
TT-N-S01-0278-W	*Heat shock protein 70*	*Marsupenaeus japonicus*	3e-52
TT-N-S01-0333-W			2e-62
TT-N-S01-0339-W			1e-114
TT-N-S01-0959-W			1e-115
TT-N-S01-0679-W	*Heat shock 70kDa protein 8 isoform 2*	*Homo sapiens*	6e-66
TT-N-S01-0003-W	*Hemomucin*	*Aedes aegypti*	2e-07
TT-N-S01-0082-W	*Influenza virus NS1A binding protein isoform a*	*Tribolium castaneum*	4e-43
TT-N-S01-0586-W	*Inhibitor of apoptosis protein*	*Bombyx mori*	3e-19
TT-N-S01-0436-W	*Latent nuclear antigen*	*Aedes aegypti*	3e-41
TT-N-S01-1016-W	*Lymphoid organ expressed yellow head virus receptor protein*	*Penaeus monodon*	3e-93
TT-N-S01-0163-W	*Myeloid leukemia factor (Myelodysplasia-myeloid leukemia factor) (dMLF)*	*Tribolium castaneum*	6e-06
TT-N-S01-0786-W	*nucleolysin tia-1*	*Aedes aegypti*	2e-22
TT-N-S01-0030-W	*Programmed cell death protein*	*Aedes aegypti*	7e-63
TT-N-S01-1025-W	*Retinaldehyde binding protein*	*Aedes aegypti*	4e-26
TT-N-S01-0779-W	*small glutamine-rich tetratricopeptide*	*Gallus gallus*	6e-19
TT-N-S01-0971-W	*transcript expressed during hematopoiesis 2*	*Bos taurus*	3e-15
TT-N-S01-0674-W	*Programmed cell death protein 7 (ES18)*	*Rattus norvegicus*	1e-09
TT-N-S01-0684-W	*Stimulated by retinoic acid 13*	*Danio rerio*	1e-10
TT-N-S01-0683-W	*Vpr (HIV-1) binding protein, isoform CRA_b*	*Homo sapiens*	4e-22

Table A3 Examples of transcripts from a testis cDNA library categorized as members of gene expression, regulation and protein synthesis (86 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-1014-W	*4Fe-4S ferredoxin, iron-sulfur binding:Protein of unknown function DUF224:FAD linked oxidase, C-terminal:FAD linked oxidase, N-terminal*	*Kineococcus radiotolerans SRS30216*	3e-13
TT-N-S01-0615-W	*actin depolymerizing factor (Cofilin/actin-depolymerizing factor homolog) (Protein D61) (Protein twinstar)*	*Aedes aegypti*	5e-44
TT-N-S01-0450-W	*Activating transcription factor 2*	*Bos taurus*	4e-14
TT-N-S01-0111-W	*Aldo-keto reductase*	*Aedes aegypti*	4e-73

Table A3 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0917-W	*Arginyl-tRNA--protein transferase 1 (Arginyltransferase 1) (Arginine-tRNA--protein transferase 1)*	*Apis mellifera*	5e-43
TT-N-S01-0075-W	*Argonaute (plant)-Like Gene family member (alg-1)*	*Caenorhabditis elegans*	2e-38
TT-N-S01-0090-W	*B-cell receptor-associated protein 37*	*Tribolium castaneum*	3e-84
TT-N-S01-0031-W	*BCS-2*	*Balanus amphitrite*	4e-11
TT-N-S01-0961-W	*cleavage and polyadenylation specificity factor*	*Aedes aegypti*	6e-87
TT-N-S01-1093-W			3e-86
TT-N-S01-0415-W	*Cleavage stimulation factor 64-kDa subunit*	*Bombyx mori*	1e-82
TT-N-S01-0475-W	*COP9 constitutive photomorphogenic homolog subunit 5 isoform 1*	*Apis mellifera*	5e-73
TT-N-S01-0413-W	*Cre (cAMP responsive element) binding protein-like 2 isoform 2*	*Gallus gallus*	1e-24
TT-N-S01-0990-W	*DNA-directed RNA polymerase III subunit 127.6 kDa polypeptide (RNA polymerase III subunit 2) (RPC2) isoform 1*	*Apis mellifera*	5e-97
TT-N-S01-0290-W	*Doublecortin and CaM kinase-like 2*	*Gallus gallus*	3e-64
TT-N-S01-0045-W	*eif3s12-prov protein (eukaryotic translation initiation factor 3, subunit 12)*	*Xenopus tropicalis*	1e-30
TT-N-S01-0055-W	*Elongation factor 1 alpha*	*Litopenaeus stylirostris*	4e-48
TT-N-S01-0058-W		*Pocillopora damicornis*	6e-81
TT-N-S01-0154-W		*Penaeus monodon*	5e-72
TT-N-S01-0159-W		*Locusta migratoria*	1e-71
TT-N-S01-0360-W		*Pocillopora damicornis*	2e-98
TT-N-S01-0416-W		*Culicoides sonorensis*	2e-85
TT-N-S01-0569-W		*Armadillidium vulgare*	2e-47
TT-N-S01-0648-W		*Schizophyllum commune*	1e-70
TT-N-S01-0947-W		*Schizophyllum commune*	1e-94
TT-N-S01-1030-W		*Litopenaeus stylirostris*	2e-24
TT-N-S01-0746-W	*Elongation factor 1 beta'*	*Bombyx mori*	2e-51
TT-N-S01-0167-W	*Elongation factor-2*	*Libinia emarginata*	1e-80
TT-N-S01-0704-W			1e-114
TT-N-S01-0221-W	*Negative elongation factor B (NELF-B)*	*Rattus norvegicus*	4e-35
TT-N-S01-0838-W	*Negative elongation factor B homolog*	*Drosophila melanogaster*	1e-25
TT-N-S01-0272-W	*Tail muscle elongation factor 1 gamma*	*Procambarus clarkii*	1e-85
TT-N-S01-1031-W			1e-114
TT-N-S01-0907-W	*Transcription elongation factor 1 homolog*	*Drosophila melanogaster*	8e-37
TT-N-S01-0587-W	*Transcription factor 2B*	*Bombyx mori*	1e-68
TT-N-S01-0342-W	*Translation elongation factor 2*	*Spodoptera exigua*	5e-53
TT-N-S01-0269-W	*Translation initiation factor 4C (1A)*	*Anopheles gambiae*	3e-63
TT-N-S01-0294-W	*Translation initiation factor 4C (1A)*	*Anopheles gambiae*	7e-65
TT-N-S01-0084-W	*Eukaryotic initiation factor 4A*	*Callinectes sapidus*	1e-118
TT-N-S01-0663-W	*Eukaryotic translation initiation factor 1A*	*Drosophila melanogaster*	4e-11
TT-N-S01-0260-W	*Eukaryotic translation initiation factor 2c*	*Aedes aegypti*	2e-30
TT-N-S01-0122-W	*Eukaryotic translation initiation factor 3 subunit 2 beta*	*Bombyx mori*	3e-58
TT-N-S01-0139-W			2e-69
TT-N-S01-0418-W			1e-42
TT-N-S01-0701-W	*Eukaryotic translation initiation factor 3, subunit 9 (eta)*	*Rattus norvegicus*	6e-18
TT-N-S01-0386-W	*Eukaryotic translation initiation factor 4 gamma, 2*	*Nasonia vitripennis*	9e-30
TT-N-S01-0707-W	*Glutaminyl-tRNA synthetase*	*Trypanosoma cruzi*	1e-47

Table A3 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0966-W	*Glycyl-tRNA synthetase*	*Aedes aegypti*	1e-67
TT-N-S01-0189-W	*Growth factor receptor-bound protein*	*Aedes aegypti*	7e-55
TT-N-S01-0230-W	*heterogeneous nuclear ribonucleoprotein*	*Aedes aegypti*	1e-66
TT-N-S01-0430-W	*IMP4, U3 small nucleolar ribonucleoprotein, homolog*	*Apis mellifera*	2e-48
TT-N-S01-0071-W	*Low molecular weight neurofilament protein*	*Xenopus laevis*	8e-04
TT-N-S01-0206-W	*LSM4 homolog, U6 small nuclear RNA associated*	*Danio rerio*	7e-46
TT-N-S01-0009-W	*Microsomal signal peptidase 12 kDa subunit-like*	*Ixodes scapularis*	3e-23
TT-N-S01-1007-W	*Multiprotein bridging factor 1*	*Bombyx mori*	3e-50
TT-N-S01-1026-W			3e-50
TT-N-S01-0306-W	*Nascent polypeptide associated complex protein alpha subunit CG8759-PB, isoform B isoform 1*	*Apis mellifera*	6e-65
TT-N-S01-0052-W	*Polyadenylate binding protein II*	*Apis mellifera*	6e-74
TT-N-S01-0916-W	*Proteasome subunit beta type 2 (Proteasome component C7-I) (Multicatalytic endopeptidase complex subunit C7-I)*	*Tribolium castaneum*	6e-64
TT-N-S01-0785-W	*Protein disulfide isomerase family A, member 6*	*Xenopus tropicalis*	2e-08
TT-N-S01-0087-W	*Protein mago nashi*	*Apis mellifera*	3e-76
TT-N-S01-0974-W	*Protein mago nashi (mago-nashi homolog, proliferation-associated) MAGOH*	*Apis mellifera*	4e-60
TT-N-S01-0257-W	*Receptor expression-enhancing protein 5 (Polyposis locus protein 1 homolog)*	*Pongo pygmaeus*	4e-54
TT-N-S01-0340-W			2e-54
TT-N-S01-0027-W	*Riboflavin kinase*	*Bos taurus*	1e-49
TT-N-S01-0199-W	*Ribonuclease H1 CG8729-PB, isoform B*	*Drosophila melanogaster*	5e-29
TT-N-S01-0039-W	*Ring finger protein 20*	*Gallus gallus*	8e-06
TT-N-S01-0847-W	*Ruvbl2-prov protein (RuvB-like DNA helicase reptin)*	*Xenopus laevis*	1e-115
TT-N-S01-0427-W	*Serine/arginine repetitive matrix 1*	*Gallus gallus*	4e-05
TT-N-S01-0667-W	*Serine/threonine protein kinase Pto (Pto-like serine/threonine kinase)*	*Lycopersicon esculentum*	3e-07
TT-N-S01-0098-W	*Small nuclear ribonucleoprotein D2*	*Macaca mulatta*	8e-29
TT-N-S01-0967-W	*Small nuclear ribonucleoprotein D2-like protein*	*Toxoptera citricida*	5e-41
TT-N-S01-0673-W	*small nuclear ribonucleoprotein E*	*Bombyx mori*	1e-31
TT-N-S01-0915-W			1e-15
TT-N-S01-0873-W	*Small nuclear ribonucleoprotein polypeptide G*	*Homo sapiens*	2e-23
TT-N-S01-0259-W	*Spermidine synthase*	*Gallus gallus*	2e-39
TT-N-S01-0941-W	*Transformer-2 protein A*	*Bombyx mori*	3e-34
TT-N-S01-0985-W	*Transformer-2 protein B*	*Bombyx mori*	6e-57
TT-N-S01-0757-W	*tRNA-dihydrouridine synthase*	*Aedes aegypti*	3e-36
TT-N-S01-0205-W	*U2 small nuclear ribonucleoprotein auxiliary factor 2 isoform 1*	*Bombyx mori*	2e-62
TT-N-S01-0565-W	*winged helix nude*	*Branchiostoma lanceolatum*	1e-32
TT-N-S01-0644-W			3e-11
TT-N-S01-0883-W	*Zinc finger protein 420*	*Danio rerio*	8e-44
TT-N-S01-0874-W	*zinc finger protein 420 isoform 1*	*Mus musculus*	2e-33
TT-N-S01-0573-W	*Zinc finger protein 501*	*Pongo pygmaeus*	4e-45
TT-N-S01-0578-W	*Zinc finger, ZZ domain containing 3*	*Apis mellifera*	3e-48

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A4 Examples of transcripts from a testis cDNA library categorized as members of internal/external structure and motility (40 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0315-W	*Acid cluster protein 33, spastic paraplegia 21 (autosomal recessive, Mast syndrome)*	*Apis mellifera*	1e-81
TT-N-S01-0222-W	*Alpha tubulin*	*Apis mellifera*	1e-70
TT-N-S01-0338-W	*ARP1 actin-related protein 1 homolog A, centractin alpha*	*Gallus gallus*	9e-36
TT-N-S01-0557-W	*cdp-diacylglycerol—glycerol-3-phosphate 3-phosphatidyltransferase*	*Aedes aegypti*	3e-24
TT-N-S01-0072-W	*Cell wall-plasma membrane linker protein homolog*	*Arabidopsis thaliana*	3e-07
TT-N-S01-0567-W	*Chromobox protein homolog 1 (Heterochromatin protein 1 homolog beta) (HP1 beta) (Modifier 1 protein)*	*Apis mellifera*	1e-58
TT-N-S01-0067-W	*Dynactin 5 (p25)*	*Strongylocentrotus purpuratus*	2e-63
TT-N-S01-1055-W	*Flotillin 2 CG32593-PB, isoform B*	*Drosophila melanogaster*	4e-48
TT-N-S01-0130-W	*Histone H2AV (H2A.F/Z)*	*Nasonia vitripennis*	2e-50
TT-N-S01-0309-W	*H2A histone family, member V isoform 1*	*Apis mellifera*	4e-64
TT-N-S01-0148-W	*H3 histone, family 3B*	*Mus musculus*	1e-69
TT-N-S01-0725-W			2e-68
TT-N-S01-0982-W			2e-69
TT-N-S01-0101-W	*Histone H1*	*Mytilus galloprovincialis*	2e-11
TT-N-S01-0376-W	*Histone h1.1*	*Oikopleura dioica*	2e-05
TT-N-S01-1015-W			6e-41
TT-N-S01-0108-W	*Histone H1-delta*	*Strongylocentrotus purpuratus*	7e-45
TT-N-S01-0911-W			1e-45
TT-N-S01-0319-W	*Histone H3*	*Arabidopsis thaliana*	9e-21
TT-N-S01-0471-W		*Entamoeba histolytica HM-1:IMSS*	9e-16
TT-N-S01-0506-W	*ICP4 protein*	*Gallid herpesvirus 3*	7e-06
TT-N-S01-0121-W	*Innexin inx1*	*Schistocerca americana*	3e-72
TT-N-S01-0246-W			3e-17
TT-N-S01-1038-W			5e-11
TT-N-S01-0264-W	*Integrin, beta-like 1*	*Danio rerio*	8e-13
TT-N-S01-1006-W	*Leucine-rich repeat family protein / extensin family protein*	*Arabidopsis thaliana*	9e-15
TT-N-S01-0773-W	*Microtubule-associated protein 1 light chain 3 alpha*	*Xenopus tropicalis*	1e-20
TT-N-S01-0119-W	*Profilin (Chickadee protein)*	*Tribolium castaneum*	1e-37
TT-N-S01-0710-W			1e-37
TT-N-S01-0476-W	*Putative membrane protein*	*Emiliania huxleyi virus 86*	2e-09
TT-N-S01-0976-W	*Rhodopsin-like receptor/ structural constituent of cell wall*	*Arabidopsis thaliana*	3e-07
TT-N-S01-0312-W	*Stromal cell derived factor 2*	*Mus musculus*	1e-60
TT-N-S01-0855-W	*Structural constituent of cell wall*	*Arabidopsis thaliana*	2e-14
TT-N-S01-0032-W	*Transmembrane protein 93 isoform 1*	*Pan troglodytes*	2e-33
TT-N-S01-0893-W	*Tubulin alpha-1 chain (Alpha-I tubulin)*	*Homarus americanus*	1e-86
TT-N-S01-0107-W	*Tubulin alpha-3 chain*	*Homarus americanus*	5e-29
TT-N-S01-0051-W	*Tubulin beta-1 chain*	*Homarus americanus*	1e-108
TT-N-S01-0647-W			1e-103
TT-N-S01-0328-W	*Tubulin gamma chain - African clawed frog*	*Gallus gallus*	6e-43
TT-N-S01-0112-W	*Tubulin-specific chaperone e*	*Aedes aegypti*	1e-44

Table A5 Examples of transcripts from a testis cDNA library categorized as members of metabolism (60 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0483-W	*26S protease regulatory subunit*	*Aedes aegypti*	1e-119
TT-N-S01-0283-W	*26S proteasome regulatory complex subunit p48B*	*Drosophila melanogaster*	2e-50
TT-N-S01-0638-W	*26S proteasome subunit P45 family protein*	*Tetrahymena thermophila SB210*	5e-67
TT-N-S01-0447-W	*Acyl-CoA oxidase (ACX3)*	*Tetrahymena thermophila SB210*	7e-66
TT-N-S01-0247-W	*Alpha-amylase*	*Corbicula fluminea*	4e-43
TT-N-S01-0696-W	*Ancient ubiquitous protein 1*	*Tribolium castaneum*	2e-13
TT-N-S01-0322-W	*ATP lipid-binding protein like protein*	*Marsupenaeus japonicus*	6e-59
TT-N-S01-0632-W			9e-58
TT-N-S01-0213-W	*ATP synthase F0 subunit 6*	*Penaeus monodon*	1e-72
TT-N-S01-0238-W			8e-86
TT-N-S01-0261-W			1e-89
TT-N-S01-0603-W			3e-97
TT-N-S01-0986-W	*ATPase inhibitor-like protein*	*Bombyx mori*	5e-22
TT-N-S01-0008-W	*Carbonyl reductase 1-like*	*Tribolium castaneum*	2e-68
TT-N-S01-0676-W	*Casein kinase II beta subunit*	*Apis mellifera*	1e-116
TT-N-S01-0188-W	*Casein kinase II, alpha 1 polypeptide*	*Tribolium castaneum*	5e-67
TT-N-S01-0534-W	*Coproporphirynogen oxidase*	*Aplysia californica*	3e-05
TT-N-S01-0217-W	*Cystathionine gamma-lyase*	*Rattus norvegicus*	1e-52
TT-N-S01-0933-W	*Der1-like domain family member 1 (Degradation in endoplasmic reticulum protein 1, DER1)*	*Bombyx mori*	1e-59
TT-N-S01-0589-W	*Endoplasmic reticulum resident protein (ERp44) (Thioredoxin domain-containing protein 4)*	*Aedes aegypti*	1e-49
TT-N-S01-0231-W	*GTP binding protein*	*Aedes aegypti*	8e-72
TT-N-S01-0572-W	*Long chain acyl-CoA synthetase*	*Oryza sativa (japonica cultivar-group)*	1e-33
TT-N-S01-0709-W	*Lysyl oxidase-like 2 CG4402-PA*	*Apis mellifera*	8e-57
TT-N-S01-0094-W	*Mitochondrial ATP synthase gamma-subunit*	*Graphocephala atropunctata*	6e-62
TT-N-S01-0730-W	*Multiple inositol polyphosphate phosphatase*	*Aedes aegypti*	8e-22
TT-N-S01-0004-W	*Multiple inositol polyphosphate phosphatase 2*	*Apis mellifera*	2e-09
TT-N-S01-0495-W		*Drosophila melanogaster*	8e-09
TT-N-S01-0678-W		*Drosophila melanogaster*	1e-07
TT-N-S01-0248-W	*NADH:ubiquinone oxidoreductase NDUFS6 13 kDa subunit*	*Aedes aegypti*	1e-23
TT-N-S01-0791-W	*Nucleoside diphosphate kinase*	*Graphocephala atropunctata*	2e-09
TT-N-S01-0140-W	*Phenylalanine ammonia lyase*	*Rhodotorula glutinis*	2e-06
TT-N-S01-0178-W	*Phosphoserine aminotransferase 1 isoform 2*	*Strongylocentrotus purpuratus*	8e-48
TT-N-S01-1067-W	*Proteasome (prosome, macropain) 26S subunit, non-ATPase, 13*	*Apis mellifera*	6e-34
TT-N-S01-0554-W	*Proteasome (prosome, macropain) subunit, alpha type, 1, isoform CRA_a*	*Homo sapiens*	3e-07
TT-N-S01-0287-W	*Proteasome alpha 3 subunit*	*Bombyx mori*	2e-73
TT-N-S01-0385-W	*26S Proteasome non-ATPase regulatory subunit 3 (Diphenol oxidase A2 component)*	*Apis mellifera*	6e-69

