การเปรียบเทียบเมทิเลชันของไลน์-1 ในคนที่สูบบุหรี่กับไม่สูบบุหรี่

นางสาวศิริพร วังศรี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาศัลยศาสตร์ช่องปากและแม็กซิลโลเฟเชียล ภาควิชาศัลยศาสตร์
คณะทันตแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2554

COMPARISON OF LINE-1 METHYLATION BETWEEN SMOKERS AND NON-SMOKERS

Miss Siriporn Wangsri

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of Master of Science Program in Oral and Maxillofacial Surgery

Department of Oral and Maxillofacial Surgery

Faculty of Dentistry

Chulalongkorn University

Academic Year 2011

Thesis Title	COMPARISON OF LINE-1 METHYLATION BETWEEN SMOKERS AND NON-SMOKERS
By	Miss Siriporn Wangsri
Field of Study	Oral and maxillofacial surgery
Thesis Advisor	Keskanya Subbalekha, D.D.S., Ph.D.
Thesis Co-advisor	Professor Apiwat Mutirangura, M.D., Ph.D.

Accepted by the Faculty of Dentistry, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

…….. Dean of the Faculty of Dentistry
(Associate Professor Wacharaporn Tasachan, D.D.S.)

THESIS COMMITTEE

…….. Chairman
(Narong Lumbikananda, D.D.S., Ph.D.)

……..…… Thesis Advisor
(Keskanya Subbalekha, D.D.S., Ph.D.)

…….. Thesis Co-advisor
(Professor Apiwat Mutirangura, M.D., Ph.D.)

…….. Examiner
(Assistant Professor Atiphan Pimkhaokham, D.D.S., Ph.D.)

…….. External Examiner
(Kriangsak Ruchusatsawat, M.S., Ph.D.)

ศิริพร วังศรี : การเปรียบเทียบเมทิเลชันของไลน์-1 ในคนที่สูบบุหรี่กับไม่สูบบุหรี่.
. (COMPARISON OF LINE-1 METHYLATION BETWEEN SMOKERS AND NON-SMOKERS) อ.ที่ปรึกษาวิทยานิพนธ์หลัก : อ.ทญ.ดร.เกศกัญญา สัพพะเลข, อ.ที่ปรึกษาวิทยานิพนธ์ร่วม : ศ.นพ.ดร.อภิวัฒน์ มุทิรางกูร, 138 หน้า.

วัตถุประสงค์ การศึกษานี้มีวัตถุประสงค์เพื่อสืบค้นความเป็นไปได้ที่การสูบบุหรี่อาจส่งเสริมให้เกิดมะเร็งโดยการลดลงของเมทิเลชั่นในไลน์-1 (hypomethylation)

วัสดุและวิธีการ เปรียบเทียบเมทิเลชั่นของไลน์-1 ในเยื่อบุช่องปาก ที่มีลักษณะทางคลินิกปกติของ คนปกติที่ไม่เป็นมะเร็งระหว่างคนที่ สูบบุหรี่กับคนที่ไม่สูบบุหรี่โดยใช้วิธีการ Combined Bisulfite Restriction Analysis of LINE-1 เมทิเลชั่นของไลน์ -1ได้ถูกแบ่งออกเป็น 4 รูปแบบ ตามสถานภาพ และตำแหน่งของเมทิลเลชั่นของคู่เบสไซโตซีนและกวานีนจากทิศทาง 5’ ไปยัง 3’ ได้แก่ $^{m}C^{m}C$, $^{u}C^{u}C$, $^{m}C^{u}C$ และ $^{u}C^{m}C$ โดย ^{m}C แสดงถึงเบสไซโตซีนที่มีเมทิลเลชั่น และ ^{u}C แสดงถึงเบสไซโตซีนที่ไม่มีเมทิลเลชั่น

ผลการศึกษา พบว่าระดับเมทิเลชั่นของไลน์-1โดยรวมไม่มีความแตกต่างระหว่างคนที่สูบบุหรี่กับคนที่ไม่สูบบุหรี่ แต่มีความแตกต่างของ รูปแบบเมทิเลชั่นของไลน์ -1 โดย ในคนที่สูบบุหรี่มีค่าร้อยละของ $^{m}C^{m}C$ และร้อยละของ $^{u}C^{u}C$ เพิ่มขึ้น ในขณะที่ ร้อยละของ $^{m}C^{u}C$ ลดลงอย่างมีนัยสำคัญ (p=0.002 p=0.015 และ p<0.0001 ตามลำดับ) ร้อยละของ $^{m}C^{u}C$ ที่ต่ำลงยังคงพบอยู่ ในผู้ที่หยุดสูบบุหรี่ แม้ว่าจะหยุดนานกว่า 1 ปีแล้วก็ตาม (p=0.001) นอกจากนี้ร้อยละของ $^{m}C^{u}C$ ยังคงลดลงตามปริมาณ การสูบบุหรี่ที่เพิ่มมากขึ้น (p=0.028) และพบว่า $^{u}C^{u}C$ ที่เพิ่มขึ้นสามารถเปลี่ยนแปลงมาจาก $^{m}C^{u}C$ และ $^{u}C^{m}C$ ในขณะที่ $^{m}C^{m}C$ ที่เพิ่มขึ้นมาจาก $^{m}C^{u}C$ เท่านั้น จากการวิเคราะห์ข้อมูลการแสดงออกของยีนใน microarrays ของเยื่อบุทางเดินหายใจของผู้ที่สูบบุหรี่ พบว่าการสูบบุหรี่มีผลต่อการกระตุ้นหรือยับยั้งการแสดงออกของยีนที่มีไลน์-1 เมื่อเปรียบเทียบกับยีนที่ไม่มีไลน์-1