Table A5 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0405-W	*Proteasome subunit alpha type*	*Aedes aegypti*	8e-99
TT-N-S01-0812-W			3e-65
TT-N-S01-0949-W			2e-80
TT-N-S01-0063-W	*Proteasome subunit alpha type 1 (Proteasome component C2) (Macropain subunit C2) (Multicatalytic endopeptidase complex subunit C2)*	*Canis familiaris*	6e-06
TT-N-S01-0369-W	*Proteasome subunit beta type 1 (Proteasome 26 kDa subunit)*	*Tribolium castaneum*	5e-26
TT-N-S01-0879-W	*Proteasome subunit, alpha type, 5*	*Apis mellifera*	1e-20
TT-N-S01-0863-W	*Protein serine/threonine kinase*	*Dictyostelium discoideum AX4*	3e-22
TT-N-S01-0118-W	*Receptor for activated protein kinase C-like*	*Blattella germanica*	1e-122
TT-N-S01-1021-W	*Rieske iron-sulfur protein 1*	*Graphocephala atropunctata*	3e-51
TT-N-S01-1009-W	*ring finger protein 20 isoform 3*	*Macaca mulatta*	1e-26
TT-N-S01-0905-W	*ring finger protein 44 isoform 3*	*Pan troglodytes*	7e-44
TT-N-S01-0755-W	*RNA binding motif protein 25 isoform 11*	*Bos taurus*	5e-08
TT-N-S01-0756-W	*RNA binding motif protein 25 isoform 11*	*Bos taurus*	4e-08
TT-N-S01-0065-W	*Selenoprotein M precursor*	*Danio rerio*	2e-16
TT-N-S01-0988-W	*Serine dehydratase-like*	*Xenopus tropicalis*	9e-20
TT-N-S01-0150-W	*Serine palmitoyl transferase LCB2 subunit*	*Drosophila melanogaster*	4e-34
TT-N-S01-0903-W	*Serine/threonine-protein kinase 23 (Muscle-specific serine kinase 1) (MSSK-1)*	*Apis mellifera*	4e-90
TT-N-S01-0281-W	*Succinate dehydrogenase complex, subunit C precursor*	*Danio rerio*	2e-28
TT-N-S01-0720-W	*THO complex 3*	*Danio rerio*	1e-36
TT-N-S01-0485-W	*Tudor domain containing 9*	*Bos taurus*	4e-07
TT-N-S01-0724-W	*Tyrosine protein kinase*	*Aedes aegypti*	2e-12
TT-N-S01-0492-W	*Ubiquinol-cytochrome c reductase core protein II*	*Strongylocentrotus purpuratus*	2e-36
TT-N-S01-0972-W	*Ubiquitin carboxyl-terminal hydrolase 5 (Ubiquitin thiolesterase 5) (Ubiquitin-specific-processing protease 5) (Deubiquitinating enzyme 5) (Isopeptidase T) isoform 2*	*Rattus norvegicus*	4e-65
TT-N-S01-0872-W	*Ubiquitin-conjugating enzyme E2*	*Aedes aegypti*	1e-58

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A6 Examples of transcripts from a testis cDNA library categorized as members of miscellaneous function (56 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0373-W	*2-Cys thioredoxin peroxidase*	*Aedes aegypti*	1e-35
TT-N-S01-0813-W			5e-63
TT-N-S01-0043-W	*ADP-Ribosylation Factor related family member (arf-3), Adp-ribosylation factor related protein 3*	*Caenorhabditis elegans*	2e-87
TT-N-S01-0373-W	*2-Cys thioredoxin peroxidase*	*Aedes aegypti*	1e-35
TT-N-S01-0813-W			5e-63
TT-N-S01-0043-W	*ADP-Ribosylation Factor related family member (arf-3), Adp-ribosylation factor related protein 3*	*Caenorhabditis elegans*	2e-87
TT-N-S01-0375-W	*Alpha-2-macroglobulin*	*Penaeus monodon*	8e-19
TT-N-S01-0438-W			7e-32

Table A6 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0332-W	*Alpha-endosulfine*	*Aedes aegypti*	1e-21
TT-N-S01-0162-W	*Alveolar soft part sarcoma chromosome region, candidate 1 long isoform isoform 4*	*Rattus norvegicus*	4e-29
TT-N-S01-0900-W	*Bromodomain containing 8 (Skeletal muscle abundant protein, SMAP, SMAP2, Thyroid hormone receptor coactivating protein 120kDa, TrCP120)*	*Apis mellifera*	3e-36
TT-N-S01-0849-W	*Calcitonin gene-related peptide-receptor component protein isoform a*	*Homo sapiens*	9e-27
TT-N-S01-0715-W	*calmodulin regulated spectrin-associated protein 1*	*Rattus norvegicus*	1e-21
TT-N-S01-0464-W	*Cement precursor protein 3B variant 2*	*Phragmatopoma californica*	9e-10
TT-N-S01-0975-W			2e-05
TT-N-S01-1056-W	*Chromosome 20 open reading frame 11*	*Homo sapiens*	3e-34
TT-N-S01-1058-W			3e-34
TT-N-S01-0880-W	*dynein light intermediate chain*	*Aedes aegypti*	6e-80
TT-N-S01-0182-W	*Endoplasmin (Heat shock protein 90 kDa beta member 1) (94 kDa glucose-regulated protein) (GRP94)*	*Mesocricetus auratus*	1e-60
TT-N-S01-0694-W	*GRN protein*	*Xenopus tropicalis*	4e-41
TT-N-S01-0320-W	*pregnancy-related serine protease HTRA3*	*Homo sapiens*	9e-10
TT-N-S01-0489-W			4e-08
TT-N-S01-0782-W			8e-10
TT-N-S01-0953-W			1e-07
TT-N-S01-0498-W	*Huntingtin interacting protein K, partial*	*Apis mellifera*	4e-27
TT-N-S01-0060-W	*Importin 7*	*Aedes aegypti*	1e-67
TT-N-S01-0074-W	*Leucine-rich repeat flightless-interacting protein 2*	*Xenopus laevis*	2e-13
TT-N-S01-0991-W	*metallothionein*	*Homarus americanus*	3e-29
TT-N-S01-0026-W	*Methyltransferase WBSCR22 (Williams-Beuren syndrome chromosome region 22 protein homolog)*	*Mus musculus*	2e-61
TT-N-S01-0214-W	*Mucin-like protein*	*Trypanosoma cruzi*	1e-18
TT-N-S01-0025-W	*Myosin 61F CG9155-PB, isoform B isoform 1*	*Apis mellifera*	6e-47
TT-N-S01-0637-W	*Nascent polypeptide associated complex protein alpha subunit CG8759-PB, isoform B isoform 1*	*Apis mellifera*	2e-42
TT-N-S01-0682-W	*oligonucleotide/oligosaccharide-binding fold containing 1(OBFC1)*	*Mus musculus*	1e-08
TT-N-S01-0901-W	*P. falciparum RESA-like protein with DnaJ domain*	*Plasmodium falciparum 3D7*	2e-18
TT-N-S01-0460-W	*Penaeus monodon clone TUZX4-6:86 microsatellite sequence*	*Penaeus monodon*	2e-71
TT-N-S01-0115-W	*Protein-glutamine gamma-glutamyltransferase K (Transglutaminase K) (TGase K)*	*Tribolium castaneum*	1e-68
TT-N-S01-0996-W	*Putative accessory gland protein*	*Gryllus bimaculatus*	7e-33
TT-N-S01-0957-W	*RAB, member of RAS oncogene family-like 3*	*Apis mellifera*	7e-64
TT-N-S01-1036-W	*Rac GTPase activating protein 1 isoform 1*	*Canis familiaris*	2e-27
TT-N-S01-1040-W			4e-28
TT-N-S01-1048-W			2e-23
TT-N-S01-0700-W	*RWD domain containing 4A*	*Macaca mulatta*	4e-42
TT-N-S01-0018-W	*Salivary gland secretion 1 CG3047-PA*	*Drosophila melanogaster*	5e-12
TT-N-S01-0956-W			7e-14

Table A6 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0843-W	*Serine/threonine-protein phosphatase 2A catalytic subunit beta isoform (PP2A-beta)*	*Tribolium castaneum*	1e-109
TT-N-S01-0636-W	*source of immunodominant MHC-associated peptides (Oligosaccharyl transferase subunit STT3B)*	*Ctenopharyngodon idella*	4e-75
TT-N-S01-0144-W	*SUMO, small ubiquitin-like modifier, SUMO, small ubiquitin-like modifier SMO-1 (10.2 kD) (smo-1)*	*Apis mellifera*	5e-38
TT-N-S01-0175-W			8e-38
TT-N-S01-0426-W			5e-38
TT-N-S01-0626-W	*Synaptobrevin-like protein 1*	*Canis familiaris*	1e-48
TT-N-S01-0414-W	*T-complex protein 1, alpha subunit(TCP-1-alpha)(CCT-alpha)*	*Delia antiqua*	5e-73
TT-N-S01-0224-W	*Tetraspanin 96F CG6120-PA*	*Drosophila melanogaster*	8e-41
TT-N-S01-0702-W	*Tetratricopeptide repeat domain 9C (TTC9)*	*Rattus norvegicus*	3e-23
TT-N-S01-0897-W	*Thioredoxin-2 CG31884-PA, isoform A*	*Apis mellifera*	2e-34
TT-N-S01-0232-W	*Thyroid hormone receptor-associated protein complex 240 kDa component (Trap240) (Thyroid hormone receptor associated protein 1)*	*Canis familiaris*	2e-57
TT-N-S01-0407-W	*TPA_inf: troponin I isoform a2*	*Drosophila pseudoobscura*	1e-26
TT-N-S01-0223-W	*Ubiquitin carboxyl-terminal hydrolase 14 (Ubiquitin thiolesterase 14) (Ubiquitin-specific processing protease 14) (Deubiquitinating enzyme 14)*	*Apis mellifera*	2e-51
TT-N-S01-0120-W	*WD-repeat protein 43*	*Canis familiaris*	5e-15
TT-N-S01-0902-W	*Zinc binding dehydrogenase*	*Aedes aegypti*	2e-27

Table A7 Examples of transcripts from a testis cDNA library categorized as members of mitochondrial protein (25 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0061-W	*Cytochrome b*	*Penaeus monodon*	1e-101
TT-N-S01-0473-W			1e-102
TT-N-S01-0983-W			1e-117
TT-N-S01-0211-W	*Cytochrome oxidase subunit I*	*Heterocarpus laevigatus*	1e-112
TT-N-S01-0466-W	*Cytochrome c oxidase subunit I*	*Marsupenaeus japonicus*	9e-44
TT-N-S01-0627-W		*Fenneropenaeus merguiensis*	2e-84
TT-N-S01-0797-W		*Marsupenaeus japonicus*	8e-76
TT-N-S01-0833-W		*Marsupenaeus japonicus*	1e-17
TT-N-S01-0877-W		*Fenneropenaeus merguiensis*	1e-113
TT-N-S01-1013-W		*Marsupenaeus japonicus*	1e-119
TT-N-S01-1018-W		*Marsupenaeus japonicus*	1e-116
TT-N-S01-0316-W	*Cytochrome c oxidase subunit II*	*Penaeus monodon*	2e-89
TT-N-S01-0592-W			1e-113
TT-N-S01-0769-W	*Cytochrome c oxidase subunit III*	*Scutigerella causeyae*	1e-28
TT-N-S01-0770-W		*Tricholepidion gertschi*	6e-21
TT-N-S01-1043-W		*Penaeus monodon*	1e-75
TT-N-S01-1050-W		*Priapulus caudatus*	6e-05
TT-N-S01-0243-W	*Histidine triad family zinc-binding protein, protein kinase C inhibitor*	*Aedes aegypti*	3e-42
TT-N-S01-0040-W	*Mitochondrial NADH dehydrogenase (ubiquinone) 1 alpha subcomplex*	*Aedes aegypti*	7e-28

Table A7 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0954-W	*Mitochondrial tumor suppressor 1 isoform 5 (MTUS1)*	*Apis mellifera*	1e-09
TT-N-S01-0323-W	*NADH dehydrogenase subunit 5*	*Penaeus monodon*	1e-102
TT-N-S01-0823-W			1e-62
TT-N-S01-0201-W	*NADH dehydrogenase subunit 6*	*Penaeus monodon*	1e-52
TT-N-S01-0048-W	*NADH-ubiquinone oxidoreductase fe-s protein 2 (ndufs2)*	*Aedes aegypti*	2e-92
TT-N-S01-0062-W	*ND3 (NADH dehydrogenase subunit 3)*	*Clibanarius albidigitus*	2e-07

Table A8 Examples of transcripts from a testis cDNA library categorized as members of transport (28 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0177-W	*Adaptor-related protein complex 1, beta 1 subunit, isoform CRA_c*	*Homo sapiens*	2e-38
TT-N-S01-0735-W	*ADP ribosylation factor 79F*	*Argas monolakensis*	7e-86
TT-N-S01-0458-W	*ADP-ribosylation-like factor 6 interacting protein 5*	*Tribolium castaneum*	4e-57
TT-N-S01-0926-W	*Cation efflux protein/ zinc transporter*	*Aedes aegypti*	2e-61
TT-N-S01-0467-W	*Cytochrome B561*	*Aedes aegypti*	7e-41
TT-N-S01-0444-W	*E1B-55kDa-associated protein 5 isoform 5*	*Pan troglodytes*	1e-51
TT-N-S01-0499-W	*Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 9*	*synthetic construct*	3e-55
TT-N-S01-1001-W	*karyopherin (importin) alpha 2*	*Ictalurus punctatus*	3e-65
TT-N-S01-0396-W	*Kinesin light chain 1 and*	*Aedes aegypti*	5e-17
TT-N-S01-0134-W	*Laminin beta chain*	*Schistocerca gregaria*	1e-43
TT-N-S01-0011-W	*Lipocalin-1 interacting membrane receptor (limr)*	*Aedes aegypti*	2e-24
TT-N-S01-0437-W	*Metaxin 2*	*Tribolium castaneum*	1e-79
TT-N-S01-0862-W	*NADH dehydrogenase subunit 1*	*Penaeus monodon*	1e-117
TT-N-S01-0325-W	*Protein transport protein SEC61 gamma subunit*	*Rattus norvegicus*	8e-26
TT-N-S01-0850-W			3e-26
TT-N-S01-0711-W	*Putative Na+/K+-ATPase alpha subunit*	*Homarus americanus*	1e-112
TT-N-S01-0698-W	*Serologically defined breast cancer antigen 84*	*Danio rerio*	5e-66
TT-N-S01-0699-W			5e-66
TT-N-S01-0390-W	*Solute carrier family 2 (facilitated glucose transporter), member 13*	*Strongylocentrotus purpuratus*	1e-50
TT-N-S01-0449-W	*Translationally controlled tumor protein*	*Penaeus monodon*	3e-91
TT-N-S01-0748-W			1e-81
TT-N-S01-1062-W			2e-91
TT-N-S01-0324-W	*Translocase of inner mitochondrial membrane*	*Artemia franciscana*	2e-44
TT-N-S01-0860-W			1e-51
TT-N-S01-0147-W	*Transposase*	*Escherichia coli*	1e-125
TT-N-S01-1004-W	*transposon protein, putative, CACTA, En/Spm sub-class*	*Oryza sativa (japonica cultivar-group)*	5e-07
TT-N-S01-0968-W	*Vacuolar ATP synthase 21 kDa proteolipid subunit*	*Bombyx mori*	2e-66
TT-N-S01-0233-W	*Vacuolar ATPase G subunit-like protein*	*Graphocephala atropunctata*	3e-36

Table A9 Examples of transcripts from a testis cDNA library categorized as members of ribosomal and rRNA (150 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0876-W	*18S ribosomal RNA*	*Litopenaeus vannamei*	1e-166
TT-N-S01-0980-W			1e-173
TT-N-S01-1086-W			1e-175
TT-N-S01-0017-W	*28S ribosomal protein S16, mitochondrial*	*Aedes aegypti*	7e-08
TT-N-S01-0898-W	*40S ribosomal protein*	*Perinereis aibuhitensis*	2e-54
TT-N-S01-0348-W	*40S ribosomal protein S13*	*Ictalurus punctatus*	2e-70
TT-N-S01-0886-W	*40S ribosomal protein S15 (RIG protein) isoform 1*	*Macaca mulatta*	3e-65
TT-N-S01-0999-W			2e-65
TT-N-S01-0559-W	*40S ribosomal protein S16*	*Rattus norvegicus*	1e-66
TT-N-S01-0267-W	*40S ribosomal protein S18*	*Spodoptera frugiperda*	9e-27
TT-N-S01-0110-W	*40S ribosomal protein S2*	*Urechis caupo*	8e-96
TT-N-S01-0708-W	*40S ribosomal protein S23*	*Argas monolakensis*	1e-75
TT-N-S01-0811-W			3e-74
TT-N-S01-0931-W			4e-76
TT-N-S01-0951-W			4e-76
TT-N-S01-0564-W	*40S ribosomal protein S25*	*Aedes aegypti*	4e-36
TT-N-S01-0091-W	*40S ribosomal protein S27*	*Homarus americanus*	1e-43
TT-N-S01-0516-W			1e-43
TT-N-S01-0732-W	*40S ribosomal protein S3a (C3 protein)*	*Aedes aegypti*	2e-89
TT-N-S01-0321-W		*Tribolium castaneum*	3e-96
TT-N-S01-0668-W		*Tribolium castaneum*	7e-99
TT-N-S01-0837-W		*Apis mellifera*	4e-66
TT-N-S01-0432-W	*40S ribosomal protein S5*	*Ornithodoros moubata*	9e-49
TT-N-S01-0921-W	*40S ribosomal protein S6*	*Spodoptera frugiperda*	1e-48
TT-N-S01-0314-W	*40S ribosomal protein S7*	*Ictalurus punctatus*	4e-71
TT-N-S01-0398-W			4e-71
TT-N-S01-0487-W			1e-70
TT-N-S01-0195-W	*40S ribosomal protein Sa*	*Ictalurus punctatus*	6e-83
TT-N-S01-0989-W			8e-49
TT-N-S01-0650-W	*60S acidic ribosomal protein P2*	*Strongylocentrotus purpuratus*	2e-31
TT-N-S01-1029-W			1e-31
TT-N-S01-0161-W	*60S ribosomal protein L13*	*Apis mellifera*	1e-40
TT-N-S01-0852-W			3e-61
TT-N-S01-0174-W	*60S ribosomal protein L28*	*Spodoptera frugiperda*	1e-31
TT-N-S01-0355-W			3e-32
TT-N-S01-0793-W			3e-31
TT-N-S01-1002-W	*60S ribosomal protein L7A*	*Ixodes pacificus*	1e-74
TT-N-S01-0777-W	*Acidic p0 ribosomal protein*	*Dascillus cervinus*	2e-60
TT-N-S01-0807-W	*Acidic ribosomal protein P1*	*Aedes aegypti*	3e-10
TT-N-S01-0302-W	*Ribosomal protein 31*	*Lonomia obliqua*	2e-69
TT-N-S01-0482-W			2e-61
TT-N-S01-0517-W			2e-66
TT-N-S01-0022-W	*Ribosomal protein L10*	*Callinectes sapidus*	1e-106
TT-N-S01-0335-W	*Ribosomal protein L13a*	*Lysiphlebus testaceipes*	1e-45
TT-N-S01-0561-W	*Ribosomal protein L14*	*Lysiphlebus testaceipes*	1e-33
TT-N-S01-0493-W	*Ribosomal protein L15e*	*Timarcha balearica*	2e-91
TT-N-S01-0560-W	*Ribosomal protein L17e*	*Agriotes lineatus*	1e-44
TT-N-S01-0946-W	*Ribosomal protein L18a variant*	*Homo sapiens*	1e-64
TT-N-S01-0406-W	*Ribosomal protein L21*	*Danio rerio*	8e-56
TT-N-S01-0446-W		*Drosophila melanogaster*	1e-14
TT-N-S01-0584-W	*Ribosomal protein L22*	*Drosophila melanogaster*	3e-38

Table A9 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0171-W	*Ribosomal protein L23*	*Homo sapiens*	1e-67
TT-N-S01-0219-W	*Ribosomal protein L23Ae*	*Cicindela campestris*	5e-09
TT-N-S01-0056-W	*Ribosomal protein L24*	*Marsupenaeus japonicus*	2e-67
TT-N-S01-0240-W		*Bombyx mori*	3e-48
TT-N-S01-0428-W		*Marsupenaeus japonicus*	2e-68
TT-N-S01-0840-W		*Marsupenaeus japonicus*	2e-67
TT-N-S01-0059-W	*Ribosomal protein L26*	*Penaeus monodon*	2e-41
TT-N-S01-0378-W	*Ribosomal protein L27*	*Suberites domuncula*	1e-39
TT-N-S01-0789-W	*Ribosomal protein L28*	*Canis familiaris*	1e-18
TT-N-S01-0237-W	*Ribosomal protein L3*	*Bombyx mori*	5e-92
TT-N-S01-0254-W			4e-57
TT-N-S01-0948-W			1e-108
TT-N-S01-0349-W		*Spodoptera frugiperda*	1e-41
TT-N-S01-0994-W	*Ribosomal protein L30*	*Argopecten irradians*	2e-47
TT-N-S01-0518-W	*Ribosomal protein L31*	*Tribolium castaneum*	1e-41
TT-N-S01-0998-W			2e-42
TT-N-S01-0202-W	*Ribosomal protein L32e*	*Hister sp. APV-2005*	2e-23
TT-N-S01-0431-W			2e-57
TT-N-S01-1011-W			1e-57
TT-N-S01-0613-W	*Ribosomal protein L36*	*Lysiphlebus testaceipes*	7e-14
TT-N-S01-1088-W	*Ribosomal protein L39*	*Ixodes scapularis*	7e-20
TT-N-S01-0384-W	*Ribosomal protein L4*	*Apis mellifera*	5e-81
TT-N-S01-0844-W		*Bombyx mori*	4e-65
TT-N-S01-1065-W		*Apis mellifera*	5e-88
TT-N-S01-0910-W	*Ribosomal protein L44e*	*Hister sp. APV-2005*	2e-46
TT-N-S01-0229-W	*Ribosomal protein L5 variant*	*Homo sapiens*	6e-92
TT-N-S01-0286-W			1e-89
TT-N-S01-0490-W			5e-72
TT-N-S01-0726-W	*Ribosomal protein L5e*	*Carabus granulatus*	5e-54
TT-N-S01-0080-W	*Ribosomal protein L7 isoform A*	*Lysiphlebus testaceipes*	1e-73
TT-N-S01-0357-W	*Ribosomal protein L8*	*Litopenaeus vannamei*	1e-122
TT-N-S01-0993-W			1e-135
TT-N-S01-1073-W			1e-130
TT-N-S01-1084-W			1e-129
TT-N-S01-0753-W	*Ribosomal protein L9*	*Culicoides sonorensis*	3e-65
TT-N-S01-0054-W	*Ribosomal protein LP1*	*Argas monolakensis*	3e-38
TT-N-S01-0215-W			5e-38
TT-N-S01-0336-W			3e-38
TT-N-S01-0718-W			9e-34
TT-N-S01-0848-W			3e-38
TT-N-S01-0207-W	*Ribosomal protein P0*	*Lysiphlebus testaceipes*	2e-10
TT-N-S01-0361-W	*Ribosomal protein P2*	*Branchiostoma belcheri*	3e-28
TT-N-S01-0716-W			6e-28
TT-N-S01-0191-W	*Ribosomal protein S10*	*Mus musculus*	4e-50
TT-N-S01-0764-W			3e-53
TT-N-S01-0152-W	*Ribosomal protein S11*	*Apis mellifera*	5e-65
TT-N-S01-0642-W	*Ribosomal protein S18*	*Cherax destructor*	1e-60
TT-N-S01-0719-W			5e-74
TT-N-S01-0117-W	*Ribosomal protein S19e*	*Eucinetus sp. APV-2005*	3e-54
TT-N-S01-0391-W	*Ribosomal protein S21*	*Branchiostoma belcheri*	1e-31
TT-N-S01-0723-W	*Ribosomal protein S24*	*Marsupenaeus japonicus*	1e-66

Table A9 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0541-W	*Ribosomal protein S26*	*Acyrthosiphon pisum*	7e-44
TT-N-S01-0687-W		*Branchiostoma belcheri*	9e-46
TT-N-S01-0583-W	*Ribosomal protein S30*	*Crassostrea gigas*	1e-16
TT-N-S01-0155-W	*Ribosomal protein S4*	*Gallus gallus*	5e-78
TT-N-S01-0562-W	*Ribosomal protein S5*	*Strongylocentrotus purpuratus*	2e-77
TT-N-S01-0350-W	*Ribosomal protein S8e*	*Georissus sp. APV-2005*	4e-95
TT-N-S01-0377-W			2e-96
TT-N-S01-0826-W			1e-33
TT-N-S01-0176-W	*Ribosomal protein S9 CG3395-PA, isoform A*	*Apis mellifera*	2e-52
TT-N-S01-0411-W			2e-84
TT-N-S01-0806-W	*S10e ribosomal protein*	*Carabus granulatus*	2e-25
TT-N-S01-0016-W	*S5e ribosomal protein*	*Dascillus cervinus*	6e-89
TT-N-S01-0037-W			1e-62
TT-N-S01-0173-W			8e-91
TT-N-S01-0354-W			3e-97
TT-N-S01-0768-W			1e-47
TT-N-S01-0854-W			3e-97
TT-N-S01-0925-W			3e-97
TT-N-S01-0157-W	*tRNA-Ile (16S ribosomal RNA gene)*	*Penaeus monodon*	0
TT-N-S01-0227-W			0
TT-N-S01-0265-W			0
TT-N-S01-0366-W			0
TT-N-S01-0388-W			0
TT-N-S01-0393-W			0
TT-N-S01-0394-W			0
TT-N-S01-0399-W			0
TT-N-S01-0486-W			0
TT-N-S01-0532-W			0
TT-N-S01-0590-W			0
TT-N-S01-0631-W			1e-92
TT-N-S01-0635-W			0
TT-N-S01-0660-W			0
TT-N-S01-0661-W			0
TT-N-S01-0681-W			0
TT-N-S01-0689-W			0
TT-N-S01-0802-W			1e-114
TT-N-S01-0808-W			1e-160
TT-N-S01-0839-W			0
TT-N-S01-0955-W			0
TT-N-S01-0978-W			0
TT-N-S01-0984-W			0
TT-N-S01-0015-W			0
TT-N-S01-0341-W	*Ubiquitin/ribosomal protein S27a fusion protein*	*Branchiostoma belcheri tsingtaunese*	5e-79
TT-N-S01-0041-W	*Ubiquitin/ribosomal protein S30e fusion protein*	*Carabus granulatus*	1e-33
TT-N-S01-0402-W		*Hister sp. APV-2005*	3e-35
TT-N-S01-0470-W		*Sphaerius sp. APV-2005*	2e-29
TT-N-S01-0995-W		*Hister sp. APV-2005 Hister sp. APV-2005*	9e-36
TT-N-S01-1037-W			2e-35
TT-N-S01-0727-W	*Ubuiquitin/ribosomal L40 fusion protein*	*Scleronephthya gracillimum*	4e-64

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

Table A10 Examples of transcripts from a testis cDNA library categorized as members of unidentified (hypothetical)-similar to other cDNA/DNA) (101 clones)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0372-W	*Predicted protein*	*Phaeosphaeria nodorum SN15*	3e-10
TT-N-S01-0249-W	*CG10109-PA*	*Tribolium castaneum*	5e-06
TT-N-S01-0881-W	*CG10600-PA*	*Tribolium castaneum*	3e-21
TT-N-S01-0666-W	*CG12279-PA*	*Rattus norvegicus*	6e-16
TT-N-S01-0810-W	*CG12301-PA*	*Tribolium castaneum*	5e-05
TT-N-S01-0913-W	*CG12659-PB*	*Tribolium castaneum*	7e-30
TT-N-S01-0109-W	*CG12859*	*Drosophila yakuba*	9e-22
TT-N-S01-0677-W	*CG13220-PA*	*Tribolium castaneum*	2e-20
TT-N-S01-0456-W	*CG13363-PA*	*Tribolium castaneum*	7e-14
TT-N-S01-0488-W	*CG13623-PA*	*Apis mellifera*	2e-38
TT-N-S01-0046-W	*CG14073-PA, isoform A*	*Apis mellifera*	4e-10
TT-N-S01-0717-W	*CG14865-PA*	*Tribolium castaneum*	6e-25
TT-N-S01-0743-W	*CG15432-PA*	*Tribolium castaneum*	3e-16
TT-N-S01-0268-W	*CG15626-PA, isoform A*	*Apis mellifera*	2e-42
TT-N-S01-0244-W	*CG17068-PA*	*Apis mellifera*	6e-07
TT-N-S01-0102-W	*CG18542-PA*	*Apis mellifera*	7e-27
TT-N-S01-1032-W	*CG33691-PB, isoform B*	*Apis mellifera*	9e-12
TT-N-S01-0542-W	*CG3654-PD*	*Apis mellifera*	4e-27
TT-N-S01-0906-W	*CG3773-PA*	*Apis mellifera*	8e-15
TT-N-S01-0598-W	*CG8677-PA*	*Apis mellifera*	6e-18
TT-N-S01-0958-W	*ENSANGP00000014082*	*Anopheles gambiae str. PEST*	5e-14
TT-N-S01-0100-W	*ENSANGP00000015829*	*Anopheles gambiae str. PEST*	6e-09
TT-N-S01-0462-W	*ENSANGP00000020130*	*Anopheles gambiae str. PEST*	3e-09
TT-N-S01-0165-W	*ENSANGP00000020267*	*Apis mellifera*	3e-11
TT-N-S01-0597-W	*ENSANGP00000030087*	*Anopheles gambiae str. PEST*	8e-05
TT-N-S01-0547-W	*Epa4p*	*Candida glabrata*	3e-10
TT-N-S01-0421-W	*Es2 CG1474-PA*	*Apis mellifera*	3e-23
TT-N-S01-1077-W	*hCG1793893*	*Homo sapiens*	8e-07
TT-N-S01-0006-W	*Hypothetical protein*	*Dictyostelium discoideum AX4*	2e-05
TT-N-S01-0050-W	*Hypothetical protein*	*Rattus norvegicus*	2e-07
TT-N-S01-0064-W	*Hypothetical protein*	*Plasmodium falciparum 3D7*	6e-06
TT-N-S01-0078-W	*Hypothetical protein*	*Plasmodium yoelii yoelii str. 17XNL*	6e-21
TT-N-S01-0088-W	*Hypothetical protein*	*Chaetomium globosum CBS 148.51*	2e-14
TT-N-S01-0095-W	*Hypothetical protein*	*Strongylocentrotus purpuratus*	2e-40
TT-N-S01-0099-W	*Hypothetical protein*	*Delftia acidovorans SPH-1*	2e-09
TT-N-S01-0104-W	*Hypothetical protein*	*Aedes aegypti*	4e-11
TT-N-S01-0141-W	*Hypothetical protein*	*Tribolium castaneum*	5e-05
TT-N-S01-0156-W	*Hypothetical protein*	*Rattus norvegicus*	4e-05
TT-N-S01-0218-W	*Hypothetical protein*	*Gallus gallus*	1e-06
TT-N-S01-0256-W	*Hypothetical protein*	*Tetrahymena thermophila SB210*	1e-12
TT-N-S01-0343-W	*Hypothetical protein*	*Gallus gallus*	8e-07
TT-N-S01-0345-W	*Hypothetical protein*	*Tetrahymena thermophila SB210*	1e-13
TT-N-S01-0454-W	*Hypothetical protein*	*Cryptococcus neoformans var. neoformans JEC21*	8e-05
TT-N-S01-0465-W	*Hypothetical protein*	*Dictyostelium discoideum AX4*	2e-15
TT-N-S01-0511-W	*Hypothetical protein*	*Dictyostelium discoideum AX4*	6e-09
TT-N-S01-0521-W	*Hypothetical protein*	*Dictyostelium discoideum AX4*	6e-06
TT-N-S01-0654-W	*Hypothetical protein*	*Geobacter uraniumreducens Rf4*	5e-21

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A10 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0686-W	*Hypothetical protein*	*Rattus norvegicus*	3e-12
TT-N-S01-0692-W	*Hypothetical protein*	*Aedes aegypti*	5e-05
TT-N-S01-0706-W	*Hypothetical protein*	*Tetrahymena thermophila SB210*	2e-07
TT-N-S01-0713-W	*Hypothetical protein*	*Bacteroides fragilis YCH46*	2e-08
TT-N-S01-0747-W	*Hypothetical protein*	*Strongylocentrotus purpuratus*	9e-06
TT-N-S01-0751-W	*Hypothetical protein*	*Rattus norvegicus*	1e-21
TT-N-S01-0834-W	*Hypothetical protein*	*Gallus gallus*	2e-05
TT-N-S01-0878-W	*Hypothetical protein*	*Plasmodium yoelii yoelii str. 17XNL*	3e-12
TT-N-S01-0884-W	*Hypothetical protein*	*Rattus norvegicus*	3e-12
TT-N-S01-0889-W	*Hypothetical protein*	*Oryza sativa (japonica cultivar-group)*	8e-07
TT-N-S01-0908-W	*Hypothetical protein*	*Plasmodium falciparum 3D7*	8e-07
TT-N-S01-0928-W	*Hypothetical protein*	*Plasmodium berghei strain ANKA*	7e-22
TT-N-S01-0935-W	*Hypothetical protein*	*Rattus norvegicus*	3e-18
TT-N-S01-0945-W	*Hypothetical protein*	*Rattus norvegicus*	8e-11
TT-N-S01-0952-W	*Hypothetical protein*	*Plasmodium yoelii yoelii str. 17XNL*	3e-12
TT-N-S01-0964-W	*Hypothetical protein*	*Rattus norvegicus*	8e-11
TT-N-S01-0973-W	*Hypothetical protein*	*Homo sapiens*	1e-17
TT-N-S01-0997-W	*Hypothetical protein*	*Tetrahymena thermophila SB210*	2e-12
TT-N-S01-1041-W	*Hypothetical protein*	*Eimeria tenella str. Houghton*	1e-07
TT-N-S01-1046-W	*Hypothetical protein*	*Apis mellifera*	3e-10
TT-N-S01-1052-W	*Hypothetical protein*	*Apis mellifera*	2e-10
TT-N-S01-1076-W	*Hypothetical protein*	*Dictyostelium discoideum AX4*	9e-20
TT-N-S01-1078-W	*Hypothetical protein*	*Tetrahymena thermophila SB210*	5e-12
TT-N-S01-0296-W	*Hypothetical protein CBG02739*	*Caenorhabditis briggsae*	3e-06
TT-N-S01-0024-W	*Hypothetical protein FLJ20580*	*Gallus gallus*	7e-34
TT-N-S01-0305-W	*Hypothetical protein XP_584232 isoform 1*	*Bos taurus*	1e-10
TT-N-S01-0662-W	*Hypothetical protein Y50E8A.i*	*Caenorhabditis elegans*	4e-06
TT-N-S01-0864-W	*Hypothetical protein, conserved*	*Eimeria tenella str. Houghton*	8e-06
TT-N-S01-0890-W	*IP06461p*	*Drosophila melanogaster*	2e-22
TT-N-S01-0146-W	*LCI5*	*Chlamydomonas reinhardtii*	2e-07
TT-N-S01-0728-W	*MGC53864 protein*	*Xenopus laevis*	1e-06
TT-N-S01-0258-W	*MGC83791 protein*	*Xenopus laevis*	1e-07
TT-N-S01-1042-W	*putative ORF2*	*Drosophila melanogaster*	3e-07
TT-N-S01-0845-W	*RIKEN cDNA 2310061C15 gene isoform 1*	*Gallus gallus*	5e-18
TT-N-S01-0210-W	*SJCHGC01974 protein*	*Schistosoma japonicum*	1e-12
TT-N-S01-1024-W	*SJCHGC09076 protein*	*Schistosoma japonicum*	3e-18
TT-N-S01-0010-W	*Unnamed protein product*	*Homo sapiens*	2e-39
TT-N-S01-0888-W	*Unknown*	*Schistosoma japonicum*	5e-08
TT-N-S01-0979-W	*Unknown*	*Frog virus 3*	8e-10
TT-N-S01-1068-W	*Unknown*	*Schistosoma japonicum*	1e-15
TT-N-S01-0858-W	*Unknown (protein for IMAGE:2639329)*	*Danio rerio*	2e-07
TT-N-S01-0645-W	*Unknown protein*	*Arabidopsis thaliana*	4e-05
TT-N-S01-0116-W	*Unnamed protein product*	*Homo sapiens*	2e-05
TT-N-S01-0194-W	*Unnamed protein product*	*Saimiriine herpesvirus 2*	2e-12
TT-N-S01-0251-W	*Unnamed protein product*	*Tetraodon nigroviridis*	2e-06
TT-N-S01-0297-W	*Unnamed protein product*	*Kluyveromyces lactis*	6e-20
TT-N-S01-0379-W	*Unnamed protein product*	*Tetraodon nigroviridis*	2e-09

Table A10 (cont.)

Clone No.	Transcripts	Closest Species	E-value
TT-N-S01-0392-W	*Unnamed protein product*	*Kluyveromyces lactis*	1e-15
TT-N-S01-0665-W	*Unnamed protein product*	*Saimiriine herpesvirus 2*	2e-12
TT-N-S01-0788-W	*Unnamed protein product*	*Tetraodon nigroviridis*	1e-10
TT-N-S01-0804-W	*Unnamed protein product*	*Kluyveromyces lactis*	3e-06
TT-N-S01-0936-W	*Unnamed protein product*	*Tetraodon nigroviridis*	9e-21
TT-N-S01-1000-W	*Unnamed protein product*	*Tetraodon nigroviridis*	6e-17
TT-N-S01-0691-W	*Y43E12A.2*	*Tribolium castaneum*	5e-21

Table A11 Unknown transcripts from a testis cDNA library (290 clones)

GENE IDENTITY	No. of Clones	Clone
Unknown genes	290	TT 001, TT 002, TT 005, TT 007, TT 019, TT 021, TT 023, TT 028, TT 033, TT 036, TT 044, TT 047, TT 053, TT 057, TT 068, TT 069, TT 070, TT 073, TT 081, TT 086, TT 093, TT 097, TT 103, TT 113, TT 114, TT 137, TT 138, TT 142, TT 143, TT 149, TT 151, TT 158, TT 160, TT 166, TT 168, TT 170, TT 172, TT 179, TT 180, TT 181, TT 184, TT 193, TT 196, TT 197, TT 200, TT 204, TT 208, TT 209, TT 212, TT 225, TT 226, TT 228, TT 236, TT 242, TT 245, TT 250, TT 253, TT 255, TT 262, TT 263, TT 266, TT 270, TT 271, TT 273, TT 274, TT 275, TT 276, TT 279, TT 280, TT 282, TT 288, TT 291, TT 292, TT 293, TT 298, TT 299, TT 303, TT 304, TT 318, TT 327, TT 329, TT 330, TT 331, TT 344, TT 346, TT 347, TT 352, TT 363, TT 365, TT 367, TT 380, TT 381, TT 387, TT 395, TT 397, TT 400, TT 401, TT 403, TT 404, TT 408, TT 409, TT 410, TT 419, TT 420, TT 422, TT 423, TT 429, TT 434, TT 439, TT 440, TT 441, TT 442, TT 445, TT 451, TT 453, TT 455, TT 461, TT 463, TT 477, TT 478, TT 479, TT 480, TT 481, TT 491, TT 496, TT 497, TT 500, TT 501, TT 503, TT 507, TT 508, TT 509, TT 513, TT 514, TT 515, TT 528, TT 533, TT 543, TT 546, TT 551, TT 552, TT 558, TT 563, TT 568, TT 570, TT 575, TT 580, TT 581, TT 593, TT 599, TT 602, TT 605, TT 606, TT 612, TT 617, TT 623, TT 624, TT 625, TT 628, TT 634, TT 639, TT 640, TT 641, TT 643, TT 651, TT 653, TT 655, TT 656, TT 659, TT 664, TT 669, TT 670, TT 672, TT 680, TT 688, TT 690, TT 693, TT 697, TT 703, TT 705, TT 712, TT 714, TT 721, TT 722, TT 729, TT 731, TT 734, TT 736, TT 740, TT 744, TT 745, TT 749, TT 754, TT 758, TT 759, TT 761, TT 763, TT 766, TT 772, TT 774, TT 775, TT 778, TT 781, TT 790, TT 803, TT 805, TT 809, TT 816, TT 818, TT 819, TT 821, TT 822, TT 824, TT 825, TT 827, TT 829, TT 830, TT 836, TT 841, TT 842, TT 851, TT 853, TT 856, TT 857, TT 859, TT 861, TT 865, TT 866, TT 869, TT 875, TT 885, TT 887, TT 892, TT 896, TT 904, TT 909, TT 914, TT 918, TT 919, TT 920, TT 922, TT 923, TT 924, TT 927, TT 929, TT 932, TT 937, TT 938, TT 939, TT 940, TT 942, TT 943, TT 944, TT 950, TT 960, TT 962, TT 963, TT 965, TT 969, TT 970, TT 981, TT 987, TT 992, TT 1003, TT 1005, TT 1008, TT 1012, TT 1017, TT 1019, TT 1023, TT 1027, TT 1028, TT 1033, TT 1034, TT 1035, TT 1047, TT 1060, TT 1064, TT 1066, TT 1071, TT 1074, TT 1081, TT 1082, TT 1085, TT 1087, TT 1094, TT 1095, TT 1105, TT 1106, TT 1114

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A12 Examples of transcripts from a heart cDNA library categorized as members of cell division/DNA synthesis, repair and replication (1 clone)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0221-LF	High mobility group protein DSP1 (Dorsal switch protein 1)	*Tribolium castaneum*	4e-56

Table A13 Examples of transcripts from a heart cDNA library categorized as members of defense and homeostasis (6 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0185-LF	*Latency associated nuclear antigen*	*Saimiriine herpesvirus 2*	9e-12
HT-N-S01-0382-LF	*Lipoamide dehydrogenase*	*Sus scrofa*	3e-51
HT-N-S01-0140-LF	*Lysozyme*	*Penaeus monodon*	5e-85
HT-N-S01-0279-LF			1e-84
HT-N-S01-0219-LF	*Plasminogen*	*Sus scrofa*	2e-13
HT-N-S01-0144-LF	*Serine proteinase inhibitor*	*Pacifastacus leniusculus*	4e-42

Table A14 Examples of transcripts from a heart cDNA library categorized as members of gene expression, regulation and protein synthesis (13 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0452-LF	*Arginine kinase*	*Fenneropenaeus chinensis*	2e-40
HT-N-S01-0093-LF	*Calponin*	*Aedes aegypti*	4e-34
HT-N-S01-0055-LF	*COG3321: Polyketide synthase modules and related proteins*	*Burkholderia mallei GB8 horse 4*	6e-05
HT-N-S01-0189-LF	*Elongation factor 2b CG2238-PB, isoform B*	*Drosophila melanogaster*	3e-61
HT-N-S01-0442-LF	*Elongation factor-1 alpha*	*Penaeus monodon*	1e-104
HT-N-S01-0143-LF	*Elongation factor-2*	*Libinia emarginata*	1e-118
HT-N-S01-0032-LF	*Fast tropomyosin isoform*	*Homarus americanus*	3e-44
HT-N-S01-0089-LF	*High mobility group 20A*	*Mus musculus*	1e-43
HT-N-S01-0012-LF	*Hydroxyproline-rich protein*	*Micrococcus sp. 28*	4e-05
HT-N-S01-0419-LF	*Muscle lim protein*	*Aedes aegypti*	2e-34
HT-N-S01-0243-LF	*Myosin 1 light chain*	*Lonomia obliqua*	1e-48
HT-N-S01-0023-LF	*Myosin light chain 1*	*Aedes aegypti*	2e-38
HT-N-S01-0417-LF	*Reticulon 4-L2*	*Takifugu rubripes*	6e-42