สรุป ดังนั้นการสูบบุหรี่อาจทำให้เกิดการเปลี่ยนแปลงเมทิเลชั่นของไลน์-1ทั้งในรูปแบบการลดและการเพิ่ม ระดับเมทิเลชั่นของไลน์-1ที่ลดลงจากการสูบบุหรี่ทำให้เกิดผลเช่นเดียวกับระดับเมทิเลชั่นที่ลดลงในมะเร็ง

ภาควิชาศัลยศาสตร์..... ลายมือชื่อนิสิต .....
สาขาวิชาศัลยศาสตร์ช่องปากและแม็กซิลโลเฟเชียล ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์หลัก.....
ปีการศึกษา 2554..... ลายมือชื่อ อ.ที่ปรึกษาวิทยานิพนธ์ร่วม.....

5276134632 : MAJOR ORAL AND MAXILLOFACIAL SURGERY
KEYWORDS : LINE-1 METHYLATION PATTERN / SMOKING / ORAL RINSE / ORAL MUCOSA / SMOKE

SIRIPORN WANGSRI : COMPARISON OF LINE-1 METHYLATION BETWEEN SMOKERS AND NON-SMOKERS. ADVISOR : KESKANYA SUBBALEKHA, D.D.S.,Ph.D., CO-ADVISORS : PROF. APIWAT MUTIRANGURA, M.D.,Ph.D., 138 pp.

Objective This study aimed to investigate the possibility that smoking may promote cancer development via LINE-1 hypomethylation.

Material and methods The LINE-1 methylation in clinically normal oral mucosa of current smokers was compared to that of non-smokers by using combined bisulphite restriction analysis. Each LINE-1 sequence was categorised into 4 patterns depending on the methylation status and location of 2 CpG dinucleotides from 5' to 3'; which included $^{m}C^{m}C$, $^{u}C^{u}C$, $^{m}C^{u}C$ and $^{u}C^{m}C$. Of these, ^{m}C and ^{u}C represent methylated and unmethylated CpG, respectively.

Results Despite there was no significant difference in the overall LINE-1 methylation level, the percentages of some methylation patterns were different. The $\%^{m}C^{m}C$ and $\%^{u}C^{u}C$ increased, while the $\%^{m}C^{u}C$ decreased in current smokers (p=0.002, 0.015 and <0.0001, respectively). Additionally, the lower $\%^{m}C^{u}C$ still persisted in persons who had stopped smoking for over 1 year (p=0.001). The $\%^{m}C^{u}C$ also decreased in the higher pack-year smokers (p=0.028). Interestingly, the $^{u}C^{u}C$ could rise from $^{m}C^{u}C$ to $^{u}C^{m}C$, while $^{m}C^{m}C$ could rise from $^{m}C^{u}C$ only. We further analysed expression microarrays from the airway epithelia of smokers and found that smoking-associated intragenic LINE-1 sporadically repressed or activated host genes, compared to genes that do not contain LINE-1.

Conclusion The smoking paradoxically increase or decrease LINE1 methylation of certain loci. Hypomethylated LINE-1 loci induced by smoking led to the same consequences as those associated with cancer.

Department : Oral and maxillofacial surgery Student's Signature .....
Field of Study : Oral and maxillofacial surgery Advisor's Signature .....
Academic Year : 2011 ..... Co-advisor's Signature .....

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude and sincere appreciation to my advisor and co-advisor, Dr. Keskanya Subbalekha and Professor Apiwat Mutirangura for their suggestion and kindness support throughout the course of this program. They deserve my deepest appreciation for their energy and endless idea in supervision and encouragement. In addition, I would like to thank Miss Paipan Phitayanon for her kindness and advice in statistic analysis.

I would like to thank my thesis committee members; Dr. Narong Lumbikananda, Dr. Kriangsak Ruchusatsawat and Assistant Professor Dr. Atiphan Pimkhaokham for their suggestion and kindness in being committee members.

I express my sincere thanks to all of my colleagues and all staffs in the