Table A15 Examples of transcripts from a heart cDNA library categorized as members of internal/external structure and motility (5 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0407-LF	*Fasciclin-like protein*	*Aplysia californica*	3e-27
HT-N-S01-0240-LF	*H3 histone, family 3B*	*Mus musculus*	2e-69
HT-N-S01-0158-LF	*Profilin (Chickadee protein)*	*Tribolium castaneum*	7e-38
HT-N-S01-0035-LF	*Tubulin beta-1 chain*	*Homarus americanus*	6e-89
HT-N-S01-0421-LF	*Tyrosine-protein phosphatase non-receptor type 13 (Protein-tyrosine phosphatase 1E) (PTP-E1) (hPTPE1) (PTP-BAS) (Protein-tyrosine phosphatase PTPL1)*	*Tribolium castaneum*	8e-50

Table A16 Examples of transcripts from a heart cDNA library categorized as members of metabolism (27 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0441-LF	*Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain*	*Danio rerio*	4e-73
HT-N-S01-0119-LF	*ADP/ATP translocase*	*Bombyx mori*	7e-73
HT-N-S01-0350-LF	*ATP lipid-binding protein like protein*	*Marsupenaeus japonicus*	1e-30
HT-N-S01-0247-LF	*ATP synthase*	*Penaeus monodon*	6e-74
HT-N-S01-0296-LF			9e-93
HT-N-S01-0013-LF	*ATP synthase F0 subunit 6*	*Penaeus monodon*	1e-10
HT-N-S01-0078-LF			4e-83
HT-N-S01-0173-LF			2e-36
HT-N-S01-0193-LF			5e-95
HT-N-S01-0195-LF			1e-40
HT-N-S01-0329-LF	*ATP synthase F0 subunit 6*	*Penaeus monodon*	3e-31
HT-N-S01-0332-LF			2e-89
HT-N-S01-0372-LF			1e-102
HT-N-S01-0385-LF			8e-42
HT-N-S01-0414-LF			5e-71
HT-N-S01-0176-LF	*COG2804: Type II secretory pathway, ATPase PulE/Tfp pilus assembly pathway, ATPase PilB*	*Pseudomonas aeruginosa UCBPP-PA14*	3e-05
HT-N-S01-0244-LF	*ERO1-like*	*Gallus gallus*	2e-41
HT-N-S01-0154-LF	*Fructose 1,6-bisphosphate aldolase*	*Oncometopia nigricans*	1e-16
HT-N-S01-0297-LF			1e-37
HT-N-S01-0231-LF	*gcdh protein (Glutaryl-Coenzyme A dehydrogenase)*	*Danio rerio*	4e-70
HT-N-S01-0291-LF	*Glutathione S-transferase*	*Anopheles gambiae*	7e-35
HT-N-S01-0317-LF	*Glyceraldehyde-3-phosphate dehydrogenase*	*Procambarus clarkii*	6e-76
HT-N-S01-0162-LF	*NADH-ubiquinone oxidoreductase Fe-S protein 7*	*Bombyx mori*	3e-62
HT-N-S01-0218-LF	*Peroxisomal 3,2-trans-enoyl-CoA isomerase (Dodecenoyl-CoA delta-isomerase) isoform 1*	*Strongylocentrotus purpuratus*	5e-53
HT-N-S01-0397-LF	*Phosphoglycerate kinase*	*Aedes aegypti*	1e-14
HT-N-S01-0427-LF	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	1e-90
HT-N-S01-0294-LF	*Trifunctional enzyme beta subunit (tp-beta)*	*Aedes aegypti*	3e-48

Table A17 Examples of transcripts from a heart cDNA library categorized as members of miscellaneous function (16 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0450-LF	*Cathepsin D*	*Aedes aegypti*	7e-86
HT-N-S01-0165-LF	*C-type lectin*	*Penaeus monodon*	7e-06
HT-N-S01-0178-LF	*Dehydrogenase/3-ketoacyl-Coenzyme A thiolase*	*Danio rerio*	5e-09
HT-N-S01-0198-LF	*GRN protein (granulin)*	*Xenopus tropicalis*	1e-33
HT-N-S01-0433-LF	*Macrophage migration inhibitory factor*	*Bombyx mori*	5e-25
HT-N-S01-0241-LF	*Neuroparsin A precursor*	*Locusta migratoria*	9e-11
HT-N-S01-0289-LF			5e-06
HT-N-S01-0222-LF	*Papilin*	*Aedes aegypti*	3e-36
HT-N-S01-0306-LF	*pxSerpin 3*	*Plutella xylostella*	9e-09
HT-N-S01-0384-LF	*Supervillin, isoform CRA_a*	*Homo sapiens*	8e-16
HT-N-S01-0204-LF	*Tetratricopeptide repeat domain 35*	*Danio rerio*	2e-47
HT-N-S01-0358-LF	*Thyroid hormone receptor interactor 12 isoform 7*	*Canis familiaris*	7e-35
HT-N-S01-0128-LF	*Troponin T*	*Libellula pulchella*	2e-12
HT-N-S01-0200-LF	*Troponin T-1*	*Drosophila melanogaster*	1E-12
HT-N-S01-0183-LF	*Trypsin 10*	*Xenopus tropicalis*	7e-17
HT-N-S01-0039-LF			3e-19

Table A18 Examples of transcripts from a heart cDNA library categorized as members of mitochondrial protein (104 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0014-LF	*Cytochrome b*	*Penaeus monodon*	2e-46
HT-N-S01-0082-LF			5e-72
HT-N-S01-0216-LF			4e-79
HT-N-S01-0237-LF			1e-117
HT-N-S01-0272-LF			8e-95
HT-N-S01-0322-LF			7e-44
HT-N-S01-0428-LF			3e-41
HT-N-S01-0435-LF			1e-117
HT-N-S01-0022-LF		*Portunus trituberculatus*	1e-43
		Triops cancriformis	
HT-N-S01-0069-LF			8e-49
HT-N-S01-0006-LF	*Cytochrome c oxidase subunit I*	*Fenneropenaeus merguiensis*	1e-76
HT-N-S01-0028-LF			2e-36
HT-N-S01-0061-LF			2e-65
HT-N-S01-0074-LF			1e-84
HT-N-S01-0097-LF			5e-52
HT-N-S01-0130-LF			4e-94
HT-N-S01-0135-LF			5e-87
HT-N-S01-0163-LF			2e-88
HT-N-S01-0238-LF			3e-73
HT-N-S01-0307-LF			3e-80
HT-N-S01-0308-LF			1e-91
HT-N-S01-0309-LF			1e-38
HT-N-S01-0327-LF			2e-39
HT-N-S01-0333-LF			8e-37

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A18 (cont.)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0359-LF	*Cytochrome c oxidase subunit I*	*Fenneropenaeus merguiensis*	7e-64
HT-N-S01-0360-LF			3e-73
HT-N-S01-0379-LF			6e-24
HT-N-S01-0386-LF			6e-79
HT-N-S01-0395-LF			6e-29
HT-N-S01-0034-LF		*Marsupenaeus japonicus*	1e-53
HT-N-S01-0083-LF			7e-75
HT-N-S01-0103-LF			6e-42
HT-N-S01-0153-LF			3e-79
HT-N-S01-0161-LF			1e-74
HT-N-S01-0167-LF			6e-90
HT-N-S01-0172-LF			8e-39
HT-N-S01-0191-LF			1e-77
HT-N-S01-0199-LF			3e-40
HT-N-S01-0211-LF			1e-100
HT-N-S01-0213-LF			1e-82
HT-N-S01-0228-LF			2e-97
HT-N-S01-0232-LF			1e-102
HT-N-S01-0264-LF			1e-102
HT-N-S01-0268-LF			1e-101
HT-N-S01-0284-LF			3e-81
HT-N-S01-0431-LF			1e-102
HT-N-S01-0438-LF			1e-102
HT-N-S01-0443-LF			1e-103
HT-N-S01-0436-LF		*Pagurus longicarpus*	5e-96
HT-N-S01-0087-LF		*Penaeus monodon*	1e-104
HT-N-S01-0280-LF			2e-94
HT-N-S01-0398-LF			1e-69
HT-N-S01-0002-LF	*Cytochrome c oxidase subunit II*	*Penaeus monodon*	2e-26
HT-N-S01-0072-LF			3e-32
HT-N-S01-0079-LF			3e-69
HT-N-S01-0095-LF			1e-110
HT-N-S01-0261-LF			1e-101
HT-N-S01-0274-LF			1e-107
HT-N-S01-0304-LF			3e-97
HT-N-S01-0328-LF			1e-106
HT-N-S01-0373-LF			2e-48
HT-N-S01-0390-LF			7e-76
HT-N-S01-0420-LF			1e-111
HT-N-S01-0007-LF	*Cytochrome c oxidase subunit III*	*Penaeus monodon*	5e-31
HT-N-S01-0040-LF			1e-103
HT-N-S01-0044-LF			5e-80
HT-N-S01-0065-LF			9e-40
HT-N-S01-0170-LF			5e-79
HT-N-S01-0215-LF			2e-64
HT-N-S01-0269-LF			2e-81
HT-N-S01-0310-LF			8e-47
HT-N-S01-0453-LF			1e-111
HT-N-S01-0041-LF	*Cytochrome c oxidase subunit Va*	*Rhyzopertha dominica*	2e-25
HT-N-S01-0104-LF			1e-40
HT-N-S01-0133-LF	*Cytochrome c oxidase subunit VIIc precursor*	*Gallus gallus*	4e-11

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A18 (cont.)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0094-LF	*Cytochrome oxidase subunit I*	*Euphausia superba*	9e-25
HT-N-S01-0312-LF		*Fenneropenaeus penicillatus*	2e-93
HT-N-S01-0357-LF		*Fenneropenaeus sp. HCCP-2002*	1e-30
HT-N-S01-0313-LF	*Cytochrome oxidase subunit II*	*Euphausia superba*	1e-95
HT-N-S01-0142-LF	*Mitochondrial cytochrome c oxidase polypeptide VIIc*	*Graphocephala atropunctata*	2e-13
HT-N-S01-0285-LF	*Mitochondrial NADH dehydrogenase (ubiquinone) 1 alpha subcomplex*	*Aedes aegypti*	6e-06
HT-N-S01-0319-LF	*NADH dehydrogenase*	*Aedes aegypti*	4e-17
HT-N-S01-0430-LF			2e-62
HT-N-S01-0155-LF	*NADH dehydrogenase subunit 1*	*Penaeus monodon*	2e-78
HT-N-S01-0179-LF			3e-53
HT-N-S01-0246-LF			2e-96
HT-N-S01-0278-LF			1e-100
HT-N-S01-0059-LF	*NADH dehydrogenase subunit 2*	*Penaeus monodon*	4e-15
HT-N-S01-0054-LF	*NADH dehydrogenase subunit 4*	*Penaeus monodon*	2e-36
HT-N-S01-0335-LF			7e-20
HT-N-S01-0371-LF			4e-36
HT-N-S01-0174-LF	*NADH dehydrogenase subunit 4L*	*Penaeus monodon*	6e-35
HT-N-S01-0362-LF			2e-29
HT-N-S01-0394-LF			5e-26
HT-N-S01-0080-LF	*NADH dehydrogenase subunit 5*	*Penaeus monodon*	5e-22
HT-N-S01-0125-LF			1e-101
HT-N-S01-0145-LF			2e-95
HT-N-S01-0151-LF			2e-85
HT-N-S01-0311-LF			2e-23
HT-N-S01-0416-LF			5e-85
HT-N-S01-0299-LF	*NADH dehydrogenase subunit 6*	*Penaeus monodon*	5e-72
HT-N-S01-0448-LF			3e-71
HT-N-S01-0063-LF	*Ubiquinol cytochrome C oxidoreductase-subunit 6.4kD-subunit*	*Aedes aegypti*	1e-08
HT-N-S01-0399-LF	*Voltage-dependent anion-selective channel protein 2 (Outer mitochondrial membrane protein porin)*	*Xenopus laevis*	4e-33

Table A19 Examples of transcripts from a heart cDNA library categorized as members of transport (1 clone)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0016-LF	*Electron-transfer-flavoprotein beta polypeptide*	*Bombyx mori*	6e-50

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A20 Examples of transcripts from a heart cDNA library categorized as members of ribosomal and rRNA (70 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0043-LF	*40S ribosomal protein S10*	*Rattus norvegicus*	5e-19
HT-N-S01-0141-LF		*Canis familiaris*	5e-45
HT-N-S01-0062-LF	*60S ribosomal protein L31*	*Spodoptera frugiperda*	3e-32
HT-N-S01-0064-LF	*60S ribosomal protein L37a*	*Ostertagia ostertagi*	1e-34
HT-N-S01-0206-LF	*Ribosomal protein L13a*	*Lysiphlebus testaceipes*	1e-64
HT-N-S01-0320-LF	*Ribosomal protein L21*	*Ixodes scapularis*	5e-47
HT-N-S01-0042-LF	*Ribosomal protein L24*	*Marsupenaeus japonicus*	2e-52
HT-N-S01-0177-LF	*Ribosomal protein S15*	*Branchiostoma belcheri tsingtaunese*	9e-65
HT-N-S01-0408-LF	*Ribosomal protein S17*	*Homo sapiens*	4e-56
HT-N-S01-0273-LF	*Ribosomal protein S9 CG3395-PA, isoform A*	*Apis mellifera*	1e-82
HT-N-S01-0212-LF	*S5e ribosomal protein*	*Dascillus cervinus*	2e-96
HT-N-S01-0075-LF	*tRNA-Ile*	*Penaeus monodon*	0
HT-N-S01-0076-LF			0
HT-N-S01-0081-LF			0
HT-N-S01-0088-LF			0
HT-N-S01-0092-LF			0
HT-N-S01-0100-LF			0
HT-N-S01-0106-LF			1e-148
HT-N-S01-0107-LF			0
HT-N-S01-0108-LF			1e-163
HT-N-S01-0139-LF			0
HT-N-S01-0147-LF			0
HT-N-S01-0148-LF			0
HT-N-S01-0149-LF			0
HT-N-S01-0150-LF			3e-86
HT-N-S01-0152-LF			0
HT-N-S01-0159-LF			1e-104
HT-N-S01-0180-LF			0
HT-N-S01-0182-LF			1e-150
HT-N-S01-0188-LF			1e-138
HT-N-S01-0192-LF			0
HT-N-S01-0194-LF			0
HT-N-S01-0202-LF			0
HT-N-S01-0205-LF			0
HT-N-S01-0207-LF			0
HT-N-S01-0208-LF			1e-146
HT-N-S01-0210-LF			0
HT-N-S01-0220-LF			0
HT-N-S01-0223-LF			0
HT-N-S01-0229-LF			0
HT-N-S01-0242-LF			0
HT-N-S01-0249-LF			0
HT-N-S01-0253-LF			0
HT-N-S01-0258-LF			0
HT-N-S01-0259-LF			0
HT-N-S01-0260-LF			0
HT-N-S01-0263-LF			0
HT-N-S01-0266-LF			0
HT-N-S01-0270-LF			0
HT-N-S01-0283-LF			0

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A20 (cont.)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0288-LF	*tRNA-Ile*	*Penaeus monodon*	2e-78
HT-N-S01-0290-LF			3e-77
HT-N-S01-0314-LF			0
HT-N-S01-0316-LF			0
HT-N-S01-0323-LF			1e-103
HT-N-S01-0331-LF			1e-159
HT-N-S01-0342-LF			1e-141
HT-N-S01-0347-LF			2e-80
HT-N-S01-0392-LF			0
HT-N-S01-0406-LF			0
HT-N-S01-0412-LF			0
HT-N-S01-0413-LF			0
HT-N-S01-0418-LF			1e-164
HT-N-S01-0422-LF			0
HT-N-S01-0429-LF			0
HT-N-S01-0437-LF			0
HT-N-S01-0444-LF			0
HT-N-S01-0445-LF			0
HT-N-S01-0446-LF			0
HT-N-S01-0451-LF			0

Table A21 Examples of transcripts from a heart cDNA library categorized as members of unidentified (hypothetical)-similar to other cDNA/DNA (40 clones)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0138-LF	CG1102-PA	*Drosophila melanogaster*	7e-23
HT-N-S01-0369-LF	CG14996-PB	*Drosophila melanogaster*	3e-41
HT-N-S01-0424-LF	CG1572-PA, isoform A isoform 1	*Tribolium castaneum*	2e-33
HT-N-S01-0282-LF	CG40410-PA	*Apis mellifera*	1e-50
HT-N-S01-0053-LF	CG5028-PA	*Drosophila melanogaster*	1e-17
HT-N-S01-0257-LF	CG5451-PA	*Drosophila melanogaster*	4e-13
HT-N-S01-0404-LF	CG7630-PA	*Drosophila melanogaster*	7e-07
HT-N-S01-0038-LF	Conserved hypothetical protein	*Neosartorya fischeri NRRL 181*	8e-08
HT-N-S01-0121-LF	DNA segment, Chr 10, Johns Hopkins University 81 expressed	*Mus musculus*	5e-48
HT-N-S01-0365-LF	ENSANGP00000009996	*Anopheles gambiae*	6e-59
HT-N-S01-0073-LF	ENSANGP00000010943	*Apis mellifera*	9e-15
HT-N-S01-0171-LF	ENSANGP00000011689	*Anopheles gambiae*	2e-08
HT-N-S01-0364-LF	ENSANGP00000011882	*Anopheles gambiae*	5e-40
HT-N-S01-0045-LF	ENSANGP00000013096	*Anopheles gambiae*	2e-31
HT-N-S01-0380-LF	ENSANGP00000013298	*Anopheles gambiae str. PEST*	4e-06
HT-N-S01-0126-LF	ENSANGP00000014057	*Apis mellifera*	8e-13
HT-N-S01-0122-LF	ENSANGP00000018202	*Anopheles gambiae*	1e-30
HT-N-S01-0196-LF	ENSANGP00000020213	*Anopheles gambiae*	4e-06
HT-N-S01-0248-LF	ENSANGP00000025500	*Anopheles gambiae str. PEST*	1e-15
HT-N-S01-0250-LF	ENSANGP00000030087	*Anopheles gambiae str. PEST*	3e-07
HT-N-S01-0157-LF	GA13958-PA	*Drosophila pseudoobscura*	5e-07
HT-N-S01-0255-LF	GA15266-PA	*Drosophila pseudoobscura*	6e-18

Table A21 (cont.)

Clone No.	Transcripts	Closest Species	E-value
HT-N-S01-0137-LF	Hypothetical 18K protein	*Carassius auratus*	5e-05
HT-N-S01-0156-LF			9e-05
HT-N-S01-0166-LF			8e-05
HT-N-S01-0265-LF			5e-05
HT-N-S01-0286-LF			5e-05
HT-N-S01-0338-LF			6e-05
HT-N-S01-0440-LF			6e-05
HT-N-S01-0086-LF	Hypothetical protein	*Trypanosoma brucei TREU927*	1e-14
HT-N-S01-0186-LF	Hypothetical protein	*Arabidopsis thaliana*	2e-19
HT-N-S01-0467-LF	Hypothetical protein	*Eimeria tenella str. Houghton*	5e-09
HT-N-S01-0368-LF	Hypothetical protein Efae03002652	*Enterococcus faecium*	8e-06
HT-N-S01-0262-LF	LOC496680 protein	*Xenopus tropicalis*	1e-13
HT-N-S01-0432-LF	Unknown	*Unknown*	1e-06
HT-N-S01-0021-LF	Unknown (protein for MGC:101057)	*Danio rerio*	2e-09
HT-N-S01-0116-LF	Unknown (protein for MGC:82165)	*Xenopus laevis*	6e-20
HT-N-S01-0301-LF			8e-17
HT-N-S01-0008-LF	unnamed protein product	*Tetraodon nigroviridis*	2e-37
HT-N-S01-0031-LF	unnamed protein product	*Tetraodon nigroviridis*	2e-08

Table A22 Unknown transcripts from a heart cDNA library (130 clones)

Gene Identity	No. of Clones	Clone
Unknown genes	130	HT 001, HT 003, HT 004, HT 005, HT 009, HT 015, HT 017, HT 018, HT 020, HT 024, HT 025, HT 026, HT 027, HT 030, HT 033, HT 036, HT 037, HT 049, HT 050, HT 051, HT 056, HT 057, HT 058, HT 060, HT 066, HT 067, HT 068, HT 070, HT 071, HT 084, HT 085, HT 090, HT 098, HT 099, HT 105, HT 110, HT 112, HT 113, HT 114, HT 115, HT 118, HT 120, HT 124, HT 127, HT 129, HT 131, HT 132, HT 134, HT 136, HT 146, HT 160, HT 164, HT 168, HT 169, HT 175, HT 181, HT 184, HT 187, HT 190, HT 197, HT 201, HT 203, HT 209, HT 214, HT 217, HT 224, HT 225, HT 226, HT 227, HT 230, HT 233, HT 234, HT 235, HT 236, HT 239, HT 251, HT 252, HT 254, HT 267, HT 271, HT 275, HT 276, HT 281, HT 287, HT 292, HT 293, HT 295, HT 298, HT 300, HT 302, HT 303, HT 305, HT 315, HT 325, HT 326, HT 330, HT 334, HT 341, HT 348, HT 361, HT 363, HT 366, HT 367, HT 370, HT 374, HT 375, HT 376, HT 377, HT 378, HT 381, HT 383, HT 387, HT 388, HT 393, HT 396, HT 403, HT 409, HT 410, HT 411, HT 415, HT 434, HT 454, HT 457, HT 458, HT 461, HT 462, HT 463, HT 464, HT 465, HT 468

ศูนย์วิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

Table A23 Examples of transcripts from the forward SSH library of testes of *P. monosdon* categorized as members of sex-related gene (1 clone)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0019-W	*Meiotic recombination protein DMC1/LIM15 homolog isoform 1*	*Canis familiaris*	506	1e-24

Table A24 Examples of transcripts from the forward SSH library of testes of *P. monosdon* categorized as members of stress response and cell defense protein (4 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0148-W	*Allergen Pen m 2*	*Penaeus monodon*	309	5e-22
TT-N-ST01-0150-W	*Antimicrobial peptide*	*Fenneropenaeus chinensis*	389	2e-22
TT-N-ST01-0147-W	*C2 domain containing protein*	*Tetrahymena thermophila SB210*	592	2e-19
TT-N-ST01-0069-W	*Myelodysplasia/myeloid leukemia factor CG8295-PD, isoform D*	*Drosophila melanogaster*	709	1e-33

Table A25 Examples of transcripts from the forward SSH library of testes of *P. monosdon* categorized as members of protein synthesis and DNA replication (14 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0059-W	*Elongation factor-1 alpha*	*Penaeus monodon*	353	7e-62
TT-N-ST01-0099-W		*Armadillidium vulgare*	557	2e-95
TT-N-ST01-0039-W	*Eukaryotic translation initiation factor 2 subunit 2*	*Bombyx mori*	600	7e-36
TT-N-ST01-0088-W	*Oncoprotein nm23*	*Ictalurus punctatus*	436	6e-34
TT-N-ST01-0156-W	*ATP-dependent RNA helicase*	*Aedes aegypti*	472	3e-76
TT-N-ST01-0085-W	*Basic leucine zipper and W2 domain-containing protein 2*	*Danio rerio*	466	2e-09
TT-N-ST01-0149-W	*C-1-tetrahydrofolate synthase, cytoplasmic (C1-THF synthase)*	*Pongo pygmaeus*	714	7e-85
TT-N-ST01-0143-W	*Ras-related nuclear protein*	*Marsupenaeus japonicus*	602	4e-50
TT-N-ST01-0178-W	*Sensitized chromosome inheritance modifier 19 CG9241-PA*	*Drosophila melanogaster*	368	3e-15
TT-N-ST01-0103-W	*Signal peptidase complex subunit 2 homolog*	*Tribolium castaneum*	352	3e-28
TT-N-ST01-0138-W	*Signal sequence receptor*	*Bombyx mori*	279	9e-14
TT-N-ST01-0077-W	*Actin-depolymerizing factor 1*	*Bombyx mori*	316	8e-21
TT-N-ST01-0109-W	*Bmsqd-2*	*Apis mellifera*	713	1e-102
TT-N-ST01-0131-W	*DEAD (Asp-Glu-Ala-Asp) box polypeptide 54 isoform 3*	*Pan troglodytes*	562	1e-19

Table A26 Examples of transcripts from the forward SSH library of testes of *P. monosdon* categorized as members of internal/external structure, motility, and transport (2 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0162-W	*Transmembrane protein*	*Pan troglodytes*	540	2e-64
TT-N-ST01-0093-W	*NTF2-related export protein (p15)*	*Tribolium castaneum*	420	4e-19

Table A27 Examples of transcripts from the forward SSH library of testes of *P. monosdon* categorized as members of metabolism (18 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0053-W	*Cytochrome c oxidase subunit 6a polypeptide 1*	*Xenopus tropicalis*	345	2e-08
TT-N-ST01-0007-W	*Cytochrome c oxidase subunit I*	*Fenneropenaeus chinensis*	467	4e-76
TT-N-ST01-0137-W		*Marsupenaeus japonicus*	712	1e-112
TT-N-ST01-0030-W	*Cytochrome c oxidase subunit III*	*Penaeus monodon*	439	1e-70
TT-N-ST01-0155-W	*26S proteasome regulatory complex ATPase RPT4*	*Aedes aegypti*	490	2e-81
TT-N-ST01-0108-W	*Proteasome (prosome, macropain) 26S subunit, ATPase, 5, isoform CRA_a*	*Homo sapiens*	199	3e-12
TT-N-ST01-0161-W	*Proteasome 26S non-ATPase subunit 12*	*Tribolium castaneum*	209	4e-18
TT-N-ST01-0074-W	*Proteasome subunit alpha type 2 (Proteasome component C3) (Macropain subunit C3) (Multicatalytic endopeptidase complex subunit C3)*	*Strongylocentrotus purpuratus*	401	4e-40
TT-N-ST01-0056-W	*Proteasome subunit, alpha type, 5*	*Apis mellifera*	707	4e-23
TT-N-ST01-0182-W	*GTP-binding protein*	*Bombyx mori*	373	5e-49
TT-N-ST01-0119-W	*Cytosolic manganese superoxide dismutase*	*Penaeus monodon*	343	2e-12
TT-N-ST01-0141-W			346	6e-13
TT-N-ST01-0125-W	*Dolichyl-diphosphooligosaccharide--proteinglycotransferase*	*Branchiostoma belcheri tsingtaunese*	427	4e-19
TT-N-ST01-0116-W	*Activated protein kinase C receptor*	*Toxoptera citricida*	410	4e-70
TT-N-ST01-0035-W	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	644	1e-113
TT-N-ST01-0058-W			702	7e-17
TT-N-ST01-0034-W	*myosin*	*Dictyostelium discoideum AX4*	596	6e-23
TT-N-ST01-0010-W	*Malate dehydrogenase 1, NAD (soluble), isoform CRA_d*	*Homo sapiens*	374	4e-43

Table A28 Known transcripts from the forward SSH library of testes of *P. monosdon* categorized as members of ribosomal and rRNA (18 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0037-W	*40S ribosomal protein*	*Perinereis aibuhitensis*	430	6e-51
TT-N-ST01-0024-W	*40S ribosomal protein S15 (RIG protein) isoform 1*	*Macaca mulatta*	425	3e-60
TT-N-ST01-0107-W	*40S ribosomal protein S2*	*Ornithodoros parkeri*	496	3e-56
TT-N-ST01-0044-W	*40S ribosomal protein S23*	*Argas monolakensis*	366	1e-62
TT-N-ST01-0072-W	*40S ribosomal protein S4*	*Spodoptera frugiperda*	422	4e-46
TT-N-ST01-0152-W	*Ribosomal protein L10*	*Callinectes sapidus*	480	1e-85
TT-N-ST01-0002-W	*Ribosomal protein L10Ae*	*Biphyllus lunatus*	618	8e-78
TT-N-ST01-0028-W	*Ribosomal protein L11e*	*Scarabaeus laticollis*	174	1e-16
TT-N-ST01-0062-W	*Ribosomal protein L12*	*Ixodes scapularis*	477	1e-54
TT-N-ST01-0167-W	*Ribosomal protein L3*	*Bombyx mori*	350	3e-51
TT-N-ST01-0081-W	*Ribosomal protein LP1*	*Argas monolakensis*	477	3e-38
TT-N-ST01-0095-W	*Ribosomal protein S2*	*Lysiphlebus testaceipes*	439	4e-74
TT-N-ST01-0021-W	*Ribosomal protein S20*	*Chlamys farreri*	402	1e-43
TT-N-ST01-0017-W	*Ribosomal protein S3*	*Gallus gallus*	467	3e-72
TT-N-ST01-0070-W	*Ribosomal protein S3*	*Danio rerio*	535	3e-81
TT-N-ST01-0071-W	*Ribosomal protein S3A*	*Spodoptera frugiperda*	545	2e-74
TT-N-ST01-0169-W	*tRNA-Ile (16S ribosomal RNA gene)*	*Penaeus monodon*	714	0
TT-N-ST01-0049-W			296	1e-153

Table A29 Examples of transcripts the forward SSH library of testes of *P. monosdon* categorized as members of unidentified (hypothetical)-similar to other cDNA/DNA) from (9 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST01-0055-W	mFLJ00348 protein	*Mus musculus*	212	3e-11
TT-N-ST01-0041-W	Hypothetical protein	*Geobacter uraniumreducens Rf4*	356	3e-11
TT-N-ST01-0151-W			281	4e-12
TT-N-ST01-0145-W	Hypothetical protein	*Dictyostelium discoideum AX4*	258	3e-05
TT-N-ST01-0157-W	Hypothetical protein	*Rattus norvegicus*	528	3e-11
TT-N-ST01-0173-W	Hypothetical protein	*Plasmodium falciparum 3D7*	545	2e-17
TT-N-ST01-0047-W	Unknown	*Schistosoma japonicum*	273	4e-08
TT-N-ST01-0073-W	Unknown	*Schistosoma japonicum*	503	5e-14
TT-N-ST01-0181-W			309	1e-15

Table A30 Clustering of unknown transcripts of the forward SSH library of testes of *P. monosdon* (112 clones)

Gene	Clustering	Clone
Unknown genes	CL1Contig1 (13 clones)	**ST01-0180** (305) [ST01-0006 (351), ST01-0029 (537), ST01-0054 (188), ST01-0094 (285), ST01-0114 (291), ST01-0124 (244), ST01-0126 (299), ST01-0128 (291), ST01-0139 (256), ST01-0164 (296), ST01-0170 (289), ST01-0174 (244*)]
	CL1Contig2 (16 clones)	**ST01-0052** (176) [ST01-0031 (230), ST01-0042 (219), ST01-0068 (236), ST01-0075 (223), ST01-0090 (243), ST01-0096 (225), ST01-0098 (236), ST01-0117 (224), ST01-0132 (207), ST01-0133 (226), ST01-0134 (213), ST01-0159 (252), ST01-0166 (243), ST01-0168 (343), ST01-0179 (218)]
	CL2Contig1 (4 clones)	**ST01-0076** (644) [ST01-0022 (537), ST01-0038 (453), ST01-0171 (453)]
	CL3Contig1 (4 clones)	**ST01-0012** (333) [ST01-0015 (265), ST01-0091 (318), ST01-0177 (308)]
	CL4Contig1 (4 clones)	**ST01-0016** (370) [ST01-0040 (267), ST01-0102 (355), ST01-0130 (278)]
	CL5Contig1 (4 clones)	**ST01-0111** (269) [ST01-0011 (281), ST01-0025 (125), ST01-0036 (304)]
	CL8Contig1 (2 clones)	**ST01-0110** (568) [ST01-0018 (296)]
	CL10Contig1 (2 clones)	**ST01-0023** (684) [ST01-0176 (684)]
	CL11Contig1 (2 clones)	**ST01-0165** (359) [ST01-0008 (358)]
	CL12Contig1 (2 clones)	**ST01-0020** (393) [ST01-0144 (393)]
	CL13Contig1 (2 clones)	**ST01-0032** (608) [ST01-0142 (611)]
	CL16Contig1 (2 clones)	**ST01-0118** (425) [ST01-0175 (424)]
	Singleton (55 clones)	ST01-0001 (350), ST01-0003 (522), ST01-0004 (413), ST01-0005 (518), ST01-0009 (266), ST01-0013 (258), ST01-0014 (440), ST01-0026 (415), ST01-0027 (243), ST01-0033 (650), ST01-0043 (398), ST01-0045 (456), ST01-0046 (353), ST01-0048 (369), ST01-0050 (382), ST01-0051 (330), ST01-0057 (475), ST01-0060 (695), ST01-0061 (245), ST01-0063 (272), ST01-0064 (311), ST01-0065 (319), ST01-0066 (519), ST01-0067 (466), ST01-0078 (248), ST01-0079 (432), ST01-0080 (271), ST01-0082 (468), ST01-0083 (229), ST01-0084 (429), ST01-0086 (375), ST01-0087 (351), ST01-0089 (237), ST01-0092 (424), ST01-0097 (310), ST01-0100 (380), ST01-0101 (354), ST01-0104 (510), ST01-0105 (229), ST01-0112 (340), ST01-0113 (283), ST01-0115 (284), ST01-0120 (229), ST01-0121 (435), ST01-0122 (708), ST01-0123 (427), ST01-0127 (373), ST01-0129 (558), ST01-0136 (209), ST01-0140 (228), ST01-0146 (353), ST01-0158 (283), ST01-0160 (416), ST01-0163 (273), ST01-0172 (648)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A31 Examples of transcripts the reverse SSH library of testes of *P. monosdon* categorized as members of sex-related gene) from (1 clone)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST02-0022-LF	*Progestin receptor membrane component 1*	*Oryzias latipes*	574	1e-47

Table A32 Examples of transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of stress response and cell defense protein (13 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST02-0024-LF	*Allergen Pen m 2*	*Penaeus monodon*	309	2e-21
TT-N-ST02-0188-LF			309	5e-22
TT-N-ST02-0054-LF	*Heat shock protein gp96*	*Strongylocentrotus purpuratus*	498	1e-21
TT-N-ST02-0078-LF	*Niemann-Pick disease type C2*	*Oreochromis mossambicus*	456	5e-06
TT-N-ST02-0108-LF			456	5e-06
TT-N-ST02-0097-LF	*Calcitonin gene-related peptide-receptor component protein isoform a*	*Homo sapiens*	510	8e-21
TT-N-ST02-0010-LF	*Cathepsin B*	*Hippoglossus hippoglossus*	238	1e-26
TT-N-ST02-0065-LF	*Small optic lobes CG1391-PB, isoform B (Calpain)*	*Apis mellifera*	580	7e-81
TT-N-ST02-0070-LF			580	7e-81
TT-N-ST02-0185-LF	*Tetraspanin 3, isoform CRA_a*	*Homo sapiens*	590	4e-10
TT-N-ST02-0056-LF	*Tetraspanin 96F CG6120-PA*	*Drosophila melanogaster*	438	5e-14
TT-N-ST02-0016-LF	*Variable surface lipoprotein*	*Mycoplasma bovis*	507	1e-08

Table A33 Examples of transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of metabolism (13 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST02-0029-LF	*Alcohol dehydrogenase*	*Bombyx mori*	323	3e-35
TT-N-ST02-0156-LF			680	2e-90
TT-N-ST02-0176-LF			680	2e-90
TT-N-ST02-0141-LF	*Cytochrome b*	*Penaeus monodon*	494	3e-80
TT-N-ST02-0007-LF	*Cytochrome c oxidase subunit I*	*Fenneropenaeus chinensis*	403	5e-68
TT-N-ST02-0020-LF			405	1e-66
TT-N-ST02-0117-LF			405	7e-68
TT-N-ST02-0075-LF	*GTP binding protein*	*Bombyx mori*	503	2e-70
TT-N-ST02-0040-LF	*Proteasome (prosome, macropain) 26S subunit, non-ATPase, 13*	*Tribolium castaneum*	706	2e-78
TT-N-ST02-0101-LF	*Proteasome 26S non-ATPase subunit 12*	*Tribolium castaneum*	560	1e-71
TT-N-ST02-0130-LF			560	1e-71
TT-N-ST02-0049-LF	*Proteasome 26S subunit subunit 4 ATPase CG5289-PA*	*Drosophila melanogaster*	647	1e-89
TT-N-ST02-0165-LF	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	628	1e-107

Table A34 Examples of transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of protein synthesis and DNA replication (22 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST02-0063-LF	*Cement precursor protein 3B variant 2*	*Phragmatopoma californica*	437	3e-08
TT-N-ST02-0106-LF	*Cement precursor protein 3B variant 3*	*Phragmatopoma californica*	620	2e-11
TT-N-ST02-0136-LF			270	3e-07
TT-N-ST02-0038-LF	*Centromere/kinetochore protein zw10 homolog*	*Apis mellifera*	423	1e-20
TT-N-ST02-0131-LF	*Drosophila melanogaster eEF1delta*	*Drosophila yakuba*	146	8e-16
TT-N-ST02-0004-LF	*Elongation factor-1 alpha*	*Armadillidium vulgare*	660	1e-113
TT-N-ST02-0033-LF			660	1e-113
TT-N-ST02-0008-LF	*Eukaryotic translation initiation factor 3 subunit 4*	*Danio rerio*	521	1e-55
TT-N-ST02-0133-LF			521	1e-55
TT-N-ST02-0092-LF	*F-box only protein 22*	*Gallus gallus*	329	6e-08
TT-N-ST02-0138-LF	*Gelsolin, cytoplasmic (Actin-depolymerizing factor) (ADF)*	*Homarus americanus*	396	3e-05
TT-N-ST02-0166-LF			396	3e-05
TT-N-ST02-0087-LF	*Helicase, lymphoid-specific isoform 2*	*Danio rerio*	408	1e-43
TT-N-ST02-0053-LF	*Mcm3-prov protein (minichromosome maintenance protein 3)*	*Xenopus laevis*	240	3e-09
TT-N-ST02-0013-LF	*Nop56 CG13849-PA, isoform A (nucleolar KKE/D repeat protein; DmNOP56)*	*Drosophila melanogaster*	348	2e-49
TT-N-ST02-0017-LF	*Nucleolin*	*Xenopus laevis*	518	3e-04
TT-N-ST02-0047-LF	*Peptidylprolyl isomerase A*	*Ictalurus punctatus*	309	8e-15
TT-N-ST02-0084-LF	*Ribosomal RNA methyltransferase*	*Aedes aegypti*	505	6e-10
TT-N-ST02-0015-LF	*RNA polymerase 1-1*	*Mus musculus*	326	2e-25
TT-N-ST02-0028-LF	*Small nuclear ribonucleoprotein D2 polypeptide 16.5kDa, isoform CRA_b*	*Homo sapiens*	313	2e-34
TT-N-ST02-0170-LF	*Small nuclear ribonucleoprotein E (snRNP-E) (Sm protein E) (Sm-E) (SmE)*	*Drosophila melanogaster*	227	2e-14
TT-N-ST02-0151-LF	*Transcription initiation factor TFIID subunit 12 (Transcription initiation factor TFIID 20/15 kDa subunits) (TAFII-20/TAFII-15)*	*Xenopus laevis*	519	1e-21

Table A35 Examples of transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of internal/external structure, motility, and transport (7 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST02-0155-LF	*ABC transporter ATP-binding protein*	*Flavobacteriales bacterium HTCC2170*	711	5e-73
TT-N-ST02-0001-LF	*Calcium-dependent chloride channel-1*	*Homo sapiens*	443	4e-11
TT-N-ST02-0009-LF	*Innexin 2*	*Penaeus monodon*	360	1e-62
TT-N-ST02-0177-LF	*Intracellular fatty acid binding protein*	*Penaeus monodon*	303	1e-156
TT-N-ST02-0071-LF	*Karyopherin (importin) alpha 4*	*Rattus norvegicus*	383	8e-08
TT-N-ST02-0193-LF	*Kinesin heavy chain*	*Loligo pealei*	570	1e-20
TT-N-ST02-0142-LF	*Ferric reductase-like protein*	*Aedes aegypti*	452	1e-28

Table A36 Examples of transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of metabolism (13 clones)

Clone No.	Transcripts	Closest Species	Size (bp)	E-value
TT-N-ST02-0029-LF	*Alcohol dehydrogenase*	*Bombyx mori*	323	3e-35
TT-N-ST02-0156-LF			680	2e-90
TT-N-ST02-0176-LF			680	2e-90
TT-N-ST02-0141-LF	*Cytochrome b*	*Penaeus monodon*	494	3e-80
TT-N-ST02-0007-LF	*Cytochrome c oxidase subunit I*	*Fenneropenaeus chinensis*	403	5e-68
TT-N-ST02-0020-LF			405	1e-66
TT-N-ST02-0117-LF			405	7e-68
TT-N-ST02-0075-LF	*GTP binding protein*	*Bombyx mori*	503	2e-70
TT-N-ST02-0040-LF	*Proteasome (prosome, macropain) 26S subunit, non-ATPase, 13*	*Tribolium castaneum*	706	2e-78
TT-N-ST02-0101-LF	*Proteasome 26S non-ATPase subunit 12*	*Tribolium castaneum*	560	1e-71
TT-N-ST02-0130-LF			560	1e-71
TT-N-ST02-0049-LF	*Proteasome 26S subunit subunit 4 ATPase CG5289-PA*	*Drosophila melanogaster*	647	1e-89
TT-N-ST02-0165-LF	*Receptor for activated protein kinase C RACK 1 isoform 1*	*Bombyx mori*	628	1e-107

Table A37 Known transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of ribosomal and rRNA (36 clones)

Clone No.	Homologue	*Species*	Size (bp)	E-value
TT-N-ST02-0139-LF	60S acidic ribosomal protein P2	*Strongylocentrotus purpuratus*	300	1e-22
TT-N-ST02-0018-LF	Ribosomal protein L12	*Ixodes scapularis*	475	1e-54
TT-N-ST02-0095-LF	Ribosomal protein L15e	*Timarcha balearica*	333	2e-22
TT-N-ST02-0002-LF	Ribosomal protein L23Ae	*Cicindela campestris*	469	5e-39
TT-N-ST02-0081-LF			469	5e-39
TT-N-ST02-0175-LF	Ribosomal protein L27	*Suberites domuncula*	371	5e-32
TT-N-ST02-0094-LF	Ribosomal protein L5e	*Carabus granulatus*	210	4e-07
TT-N-ST02-0132-LF	Ribosomal protein LP1	*Argas monolakensis*	427	3e-38
TT-N-ST02-0044-LF	Ribosomal protein S20	*Chlamys farreri*	277	2e-19
TT-N-ST02-0128-LF	Ribosomal protein S3	*Lethenteron japonicum*	538	5e-82
TT-N-ST02-0167-LF			538	5e-82
TT-N-ST02-0060-LF	16S ribosomal RNA	*Penaeus monodon*	314	1e-175
TT-N-ST02-0091-LF			273	1e-148
TT-N-ST02-0014-LF	tRNA-Ile	*Penaeus monodon*	533	0
TT-N-ST02-0019-LF			386	0
TT-N-ST02-0021-LF			550	0
TT-N-ST02-0026-LF			470	0
TT-N-ST02-0035-LF			237	1e-125
TT-N-ST02-0045-LF			548	0
TT-N-ST02-0059-LF			470	0
TT-N-ST02-0066-LF			297	1e-160
TT-N-ST02-0067-LF			571	0
TT-N-ST02-0072-LF			584	0
TT-N-ST02-0073-LF			297	1e-163

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A37 (cont.)

Clone No.	Homologue	*Species*	Size (bp)	E-value
TT-N-ST02-0080-LF	tRNA-Ile	*Penaeus monodon*	295	1e-159
TT-N-ST02-0093-LF			560	0
TT-N-ST02-0109-LF			550	0
TT-N-ST02-0120-LF			571	0
TT-N-ST02-0121-LF			360	0
TT-N-ST02-0145-LF			470	0
TT-N-ST02-0146-LF			297	1e-163
TT-N-ST02-0149-LF			523	0
TT-N-ST02-0150-LF			558	0
TT-N-ST02-0169-LF			502	0
TT-N-ST02-0173-LF			297	1e-160
TT-N-ST02-0179-LF			364	0

Table A38 Examples of transcripts from the reverse SSH library of testes of *P. monosdon* categorized as members of unidentified (hypothetical)-similar to other cDNA/DNA) (10 clones)

Clone No.	Homologue	*Species*	Size (bp)	E-value
TT-N-ST02-0122-LF	CG7876-PA	*Drosophila melanogaster*	507	5e-09
TT-N-ST02-0048-LF	Hypothetical protein	*Prochlorococcus marinus str. NATL2A*	412	7e-12
TT-N-ST02-0055-LF	Hypothetical protein	*Cytophaga hutchinsonii ATCC 33406*	702	1e-40
TT-N-ST02-0114-LF			710	4e-80
TT-N-ST02-0086-LF	Hypothetical protein	*Leptospira interrogans serovar Lai str. 56601*	507	3e-09
TT-N-ST02-0118-LF	Hypothetical protein	*Rattus norvegicus*	230	1e-18
TT-N-ST02-0157-LF	Hypothetical protein	*Schizosaccharomyces pombe 972h-*	480	8e-08
TT-N-ST02-0172-LF	Hypothetical protein	*Leptospira interrogans serovar Lai str. 56601*	507	3e-09
TT-N-ST02-0192-LF	Hypothetical protein	*Plasmodium falciparum 3D7*	309	2e-09
TT-N-ST02-0180-LF	unknown	*Homo sapiens*	542	5e-08

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table A39 Clustering of unknown transcripts of the reverse SSH library of testes of *P. monosdon* (87 clones)

Gene	Clustering	Clone
Unknown genes	CL3Contig1 (3 clones)	**ST02-0116** (273) [ST02-0079 (265), ST02-0144 (191)]
	CL5Contig1 (3 clones)	**ST02-0162** (375) [ST02-0050 (375), ST02-0181 (375)]
	CL7Contig1 (3 clones)	**ST02-0161** (260) [ST02-0103 (203), ST02-0147 (143)]
	CL10Contig1 (2 clones)	**ST02-0062** (304) [ST02-0143 (304)]
	CL11Contig1 (2 clones)	**ST02-0051** (291) [ST02-0069 (291)]
	CL12Contig1 (2 clones)	**ST02-0064** (531) [ST02-0126 (420)]
	CL14Contig1 (2 clones)	**ST02-0194** (444) [ST02-0088 (444)]
	CL15Contig1 (2 clones)	**ST02-0098** (176) [ST02-0140 (176)]
	CL19Contig1 (2 clones)	**ST02-0164** (537) [ST02-0027 (537)]
	CL20Contig1 (2 clones)	**ST02-0107** (540) [ST02-0052 (540)]
	CL23Contig1 (2 clones)	**ST02-0068** (347) [ST02-0186 (347)]
	CL26Contig1 (2 clones)	**ST02-0182** (518) [ST02-0174 (544)]
Unknown genes	Singleton (60 clones)	ST02-0003 (414), ST02-0005 (480), ST02-0006 (392), ST02-0011 (464), ST02-0012 (308), ST02-0023 (303), ST02-0025 (713), ST02-0030 (715), ST02-0031 (358), ST02-0032 (380), ST02-0034 (226), ST02-0036 (685), ST02-0037 (424), ST02-0041 (385), ST02-0042 (693), ST02-0043 (293), ST02-0046 (588), ST02-0057 (392), ST02-0058 (589), ST02-0061 (324), ST02-0074 (312), ST02-0076 (395), ST02-0077 (566), ST02-0082 (240), ST02-0083 (231), ST02-0085 (670), ST02-0096 (468), ST02-0099 (298), ST02-0100 (213), ST02-0102 (185), ST02-0104 (558), ST02-0105 (417), ST02-0110 (263), ST02-0111 (709), ST02-0112 (664), ST02-0113 (645), ST02-0115 (175), ST02-0119 (395), ST02-0125 (684), ST02-0127 (504), ST02-0129 (310), ST02-0134 (266), ST02-0135 (456), ST02-0137 (322), ST02-0148 (341), ST02-0152 (575), ST02-0153 (341), ST02-0154 (358), ST02-0158 (621), ST02-0163 (400), ST02-0168 (463), ST02-0171 (363), ST02-0178 (261), ST02-0183 (380), ST02-0184 (375), ST02-0187 (365), ST02-0189 (283), ST02-0190 (276), ST02-0191 (250), ST02-0195 (365)

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Appendix B

Table B1 Raw data and relative expression levels of *PMTST1* in testes of different groups of male *P. monodon* based on semiquantitative RT-PCR analysis

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		PMTST1	***EF-1α***			
CJ-TT	JNTT 1	1.47E+07	1.83E+07	0.82077	0.82402	0.05587
	JNTT 2	1.63E+07	1.80E+07	0.70359		
	JNTT 3	1.69E+07	1.80E+07	0.80348		
	JNTT 4	1.62E+07	1.87E+07	0.85170		
	JNTT 6	1.52E+07	1.89E+07	0.83980		
DB-TT	BUMTT 1	1.54E+07	2.00E+07	0.84886		
	BUMTT 3	1.65E+07	2.02E+07	0.85020		
	BUMTT 4	1.67E+07	2.00E+07	0.87375		
	BUMTT 5	1.58E+07	2.03E+07	0.76869	0.77285	0.07841
	BUMTT 6	1.21E+07	2.01E+07	0.81582		
	BUMTT 7	1.61E+07	2.01E+07	0.83323		
	BUMTT 8	1.58E+07	1.95E+07	0.77524		
WB-TT	RLTT 32	1.63E+07	1.99E+07	0.60268		
	RLTT 33	1.46E+07	2.08E+07	0.80167		
	RLTT 34	1.72E+07	2.14E+07	0.81263		
	RLTT 35	1.80E+07	2.12E+07	0.80371	0.86441	0.06120
	RLTT 21	1.77E+07	2.11E+07	0.90753		
	RLTT 12	1.73E+07	2.03E+07	0.93932		
	RLTT 25	1.62E+07	1.90E+07	0.86880		
	RLTT 29	1.71E+07	1.95E+07	0.80272		
1-5 DAM	BTT 2-1	2.12E+07	2.25E+07	0.94312	0.93774	0.11284
	BTT 3-1	1.64E+07	2.30E+07	0.71493		
	BTT 3-2	2.43E+07	2.38E+07	1.02068		
	BTT 5-1	2.23E+07	2.32E+07	0.96029		
	BTT 5-2	2.31E+07	2.29E+07	1.00660		
	BTT 5-3	2.33E+07	2.38E+07	0.98081		
6-9 DAM	BTT 6-1	2.25E+07	2.36E+07	0.95053	0.95862	0.05899
	BTT 6-2	2.37E+07	2.34E+07	1.01121		
	BTT 7-1	1.99E+07	2.27E+07	0.87884		
	BTT 8-1	2.11E+07	2.12E+07	0.99388		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table B1 (cont.)

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		PMTST1	*EF-1α*			
10-16 DAM	BTT 10-1	2.49E+07	2.11E+07	1.17857	1.12618	0.04479
	BTT 13-1	2.46E+07	2.22E+07	1.10832		
	BTT 14-1	2.40E+07	2.20E+07	1.08735		
	BTT 14-3	2.53E+07	2.22E+07	1.13903		
	BTT 14-4	2.24E+07	2.11E+07	1.06103		
	BTT 15-1	2.35E+07	2.09E+07	1.12750		
	BTT 16-1	2.37E+07	2.01E+07	1.18145		

Table B2 Raw data and relative expression levels of *MIPP2* in testes of different groups of male *P. monodon* based on semiquantitative RT-PCR analysis

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		MIPP2	*EF-1α*			
CJ-TT	JNTT 1	1.69E+07	1.83E+07	0.92191	0.99723	0.05316
	JNTT 2	1.86E+07	1.80E+07	1.03440		
	JNTT 3	1.86E+07	1.80E+07	1.03626		
	JNTT 4	1.79E+07	1.87E+07	0.95985		
	JNTT 6	1.96E+07	1.89E+07	1.03372		
DB-TT	BUMTT 1	1.56E+07	2.00E+07	0.77883	0.85748	0.07323
	BUMTT 2	1.54E+07	2.01E+07	0.76581		
	BUMTT 3	1.83E+07	2.02E+07	0.90928		
	BUMTT 4	1.80E+07	2.00E+07	0.89905		
	BUMTT 5	1.66E+07	2.03E+07	0.81597		
	BUMTT 6	1.68E+07	2.01E+07	0.83333		
	BUMTT 7	1.83E+07	2.01E+07	0.91217		
	BUMTT 8	1.59E+07	1.95E+07	0.81562		
	BUMTT 9	1.76E+07	1.78E+07	0.98731		
WB-TT	RLTT 32	1.79E+07	1.99E+07	0.89894	0.88757	0.07401
	RLTT 33	1.87E+07	2.08E+07	0.89831		
	RLTT 34	1.86E+07	2.14E+07	0.86770		
	RLTT 35	1.54E+07	2.12E+07	0.72716		
	RLTT 21	1.95E+07	2.11E+07	0.92479		
	RLTT 12	1.87E+07	2.03E+07	0.91877		
	RLTT 25	1.85E+07	1.90E+07	0.97019		
	RLTT 29	1.75E+07	1.95E+07	0.89471		

Table B2 (cont.)

Sample Groups		Densities of bands		Ratio of gene / ***EF-1α***	Average	STD
		MIPP2	***EF-1α***			
1-5 DAM	BTT 2-1	1.22E+07	2.07E+07	0.58918	0.65603	0.09229
	BTT 3-1	1.08E+07	2.09E+07	0.51711		
	BTT 3-2	1.39E+07	2.16E+07	0.64143		
	BTT 5-1	1.64E+07	2.13E+07	0.76950		
	BTT 5-2	1.50E+07	2.13E+07	0.70705		
	BTT 5-3	1.53E+07	2.15E+07	0.71192		
6-9 DAM	BTT 6-1	1.18E+07	2.08E+07	0.56765	0.58386	0.02735
	BTT 6-2	1.21E+07	2.05E+07	0.58909		
	BTT 7-1	1.15E+07	2.07E+07	0.55858		
	BTT 8-1	1.25E+07	2.02E+07	0.62011		
10-16 DAM	BTT 10-1	1.54E+07	2.03E+07	0.75863	0.77022	0.07924
	BTT 13-1	1.63E+07	2.09E+07	0.77889		
	BTT 14-1	1.44E+07	1.95E+07	0.73790		
	BTT 14-3	1.59E+07	1.99E+07	0.79856		
	BTT 14-4	1.17E+07	1.89E+07	0.61847		
	BTT 15-1	1.59E+07	1.86E+07	0.85171		
	BTT 16-1	1.48E+07	1.75E+07	0.84737		

Table B3 Raw data and relative expression levels of *Tra-2* in testes of different groups of male *P. monodon* based on semiquantitative RT-PCR analysis

Sample Groups		Densities of bands		Ratio of gene / ***EF-1α***	Average Density	STD
		Tra-2	***EF-1α***			
CJ-TT	JNTT 1	1.14E+06	1.83E+07	0.06202	0.06871	0.00458
	JNTT 2	1.30E+06	1.80E+07	0.07261		
	JNTT 3	1.28E+06	1.80E+07	0.07148		
	JNTT 4	1.23E+06	1.87E+07	0.06588		
	JNTT 6	1.36E+06	1.89E+07	0.07157		
DB-TT	BUMTT 1	8.27E+05	2.00E+07	0.04134	0.05891	0.01178
	BUMTT 2	8.70E+05	2.01E+07	0.04332		
	BUMTT 3	1.34E+06	2.02E+07	0.06663		
	BUMTT 4	1.29E+06	2.00E+07	0.06457		
	BUMTT 5	1.32E+06	2.03E+07	0.06486		
	BUMTT 6	1.16E+06	2.01E+07	0.05767		
	BUMTT 7	1.32E+06	2.01E+07	0.06562		

Table B3 (cont.)

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average Density	STD
		Tra-2	*EF-1α*			
DB-TT	BUMTT 8	9.72E+05	1.95E+07	0.04986		
	BUMTT 9	1.36E+06	1.78E+07	0.07631		
WB-TT	RLTT 32	1.25E+06	1.99E+07	0.06306	0.06472	0.00562
	RLTT 33	1.43E+06	2.08E+07	0.06868		
	RLTT 34	1.31E+06	2.14E+07	0.06130		
	RLTT 35	1.15E+06	2.12E+07	0.05431		
	RLTT 21	1.46E+06	2.11E+07	0.06919		
	RLTT 12	1.22E+06	2.03E+07	0.05994		
	RLTT 25	1.34E+06	1.90E+07	0.07047		
	RLTT 29	1.38E+06	1.95E+07	0.07080		
CJ-OV	JOV 5	1.15E+06	1.94E+07	0.05965	0.06632	0.00615
	JOV 6	1.24E+06	1.89E+07	0.06571		
	JOV 7	1.19E+06	1.83E+07	0.06535		
	JOV 8	1.32E+06	1.78E+07	0.07455		
WB-OV	RLOV 1	1.08E+06	1.78E+07	0.06057	0.06249	0.00668
	RLOV 33	1.03E+06	1.88E+07	0.05487		
	RLOV 6	1.21E+06	1.90E+07	0.06363		
	RLOV 31	1.35E+06	1.90E+07	0.07090		
1-5 DAM	BTT 2-1	8.62E+05	1.93E+07	0.04458	0.04719	0.00530
	BTT 3-1	7.95E+05	1.99E+07	0.03993		
	BTT 3-2	9.84E+05	2.00E+07	0.04910		
	BTT 5-1	8.98E+05	1.99E+07	0.04512		
	BTT 5-2	1.10E+06	1.98E+07	0.05558		
	BTT 5-3	9.83E+05	2.01E+07	0.04883		
6-9 DAM	BTT 6-1	8.66E+05	2.00E+07	0.04327	0.04680	0.00333
	BTT 6-2	9.92E+05	1.95E+07	0.05090		
	BTT 7-1	8.83E+05	1.96E+07	0.04514		
	BTT 8-1	9.20E+05	1.92E+07	0.04786		
10-16 DAM	BTT 10-1	1.07E+06	1.89E+07	0.05657	0.05850	0.00531
	BTT 13-1	1.15E+06	1.96E+07	0.05848		
	BTT 14-1	1.07E+06	1.84E+07	0.05810		
	BTT 14-3	1.26E+06	1.89E+07	0.06657		
	BTT 14-4	9.03E+05	1.82E+07	0.04958		
	BTT 15-1	1.02E+06	1.78E+07	0.05734		
	BTT 16-1	1.05E+06	1.67E+07	0.06289		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table B4 Raw data and relative expression levels of *prohibitin-2* in testes of different groups of male *P. monodon* based on semiquantitative RT-PCR analysis

Sample Groups		**Densities of bands**		**Ratio of gene / *EF-1α***	**Average**	**STD**
		Prohibitin-2	***EF-1α***			
CJ-TT	JNTT 1	2.36E+07	1.83E+07	1.28736	1.30465	0.07110
	JNTT 2	2.52E+07	1.80E+07	1.40302		
	JNTT 3	2.42E+07	1.80E+07	1.34586		
	JNTT 4	2.28E+07	1.87E+07	1.22038		
	JNTT 6	2.40E+07	1.89E+07	1.26662		
DB-TT	BUMTT 1	2.46E+07	2.00E+07	1.23102	1.23681	0.07199
	BUMTT 2	2.31E+07	2.01E+07	1.15243		
	BUMTT 3	2.55E+07	2.02E+07	1.26386		
	BUMTT 4	2.36E+07	2.00E+07	1.18079		
	BUMTT 5	2.35E+07	2.03E+07	1.15421		
	BUMTT 6	2.52E+07	2.01E+07	1.25442		
	BUMTT 7	2.50E+07	2.01E+07	1.24261		
	BUMTT 8	2.47E+07	1.95E+07	1.26531		
	BUMTT 9	2.47E+07	1.78E+07	1.38667		
WB-TT	RLTT 32	2.10E+07	1.99E+07	1.05407	1.05626	0.11050
	RLTT 33	1.78E+07	2.08E+07	0.85897		
	RLTT 34	2.47E+07	2.14E+07	1.15552		
	RLTT 35	2.28E+07	2.12E+07	1.07748		
	RLTT 21	2.33E+07	2.11E+07	1.10581		
	RLTT 12	1.92E+07	2.03E+07	0.94337		
	RLTT 25	2.38E+07	1.90E+07	1.24927		
	RLTT 29	1.97E+07	1.95E+07	1.00557		
CJ-OV	JOV 5	2.43E+07	1.94E+07	1.25446	1.28675	0.05157
	JOV 6	2.43E+07	1.89E+07	1.28752		
	JOV 7	2.28E+07	1.83E+07	1.24577		
	JOV 8	2.42E+07	1.78E+07	1.35925		
WB-OV	RLOV 1	2.42E+07	1.78E+07	1.36030	1.28794	0.05064
	RLOV 33	2.37E+07	1.88E+07	1.26160		
	RLOV 6	2.44E+07	1.90E+07	1.28367		
	RLOV 31	2.37E+07	1.90E+07	1.24618		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table B4 (cont.)

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		Prohibitin-2	***EF-1α***			
1-5 DAM	BTT 2-1	1.70E+07	2.07E+07	0.82317	0.91849	0.14062
	BTT 3-1	1.48E+07	2.09E+07	0.70726	0.96074	0.10646
	BTT 3-2	2.18E+07	1.95E+07	1.11886		
	BTT 5-1	1.93E+07	2.07E+07	0.93387		
	BTT 5-2	1.93E+07	2.05E+07	0.94575		
	BTT 5-3	2.00E+07	2.03E+07	0.98206		
6-9 DAM	BTT 6-1	1.58E+07	2.09E+07	0.75437	0.86087	0.09377
	BTT 6-2	1.93E+07	1.99E+07	0.97295		
	BTT 7-1	1.68E+07	2.04E+07	0.82304		
	BTT 8-1	1.86E+07	2.08E+07	0.89312		
10-16 DAM	BTT 10-1	2.03E+07	1.99E+07	1.02335	0.99387	0.08345
	BTT 13-1	1.76E+07	2.03E+07	0.86638		
	BTT 14-1	1.93E+07	1.88E+07	1.02594		
	BTT 14-3	2.19E+07	2.09E+07	1.04779		
	BTT 14-4	1.65E+07	1.87E+07	0.88231		
	BTT 15-1	1.94E+07	1.87E+07	1.03778		
	BTT 16-1	1.84E+07	1.71E+07	1.07357		

Table B5 Raw data and relative expression levels of *Trap240* in testes of different groups of male *P. monodon* based on semiquantitative RT-PCR analysis

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		Trap240	***EF-1α***			
CJ-TT	JNTT 1	1.15E+06	1.83E+07	0.06298	0.06428	0.00300
	JNTT 2	1.22E+06	1.80E+07	0.06817		
	JNTT 3	1.19E+06	1.80E+07	0.06593		
	JNTT 4	1.12E+06	1.87E+07	0.06022		
	JNTT 6	1.21E+06	1.89E+07	0.06412		
DB-TT	BUMTT 1	9.24E+05	2.00E+07	0.04621	0.05502	0.00716
	BUMTT 2	9.48E+05	2.01E+07	0.04721		
	BUMTT 3	1.22E+06	2.02E+07	0.06065		
	BUMTT 4	1.10E+06	2.00E+07	0.05517		
	BUMTT 5	1.14E+06	2.03E+07	0.05623		
	BUMTT 6	9.87E+05	2.01E+07	0.04906		
	BUMTT 7	1.25E+06	2.01E+07	0.06217		

Table B5 (cont.)

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		Trap240	***EF-1α***			
DB-TT	BUMTT 8	1.01E+06	1.95E+07	0.05165		
	BUMTT 9	1.19E+06	1.78E+07	0.06685		
WB-TT	RLTT 32	1.12E+06	1.99E+07	0.05625	0.05569	0.00280
	RLTT 33	1.20E+06	2.08E+07	0.05768		
	RLTT 34	1.24E+06	2.14E+07	0.05776		
	RLTT 35	1.08E+06	2.12E+07	0.05101		
	RLTT 21	1.14E+06	2.11E+07	0.05392		
	RLTT 12	1.18E+06	2.03E+07	0.05806		
	RLTT 25	1.09E+06	1.90E+07	0.05704		
	RLTT 29	1.05E+06	1.95E+07	0.05379		
CJ-OV	JOV 5	1.40E+06	1.94E+07	0.07215	0.07053	0.00222
	JOV 6	1.27E+06	1.89E+07	0.06726		
	JOV 7	1.31E+06	1.83E+07	0.07155		
	JOV 8	1.26E+06	1.78E+07	0.07118		
WB-OV	RLOV 1	1.12E+06	1.78E+07	0.06323	0.06780	0.00541
	RLOV 33	1.19E+06	1.88E+07	0.06319		
	RLOV 6	1.35E+06	1.90E+07	0.07105		
	RLOV 31	1.40E+06	1.90E+07	0.07372		
1-5 DAM	BTT 2-1	1.01E+06	1.93E+07	0.05240	0.05029	0.00601
	BTT 3-1	7.94E+05	1.99E+07	0.03986		
	BTT 3-2	9.85E+05	2.00E+07	0.04914		
	BTT 5-1	9.93E+05	1.99E+07	0.04989		
	BTT 5-2	1.15E+06	1.98E+07	0.05812		
	BTT 5-3	1.05E+06	2.01E+07	0.05235		
6-9 DAM	BTT 6-1	9.87E+05	2.00E+07	0.04932	0.05208	0.00321
	BTT 6-2	1.10E+06	1.95E+07	0.05654		
	BTT 7-1	9.83E+05	1.96E+07	0.05024		
	BTT 8-1	1.00E+06	1.92E+07	0.05222		
10-16 DAM	BTT 10-1	1.05E+06	1.89E+07	0.05537	0.05957	0.00572
	BTT 13-1	1.23E+06	1.96E+07	0.06287		
	BTT 14-1	1.05E+06	1.84E+07	0.05706		
	BTT 14-3	1.27E+06	1.89E+07	0.06717		
	BTT 14-4	9.13E+05	1.82E+07	0.05011		
	BTT 15-1	1.11E+06	1.78E+07	0.06224		
	BTT 16-1	1.04E+06	1.67E+07	0.06219		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table B6 Raw data and relative expression levels of *HSP70-2* in testes of different groups of male *P. monodon* based on semiquantitative RT-PCR analysis

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		HSP70-2	***EF-1α***			
CJ-TT	JNTT 1	1.69E+07	1.83E+07	0.92183	0.97048	0.07372
	JNTT 2	1.84E+07	1.80E+07	1.02663		
	JNTT 3	1.90E+07	1.80E+07	1.05876		
	JNTT 4	1.64E+07	1.87E+07	0.87879		
	JNTT 6	1.83E+07	1.89E+07	0.96637		
DB-TT	BUMTT 1	1.55E+07	2.00E+07	0.77412	0.81430	0.15826
	BUMTT 2	1.24E+07	2.01E+07	0.61962		
	BUMTT 3	1.84E+07	2.02E+07	0.91263		
	BUMTT 4	1.85E+07	2.00E+07	0.92547		
	BUMTT 5	1.35E+07	2.03E+07	0.66317		
	BUMTT 6	1.41E+07	2.01E+07	0.69933		
	BUMTT 7	1.74E+07	2.01E+07	0.86573		
	BUMTT 8	1.45E+07	1.95E+07	0.74548		
	BUMTT 9	2.00E+07	1.78E+07	1.12314		
WB-TT	RLTT 32	1.60E+07	1.99E+07	0.80633	0.82161	0.09514
	RLTT 33	1.61E+07	2.08E+07	0.77621		
	RLTT 34	1.78E+07	2.14E+07	0.83030		
	RLTT 35	1.27E+07	2.12E+07	0.59796		
	RLTT 21	1.75E+07	2.11E+07	0.82933		
	RLTT 12	1.75E+07	2.03E+07	0.86077		
	RLTT 25	1.89E+07	1.90E+07	0.99461		
	RLTT 29	1.71E+07	1.95E+07	0.87736		
CJ-OV	JOV 5	1.24E+07	1.94E+07	0.63840	0.69957	0.07183
	JOV 6	1.51E+07	1.89E+07	0.79722		
	JOV 7	1.19E+07	1.83E+07	0.65363		
	JOV 8	1.26E+07	1.78E+07	0.70902		
WB-OV	RLOV 1	1.03E+07	1.78E+07	0.57864	0.58909	0.03556
	RLOV 33	1.05E+07	1.88E+07	0.55827		
	RLOV 6	1.10E+07	1.90E+07	0.57904		
	RLOV 31	1.22E+07	1.90E+07	0.64041		

Table B6 (cont.)

Sample Groups		Densities of bands		Ratio of gene / *EF-1α*	Average	STD
		HSP70-2	***EF-1α***			
1-5 DAM	BTT 2-1	1.14E+07	1.65E+07	0.69295	0.73718	0.13009
	BTT 3-1	9.80E+06	1.78E+07	0.54945		
	BTT 3-2	1.69E+07	1.82E+07	0.92696		
	BTT 5-1	1.44E+07	1.78E+07	0.80875		
	BTT 5-2	1.37E+07	1.77E+07	0.77652		
	BTT 5-3	1.20E+07	1.79E+07	0.66844		
6-9 DAM	BTT 6-1	1.16E+07	1.75E+07	0.66534	0.65285	0.05954
	BTT 6-2	1.24E+07	1.70E+07	0.72849		
	BTT 7-1	1.03E+07	1.75E+07	0.58788		
	BTT 8-1	1.10E+07	1.74E+07	0.62967		
10-16 DAM	BTT 10-1	1.49E+07	1.68E+07	0.88651	0.87787	0.12603
	BTT 13-1	1.65E+07	1.77E+07	0.93425		
	BTT 14-1	1.41E+07	1.62E+07	0.86825		
	BTT 14-3	1.38E+07	1.55E+07	0.89002		
	BTT 14-4	1.04E+07	1.65E+07	0.62980		
	BTT 15-1	1.39E+07	1.57E+07	0.88453		
	BTT 16-1	1.57E+07	1.50E+07	1.05173		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Appendix C

Table C1 Data on expression levels of *SUMO-1* in testes of different groups of male *P. monodon* based real-time PCR analysis

Sample Name	Target Name		Tgt Cp	Ref. Cp	Ratios	Average	SD
	Targets	References	Mean	Mean	(Tgt/Ref.)		
BTT 2-1	SUMO-1	EF	24.98	19.54	0.0230	0.0237	0.0028
BTT 5-1	SUMO-1	EF	25.24	19.61	0.0202		
BTT 5-2	SUMO-1	EF	24.89	19.67	0.0268		
BTT 5-3	SUMO-1	EF	24.80	19.47	0.0248		
BTT 6-1	SUMO-1	EF	25.92	19.66	0.0131	0.0135	0.0010
BTT 7-1	SUMO-1	EF	25.54	19.45	0.0147		
BTT 8-1	SUMO-1	EF	25.91	19.63	0.0128		
BTT 10-1	SUMO-1	EF	23.54	19.19	0.0489	0.0509	0.0128
BTT 13-1	SUMO-1	EF	23.64	19.27	0.0484		
BTT 14-1	SUMO-1	EF	24.48	19.62	0.0343		
BTT 14-2	SUMO-1	EF	24.34	19.76	0.0417		
BTT 14-3	SUMO-1	EF	24.12	19.91	0.0539		
BTT 15-1	SUMO-1	EF	22.89	19.16	0.0754		
BTT 16-1	SUMO-1	EF	23.66	19.44	0.0536		
JNTT 1	SUMO-1	EF	23.57	18.99	0.0418	0.0515	0.0105
JNTT 3	SUMO-1	EF	23.24	19.24	0.0627		
JNTT 4	SUMO-1	EF	24.34	20.02	0.0501		
BUMTT 1	SUMO-1	EF	24.46	19.55	0.0334	0.0263	0.0081
BUMTT 2	SUMO-1	EF	25.91	20.05	0.0172		
BUMTT 4	SUMO-1	EF	23.74	18.90	0.0348		
BUMTT 5	SUMO-1	EF	24.27	19.27	0.0312		
BUMTT 6	SUMO-1	EF	25.78	19.85	0.0164		
BUMTT 8	SUMO-1	EF	25.47	20.14	0.0248		
RLTT 32	SUMO-1	EF	23.70	20.11	0.0831	0.0618	0.0176
RLTT 33	SUMO-1	EF	24.19	20.64	0.0854		
RLTT 34	SUMO-1	EF	23.41	19.19	0.0537		
RLTT 12	SUMO-1	EF	23.10	18.88	0.0539		
RLTT 25	SUMO-1	EF	23.69	19.28	0.0471		
RLTT 29	SUMO-1	EF	23.55	19.16	0.0476		
JOV 5	SUMO-1	EF	19.98	15.28	0.0385	0.0380	0.0043
JOV 6	SUMO-1	EF	21.21	16.70	0.0439		
JOV 7	SUMO-1	EF	20.77	15.90	0.0343		
JOV 8	SUMO-1	EF	19.72	14.90	0.0354		
RLOV 33	SUMO-1	EF	20.56	16.30	0.0521	0.0456	0.0057
RLOV 6	SUMO-1	EF	19.53	15.00	0.0433		
RLOV 31	SUMO-1	EF	19.55	14.95	0.0414		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C2 Data on expression levels of *cyclophilin A* (*CYA*)in testes of different groups of male *P. monodon* based real-time PCR analysis

Sample Name	Target Name		Tgt Cp	Ref. Cp	Ratios	Average	SD
	Targets	References	Mean	Mean	(Tgt/Ref.)		
BTT 2-1	CYA	EF	21.86	19.54	0.2006	0.1585	0.0457
BTT 3-2	CYA	EF	21.95	19.59	0.1954		
BTT 5-1	CYA	EF	22.70	19.61	0.1178		
BTT 5-2	CYA	EF	22.73	19.67	0.1201		
BTT 6-1	CYA	EF	23.96	19.66	0.0509	0.0618	0.0217
BTT 7-1	CYA	EF	22.97	19.45	0.0868		
BTT 8-1	CYA	EF	24.02	19.63	0.0477		
BTT 10-1	CYA	EF	21.64	19.19	0.1821	0.2170	0.0814
BTT 13-1	CYA	EF	21.17	19.27	0.2688		
BTT 14-1	CYA	EF	22.13	19.62	0.1755		
BTT 14-3	CYA	EF	21.52	19.91	0.3283		
BTT 14-4	CYA	EF	22.83	19.58	0.1048		
BTT 15-1	CYA	EF	20.91	19.16	0.2962		
BTT 16-1	CYA	EF	22.06	19.44	0.1630		
JNTT 1	CYA	EF	21.27	18.99	0.2058	0.2618	0.0424
JNTT 2	CYA	EF	21.19	19.49	0.3086		
JNTT 3	CYA	EF	21.17	19.24	0.2628		
JNTT 4	CYA	EF	21.91	20.02	0.2700		
BUMTT 2	CYA	EF	24.04	20.05	0.0626	0.0810	0.0250
BUMTT 4	CYA	EF	22.07	18.90	0.1113		
BUMTT 5	CYA	EF	22.52	19.27	0.1052		
BUMTT 6	CYA	EF	23.89	19.85	0.0606		
BUMTT 8	CYA	EF	23.56	20.14	0.0652		
RLTT 32	CYA	EF	22.00	20.11	0.2700	0.2202	0.0664
RLTT 33	CYA	EF	22.33	20.64	0.3101		
RLTT 12	CYA	EF	21.57	18.88	0.1553		
RLTT 25	CYA	EF	21.78	19.28	0.1771		
RLTT 29	CYA	EF	21.57	19.16	0.1885		
JOV 5	CYA	EF	16.14	15.28	0.5500	0.4363	0.1445
JOV 6	CYA	EF	18.39	16.70	0.3103		
JOV 7	CYA	EF	16.71	15.90	0.5723		
JOV 8	CYA	EF	16.58	14.90	0.3126		
RLOV 1	CYA	EF	17.59	17.52	0.9557	0.8062	0.1062
RLOV 33	CYA	EF	16.66	16.30	0.7781		
RLOV 6	CYA	EF	15.35	15.00	0.7861		
RLOV 31	CYA	EF	15.45	14.95	0.7047		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C3 Data on expression levels of *sSaposin* in testes of different groups of male *P. monodon* based real-time PCR analysis

Sample Name	Target Name		Tgt Cp	Ref. Cp	Ratios	Average	SD
	Targets	References	Mean	Mean	(Tgt/Ref.)		
BTT 2-1	Saposin	EF	22.59	19.54	0.1205	0.0900	0.0287
BTT 3-1	Saposin	EF	23.41	19.72	0.0545		
BTT 5-1	Saposin	EF	23.00	19.61	0.0954		
BTT 5-2	Saposin	EF	23.58	19.67	0.0666		
BTT 5-3	Saposin	EF	22.11	19.47	0.1131		
BTT 6-2	Saposin	EF	22.66	19.11	0.0857	0.0898	0.0094
BTT 7-1	Saposin	EF	23.03	19.45	0.0832		
BTT 8-1	Saposin	EF	22.94	19.63	0.1005		
BTT 10-1	Saposin	EF	22.38	19.19	0.1091	0.1218	0.0318
BTT 13-1	Saposin	EF	22.27	19.27	0.1250		
BTT 14-1	Saposin	EF	22.90	19.62	0.1024		
BTT 14-4	Saposin	EF	22.22	19.58	0.1134		
BTT 15-1	Saposin	EF	21.60	19.16	0.1837		
BTT 16-1	Saposin	EF	22.81	19.44	0.0970		
JNTT 1	Saposin	EF	21.93	18.99	0.1299	0.1502	0.0240
JNTT 3	Saposin	EF	22.00	19.24	0.1477		
JNTT 4	Saposin	EF	22.46	20.02	0.1845		
JNTT 6	Saposin	EF	22.29	19.44	0.1388		
BUMTT 2	Saposin	EF	22.69	20.05	0.1605	0.1837	0.0214
BUMTT 4	Saposin	EF	21.51	18.90	0.1633		
BUMTT 5	Saposin	EF	21.52	19.27	0.2100		
BUMTT 7	Saposin	EF	23.17	20.76	0.1882		
BUMTT 8	Saposin	EF	21.97	20.14	0.1965		
RLTT 34	Saposin	EF	21.98	19.19	0.1452	0.1626	0.0338
RLTT 35	Saposin	EF	21.69	19.44	0.2104		
RLTT 12	Saposin	EF	21.88	18.88	0.1255		
RLTT 25	Saposin	EF	21.73	19.28	0.1830		
RLTT 29	Saposin	EF	21.91	19.16	0.1487		
JOV 5	Saposin	EF	20.65	15.28	0.0242	0.0288	0.0069
JOV 6	Saposin	EF	21.39	16.70	0.0387		
JOV 7	Saposin	EF	21.27	15.90	0.0242		
JOV 8	Saposin	EF	20.06	14.90	0.0279		
RLOV 1	Saposin	EF	22.04	17.52	0.0436	0.0396	0.0064
RLOV 33	Saposin	EF	20.74	16.30	0.0460		
RLOV 6	Saposin	EF	19.77	15.00	0.0367		
RLOV 31	Saposin	EF	19.91	14.95	0.0321		

Table C4 Data on expression levels of *spermatogonial stem-cell renewal factor* in testes of different groups of male *P. monodon* based real-time PCR analysis

Sample Name	Target Name		Tgt Cp	Ref. Cp	Ratios	Average	SD
	Targets	References	Mean	Mean	(Tgt/Ref.)		
BTT 2-1	Sperm	EF	27.90	19.54	3.04E-03	0.0029	0.0006
BTT 3-1	Sperm	EF	28.13	19.72	2.08E-03		
BTT 3-2	Sperm	EF	27.53	19.59	4.08E-03		
BTT 5-1	Sperm	EF	28.22	19.61	2.55E-03		
BTT 5-2	Sperm	EF	27.75	19.67	3.68E-03		
BTT 5-3	Sperm	EF	27.39	19.47	2.91E-03		
BTT 6-2	Sperm	EF	27.76	19.11	2.50E-03	0.0026	0.0002
BTT 7-1	Sperm	EF	27.94	19.45	2.77E-03		
BTT 8-1	Sperm	EF	28.31	19.63	2.44E-03		
BTT 10-1	Sperm	EF	26.91	19.19	4.74E-03	0.0049	0.0010
BTT 13-1	Sperm	EF	26.84	19.27	5.28E-03		
BTT 14-1	Sperm	EF	27.79	19.62	3.46E-03		
BTT 14-3	Sperm	EF	27.16	19.91	6.56E-03		
BTT 15-1	Sperm	EF	27.01	19.16	4.33E-03		
BTT 16-1	Sperm	EF	27.11	19.44	4.90E-03		
JNTT 2	Sperm	EF	25.82	19.49	1.24E-02	0.0108	0.0015
JNTT 3	Sperm	EF	25.90	19.24	9.87E-03		
JNTT 4	Sperm	EF	26.45	20.02	1.17E-02		
JNTT 6	Sperm	EF	26.18	19.44	9.34E-03		
BUMTT 1	Sperm	EF	26.89	19.55	6.17E-03	0.0069	0.0018
BUMTT 2	Sperm	EF	27.70	20.05	4.95E-03		
BUMTT 4	Sperm	EF	25.94	18.90	7.61E-03		
BUMTT 5	Sperm	EF	26.69	19.27	5.84E-03		
BUMTT 7	Sperm	EF	27.38	20.76	1.02E-02		
BUMTT 8	Sperm	EF	26.90	20.14	6.43E-03		
RLTT 33	Sperm	EF	26.62	20.64	1.58E-02	0.0115	0.0027
RLTT 35	Sperm	EF	26.10	19.44	9.90E-03		
RLTT 12	Sperm	EF	25.21	18.88	1.25E-02		
RLTT 25	Sperm	EF	25.99	19.28	9.56E-03		
RLTT 29	Sperm	EF	25.85	19.16	9.72E-03		
JOV 5	Sperm	EF	27.89	15.28	1.60E-04	0.0003	0.0002
JOV 6	Sperm	EF	27.51	16.70	5.55E-04		
JOV 7	Sperm	EF	27.81	15.90	2.61E-04		
RLOV 33	Sperm	EF	27.20	16.30	5.23E-04	0.0008	0.0004
RLOV 6	Sperm	EF	25.61	15.00	6.37E-04		
RLOV 31	Sperm	EF	24.68	14.95	1.18E-03		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C5 Quantitative expression level data of *spermatogonial stem-cell renewal factor* in testes of different groups of male *P. monodon* based on real-time PCR analysis

Sample Name	Type	CP	Concentration	Average of each sample	Average of each group	SD
BTT 2-1	Sperm	27.9	8.64E+02	8.65E+02	958.33	276.96
BTT 2-1	Sperm	27.9	8.65E+02			
BTT 3-1	Sperm	28.17	6.67E+02	6.97E+02		
BTT 3-1	Sperm	28.09	7.26E+02			
BTT 3-2	Sperm	27.63	1.11E+03	1.21E+03		
BTT 3-2	Sperm	27.43	1.31E+03			
BTT 5-1	Sperm	28.14	6.89E+02	6.37E+02		
BTT 5-1	Sperm	28.31	5.84E+02			
BTT 5-2	Sperm	27.77	9.75E+02	9.88E+02		
BTT 5-2	Sperm	27.74	1.00E+03			
BTT 5-3	Sperm	27.33	1.43E+03	1.36E+03		
BTT 5-3	Sperm	27.46	1.28E+03			
BTT 6-2	Sperm	27.79	9.56E+02	9.88E+02	804.17	189.39
BTT 6-2	Sperm	27.72	1.02E+03			
BTT 7-1	Sperm	27.98	7.99E+02	8.32E+02		
BTT 7-1	Sperm	27.9	8.65E+02			
BTT 8-1	Sperm	28.15	6.83E+02	5.93E+02		
BTT 8-1	Sperm	28.46	5.02E+02			
BTT 10-1	Sperm	26.86	2.09E+03	2.02E+03	1717.33	402.87
BTT 10-1	Sperm	26.95	1.94E+03			
BTT 13-1	Sperm	26.81	2.16E+03	2.12E+03		
BTT 13-1	Sperm	26.87	2.08E+03			
BTT 14-1	Sperm	27.72	1.02E+03	9.54E+02		
BTT 14-1	Sperm	27.87	8.88E+02			
BTT 14-3	Sperm	27.06	1.79E+03	1.65E+03		
BTT 14-3	Sperm	27.27	1.50E+03			
BTT 15-1	Sperm	26.99	1.89E+03	1.86E+03		
BTT 15-1	Sperm	27.02	1.83E+03			
BTT 16-1	Sperm	27.1	1.72E+03	1.71E+03		
BTT 16-1	Sperm	27.12	1.70E+03			
RLTT 32	Sperm	25.72	4.73E+03	4.59E+03	3902.50	470.57
RLTT 32	Sperm	25.81	4.45E+03			
RLTT 33	Sperm	26.13	3.54E+03	3.42E+03		
RLTT 33	Sperm	26.23	3.30E+03			
RLTT 35	Sperm	26.03	3.81E+03	3.63E+03		
RLTT 35	Sperm	26.17	3.45E+03			
RLTT 21	Sperm	26.11	3.59E+03	3.53E+03		
RLTT 21	Sperm	26.16	3.46E+03			

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C5 (cont.)

Sample Name	Type	CP	Concentration	Average of each sample	Average of each group	SD
RLTT 25	Sperm	25.93	4.09E+03	3.92E+03		
RLTT 25	Sperm	26.05	3.74E+03			
RLTT 29	Sperm	25.82	4.42E+03	4.34E+03		
RLTT 29	Sperm	25.87	4.25E+03			
BUMTT 1	Sperm	26.80	2.19E+03	2.04E+03	2126.36	683.00
BUMTT 1	Sperm	26.99	1.89E+03			
BUMTT 2	Sperm	27.76	9.79E+02	1.03E+03		
BUMTT 2	Sperm	27.65	1.09E+03			
BUMTT 3	Sperm	26.09	2.91E+03	2.93E+03		
BUMTT 3	Sperm	26.07	2.94E+03			
BUMTT 4	Sperm	26.03	3.03E+03	2.97E+03		
BUMTT 4	Sperm	26.09	2.90E+03			
BUMTT 5	Sperm	26.67	2.41E+03	2.38E+03		
BUMTT 5	Sperm	26.71	2.34E+03			
BUMTT 7	Sperm	27.17	1.63E+03	1.53E+03		
BUMTT 7	Sperm	27.33	1.42E+03			
BUMTT 8	Sperm	26.83	2.14E+03	2.02E+03		
BUMTT 8	Sperm	26.98	1.90E+03			
JNTT 1	Sperm	26.03	3.79E+03	3.65E+03	3693.00	592.36
JNTT 1	Sperm	26.14	3.51E+03			
JNTT 2	Sperm	25.80	4.48E+03	4.40E+03		
JNTT 2	Sperm	25.85	4.32E+03			
JNTT 3	Sperm	25.85	4.31E+03	4.16E+03		
JNTT 3	Sperm	25.95	4.01E+03			
JNTT 4	Sperm	26.43	2.88E+03	2.84E+03		
JNTT 4	Sperm	26.47	2.80E+03			
JNTT 6	Sperm	26.16	3.48E+03	3.42E+03		
JNTT 6	Sperm	26.21	3.35E+03			
JOV 5	Sperm	27.80	9.48E+02	8.77E+02	1014.50	177.12
JOV 5	Sperm	27.98	8.05E+02			
JOV 6	Sperm	27.45	1.29E+03	1.23E+03		
JOV 6	Sperm	27.57	1.16E+03			
JOV 7	Sperm	27.78	9.64E+02	9.42E+02		
JOV 7	Sperm	27.83	9.20E+02			
RLOV 1	Sperm	25.48	5.15E+03	5.59E+03	7023.33	1636.53
RLOV 1	Sperm	25.26	6.02E+03			
RLOV 6	Sperm	25.63	6.35E+03	6.44E+03		
RLOV 6	Sperm	25.60	6.53E+03			
RLOV 31	Sperm	24.69	9.00E+03	9.05E+03		
RLOV 31	Sperm	24.67	9.09E+03			

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C6 Data on expression levels of *Dmc1* in testes of different groups of male *P. monodon* based real-time PCR analysis

Sample Name	Target Name		Tgt Cp	Ref. Cp	Ratios	Average	SD
	Targets	References	Mean	Mean	(Tgt/Ref.)		
BTT 2-1	Dmc1	EF	31.25	18.8	1.78E-04	0.000312	0.000126
BTT 5-1	Dmc1	EF	30.14	18.57	3.29E-04		
BTT 5-3	Dmc1	EF	28.88	18.16	4.28E-04		
BTT 6-1	Dmc1	EF	31.32	18.83	1.73E-04	0.000166	0.000061
BTT 6-2	Dmc1	EF	31.08	18.68	1.33E-04		
BTT 7-1	Dmc1	EF	30.47	18.49	2.48E-04		
BTT 8-1	Dmc1	EF	31.95	18.79	1.09E-04		
BTT 10-1	Dmc1	EF	29.01	18.81	8.49E-04	0.001039	0.000313
BTT 14-1	Dmc1	EF	29.39	19.01	7.50E-04		
BTT 14-3	Dmc1	EF	27.92	18.75	1.24E-03		
BTT 14-4	Dmc1	EF	28.41	18.24	8.67E-04		
BTT 16-1	Dmc1	EF	27.94	18.55	1.49E-03		
JNTT 1	Dmc1	EF	27.35	17.97	1.51E-03	0.001820	0.000320
JNTT 3	Dmc1	EF	27.3	18.43	2.15E-03		
JNTT 6	Dmc1	EF	27.84	18.72	1.80E-03		
BUMTT 1	Dmc1	EF	30.18	18.76	3.66E-04	0.000627	0.000299
BUMTT 4	Dmc1	EF	28.05	17.84	8.44E-04		
BUMTT 5	Dmc1	EF	30.02	18.57	3.60E-04		
BUMTT 7	Dmc1	EF	29.46	20.04	1.03E-03		
BUMTT 8	Dmc1	EF	29.08	18.69	5.34E-04		
RLTT 32	Dmc1	EF	27.14	19.15	2.80E-03	0.002123	0.000476
RLTT 33	Dmc1	EF	27.97	19.81	2.47E-03		
RLTT 34	Dmc1	EF	27.61	18.28	1.56E-03		
RLTT 21	Dmc1	EF	27.7	18.65	1.89E-03		
RLTT 12	Dmc1	EF	27.26	18.09	1.73E-03		
RLTT 29	Dmc1	EF	27.04	18.27	2.29E-03		
JOV 5	Dmc1	EF	31.3	14.39	8.13E-06	0.000013	0.000006
JOV 7	Dmc1	EF	30.53	14.88	1.94E-05		
JOV 8	Dmc1	EF	29.66	13.67	1.20E-05		
RLOV 1	Dmc1	EF	31.08	16.54	8.08E-07	0.000004	0.000003
RLOV 6	Dmc1	EF	31.8	14.18	4.98E-06		
RLOV 31	Dmc1	EF	31.44	14.03	5.73E-06		

Table C7 Quantitative expression level of *Dmc1* in testes of different groups of male *P. monodon* based on real-time PCR analysis

Sample Name	Type	CP	Concentration	Average of each sample	Average of each group	SD
BTT 2-1	Dmc1	31.44	1.03E+02	1.18E+02	282.50	171.31
BTT 2-1	Dmc1	31.06	1.32E+02			
BTT 5-1	Dmc1	30.12	2.42E+02	2.39E+02		
BTT 5-1	Dmc1	30.17	2.35E+02			
BTT 5-3	Dmc1	29.15	4.74E+02	4.92E+02		
BTT 5-3	Dmc1	29.05	5.09E+02			
BTT 6-1	Dmc1	31.21	1.26E+02	1.18E+02	128.60	46.79
BTT 6-1	Dmc1	31.44	1.09E+02			
BTT 6-2	Dmc1	31.18	1.22E+02	1.26E+02		
BTT 6-2	Dmc1	31.08	1.30E+02			
BTT 7-1	Dmc1	30.33	2.11E+02	1.94E+02		
BTT 7-1	Dmc1	30.61	1.76E+02			
BTT 8-1	Dmc1	31.58	9.49E+01	7.74E+01		
BTT 8-1	Dmc1	32.33	5.99E+01			
BTT 13-1	Dmc1	26.86	2.20E+03	2.26E+03	1461.17	617.65
BTT 13-1	Dmc1	26.79	2.31E+03			
BTT 14-3	Dmc1	27.98	1.03E+03	1.07E+03		
BTT 14-3	Dmc1	27.87	1.11E+03			
BTT 16-1	Dmc1	28.03	9.97E+02	1.06E+03		
BTT 16-1	Dmc1	27.86	1.12E+03			
JNTT 1	Dmc1	27.45	1.48E+03	1.59E+03	1266.75	455.82
JNTT 1	Dmc1	27.24	1.70E+03			
JNTT 3	Dmc1	27.55	1.38E+03	1.67E+03		
JNTT 3	Dmc1	27.04	1.95E+03			
JNTT 4	Dmc1	28.71	6.63E+02	6.72E+02		
JNTT 4	Dmc1	28.67	6.81E+02			
JNTT 6	Dmc1	27.87	1.11E+03	1.14E+03		
JNTT 6	Dmc1	27.80	1.17E+03			
BUMTT 1	Dmc1	30.08	2.49E+02	2.34E+02	352.13	121.98
BUMTT 1	Dmc1	30.27	2.19E+02			
BUMTT 5	Dmc1	29.54	3.58E+02	2.74E+02		
BUMTT 5	Dmc1	30.49	1.90E+02			
BUMTT 7	Dmc1	29.22	4.47E+02	4.10E+02		
BUMTT 7	Dmc1	29.48	3.73E+02			
BUMTT 8	Dmc1	28.99	5.20E+02	4.91E+02		
BUMTT 8	Dmc1	29.17	4.61E+02			
RLTT 32	Dmc1	27.10	1.87E+03	1.83E+03	1701.00	226.88
RLTT 32	Dmc1	27.18	1.78E+03			

Table C7 (cont.)

Sample Name	Type	CP	Concentration	Average of each sample	Average of each group	SD
RLTT 33	Dmc1	28.08	1.68E+03	1.72E+03		
RLTT 33	Dmc1	28.00	1.76E+03			
RLTT 34	Dmc1	27.53	1.40E+03	1.33E+03		
RLTT 34	Dmc1	27.69	1.26E+03			
RLTT 12	Dmc1	27.24	1.71E+03	1.68E+03		
RLTT 12	Dmc1	27.29	1.65E+03			
RLTT 29	Dmc1	27.10	1.87E+03	1.95E+03		
RLTT 29	Dmc1	26.98	2.03E+03			
JOV 5	Dmc1	31.38	1.07E+02	1.14E+02	211.17	101.75
JOV 5	Dmc1	31.21	1.20E+02			
JOV 7	Dmc1	30.21	2.29E+02	1.90E+02		
JOV 7	Dmc1	30.85	1.50E+02			
JOV 8	Dmc1	29.64	3.36E+02	3.31E+02		
JOV 8	Dmc1	29.68	3.25E+02			
RLOV 1	Dmc1	30.90	2.53E+02	2.27E+02	205.33	33.83
RLOV 1	Dmc1	31.26	2.01E+02			
RLOV 6	Dmc1	31.48	1.75E+02	1.69E+02		
RLOV 6	Dmc1	31.59	1.63E+02			
RLOV 31	Dmc1	31.03	2.32E+02	2.20E+02		
RLOV 31	Dmc1	31.21	2.08E+02			

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C8 Data on expression levels of *PGMRC1* in dopamine-treated testes of juvenile male *P. monodon* based real-time PCR analysis

Sample Name	PGMRC1			EF			[log(Qty of PGMRC1)/log(Qty of EF)]*100	Average	SD
	Ct	Qty	log (Qty)	Ct	Qty	log (Qty)			
JM25	19.74	2.53E+02	2.4032	17.59	3.25E+06	6.5119	3.69E+01	38.74	2.06
JM26	19.12	6.57E+02	2.8174	16.9	7.54E+06	6.8774	4.10E+01		
JM27	19.86	2.13E+02	2.3293	18.42	1.18E+06	6.0719	3.84E+01		
3 hr post injection									
NSM3-2	19.75	2.52E+02	2.4012	16.38	1.42E+07	7.1523	3.36E+01	33.99	3.16
NSM3-5	19.25	5.37E+02	2.7302	16.08	2.05E+07	7.3118	3.73E+01		
NSM3-7	20.03	1.63E+02	2.2133	16.43	1.34E+07	7.1271	3.11E+01		
DA10-6M3-3	19.29	5.08E+02	2.7055	17.93	2.14E+06	6.3304	4.27E+01	44.84	3.19
DA10-6M3-4	19.27	5.21E+02	2.7169	19.31	3.98E+05	5.5999	4.85E+01		
DA10-6M3-5	20.3	1.08E+02	2.0321	21.01	4.96E+04	4.6955	4.33E+01		
6 hr post injection									
NSM6-1	19.11	6.66E+02	2.8236	18.81	7.34E+05	5.8657	4.81E+01	46.64	2.91
NSM6-3	19.15	6.24E+02	2.7950	19	5.81E+05	5.7642	4.85E+01		
NSM6-4	18.65	1.35E+03	3.1303	16.23	1.71E+07	7.2330	4.33E+01		
DA10-6M6-2	18.51	1.66E+03	3.2201	16.26	1.66E+07	7.2201	4.46E+01	43.08	2.05
DA10-6M6-3	18.63	1.39E+03	3.1430	16.36	1.45E+07	7.1614	4.39E+01		
DA10-6M6-4	18.92	8.88E+02	2.9484	16.23	1.72E+07	7.2355	4.07E+01		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C8 (cont.)

Sample Name	PGMRC1			EF			[log(Qty of PGMRC1)/log(Qty of EF)]*100	Average	SD
	Ct	Qty	log (Qty)	Ct	Qty	log (Qty)			
12 hr post injection									
NSM12-2	18.47	1.76E+03	3.2455	15.49	4.24E+07	7.6274	4.26E+01	42.71	2.29
NSM12-3	20.63	6.61E+01	1.8199	21.4	3.11E+04	4.4928	4.05E+01		
NSM12-5	19.23	5.56E+02	2.7451	18.39	1.23E+06	6.0899	4.51E+01		
DA10-6M12-2	19.24	5.47E+02	2.7378	17.52	3.54E+06	6.5490	4.18E+01	40.63	1.60
DA10-6M12-3	19.36	4.57E+02	2.6596	17.72	2.77E+06	6.4425	4.13E+01		
DA10-6M12-5	20.15	1.36E+02	2.1325	19.51	3.12E+05	5.4942	3.88E+01		
24 hr post injection									
NSM24-2	18.62	1.42E+03	3.1523	15.99	2.28E+07	7.3579	4.28E+01	30.02	11.99
NSM24-3	24.29	2.48E-01	-0.6051	33.92	7.08E-03	-2.1500	2.81E+01		
NSM24-4	24.16	3.00E-01	-0.5230	35.03	1.82E-03	-2.7399	1.91E+01		
DA10-6M24-1	21.23	2.61E+01	1.4161	22.06	1.38E+04	4.1399	3.42E+01	23.46	15.20
DA10-6M24-2	22.7	2.78E+00	0.4446	23.27	3.15E+03	3.4983	1.27E+01		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C9 Data on expression levels of *Dmc1* in dopamine-treated testes of juvenile male *P. monodon* based real-time PCR analysis

Sample Name	Dmc1			EF			[log(Qty of Dmc1)/log(Qty of EF)]*100	Average	SD
	Ct	Qty	log (Qty)	Ct	Qty	log (Qty)			
JM23	25.09	3.96E+04	4.5980	17.58	2.05E+06	6.3107	7.29E+01	72.63	0.69
JM25	25.28	3.51E+04	4.5450	17.90	1.63E+06	6.2109	7.32E+01		
JM26	24.50	5.90E+04	4.7709	16.40	4.36E+06	6.6390	7.19E+01		
3 hr post injection									
NSM3-2	26.74	1.31E+04	4.1182	18.62	1.01E+06	6.0063	6.86E+01	72.94	3.88
NSM3-3	25.48	3.07E+04	4.4865	18.51	1.09E+06	6.0374	7.43E+01		
NSM3-5	25.35	3.36E+04	4.5260	18.79	9.10E+05	5.9588	7.60E+01		
DA 10-6M3-2	30.71	9.20E+02	2.9638	19.07	7.65E+05	5.8834	5.04E+01	69.80	17.25
DA 10-6M3-3	24.92	4.72E+04	4.6743	18.04	1.49E+06	6.1732	7.57E+01		
DA 10-6M3-4	22.43	2.99E+05	5.4761	16.64	3.74E+06	6.5729	8.33E+01		
6 hr post injection									
NSM6-1	25.66	2.71E+04	4.4332	17.66	1.91E+06	6.2810	7.06E+01	72.07	3.93
NSM6-3	26.47	1.57E+04	4.1956	18.40	1.18E+06	6.0719	6.91E+01		
NSM6-4	24.52	5.84E+04	4.7663	17.85	1.69E+06	6.2279	7.65E+01		
DA 10-6M6-2	24.59	5.56E+04	4.7450	17.90	1.65E+06	6.2162	7.63E+01	68.51	8.78
DA 10-6M6-3	25.55	2.93E+04	4.4663	17.40	2.31E+06	6.3636	7.02E+01		
DA 10-6M6-4	27.84	6.26E+03	3.7963	17.18	2.71E+06	6.4330	5.90E+01		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Table C9 (cont.)

Sample Name	Dmc1			EF			[log(Qty of Dmc1)/log(Qty of EF)]*100	Average	SD
	Ct	Qty	log (Qty)	Ct	Qty	log (Qty)			
12 hr post injection									
NSM12-2	23.68	1.05E+05	5.0223	16.70	3.56E+06	6.5508	7.67E+01	65.37	13.07
NSM12-3	26.45	1.60E+04	4.2030	18.15	1.40E+06	6.1461	6.84E+01		
NSM12-4	30.44	1.10E+03	3.0403	18.72	9.03E+05	5.9556	5.10E+01		
DA 10-6M12-2	23.58	1.11E+05	5.0440	18.05	1.50E+06	6.1746	8.17E+01	81.27	4.70
DA 10-6M12-3	24.67	5.54E+04	4.7435	17.94	1.63E+06	6.2109	7.64E+01		
10-6M12-4	21.84	3.53E+05	5.5482	16.99	2.96E+06	6.4706	8.57E+01		
24 hr post injection									
NSM24-1	32.78	1.21E+02	2.0830	33.06	8.21E+01	1.9142	1.09E+02	100.39	14.52
NSM24-3	29.70	2.03E+03	3.3067	30.20	1.10E+03	3.0414	1.09E+02		
NSM24-4	35.34	4.46E+01	1.6488	33.15	9.37E+01	1.9718	8.36E+01		
DA 10-6M24-1	28.40	4.30E+03	3.6337	19.37	6.22E+05	5.7936	6.27E+01	68.79	8.86
DA 10-6M24-2	28.79	3.33E+03	3.5226	20.61	2.78E+05	5.4441	6.47E+01		
DA 10-6M24-3	25.54	3.39E+04	4.5306	19.57	5.47E+05	5.7378	7.90E+01		

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

Biography

Miss Rungnapa Leelatanawit was born on November 7, 1978 in Bangkok. She graduated with the degree of Bachelor of Science from Department of Industrial Technology, Silapakorn University in 2000 and the degree of Master of Science from the program in Biotechnology, Chulalongkorn University in 2003. She has studied for the Degree of Doctor of Philosophy in Biotechnology, Chulalongkorn University since 2004.

Publications from this thesis

International publications

1. **Leelatanawit, R.**, Sittikankaew, K., Yocawibun, P., Klinbunga, S., Roytrakul, S., Aoki, T., Hirono, I., and Menasveta, P. Identification, Characterization and Expression of Sex-Related Genes in Testes of the Giant Tiger Shrimp *Penaeus monodon*. Comp. Biochem. Physiol. (in press).

2. **Leelatanawit, R.**, Klinbunga, S., Aoki, T., Hirono, I., Valyasevi, R., and Menasveta, P. Suppression Subtractive Hybridization (SSH) for Isolation and Characterization of Genes Related to Testicular Development of the Giant Tiger Shrimp *Penaeus monodon*. BMB Reports (accepted).

Proceeding

1. Klinbunga, S., Kamnamtong, B., **Leelatanawit, R.**, Preechaphol, R. and Menasveta, P. (2006). Isolation of sex-related genes in ovaries and testes of the giant tiger shrimp (*Penaeus monodon*). In: *Comparative Endocrinology and Biodiversity in Asia and Oceania.* (Tangpraprutgul, T., Malaivijitnond, S., hanchao, C. and Kitana, N., eds). pp. 143-147.

Conferences

1. **Leelatanawit, R.**, Sittikankeaw, K., Pasertluk, S., Thamniemdee, N., Klinbunga, S., Tassanakajon, A. and Menasveta, P. (2005). Expressed sequence tag (EST) analysis of genes expressed in ovaries of the giant tiger shrimp (*Penaeus monodon*). 31st

Congress on Science and Technology of Thailand, 18–20 October 2005, Nakornratchasrima, Thailand (Oral presentation).

2. **Leelatanawit, R.**, Sittikankeaw, K., Pasertluk, S., Thamniemdee, N., Klinbunga, S., Tassanakajon, A.and Menasveta (2005). Expressed Sequence Tag (EST) Analysis of Genes Expressed in Ovaries, Testes and Heart of the Giant Tiger Shrimp (*Penaeus monodon*). International Shrimp Symposium, BIOTHAILAND 2005, 2-5 November 2005, Bangkok, Thailand (Oral presentation).

3. **Leelatanawit, R.**, Sittikankeaw, K., Klinbunga, S. and Menasveta, P. (2007). Isolation of Genes Involving Testicular Development of the Giant Tiger Shrimp (*Penaeus monodon*) by EST and SSH Analyses. 6th National Symposium on Marine Shrimps, 29–30 March 2007, National Center for Genetic Engineering and Biotechnology (BIOTEC), NSTDA, Thailand (Oral presentation).

4. **Leelatanawit, R.**, Sittikankeaw, K., Klinbunga, S. and Menasveta, P. (2007). Isolation of genes involving testicular development of the Giant Tiger Shrimp (*Penaeus monodon*) by EST and SSH analyses. 33st Congress on Science and Technology of Thailand, 18–19 October 2007, Walailak University, Nakhon Si Thammarat, Thailand (Oral presentation).

5. **Leelatanawit, R.**, Klinbunga, S., Aoki, T., Hirono, I. and Menasveata, P. (2008). Identification, characterization and *in vitro* expression of genes functionally related to testicular development of the giant tiger shrimp *Penaeus monodon* WFC2008: 5th World Fisheries Congress. 20-24 October 2008, Yokohama, Japan (Oral presentation).

ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย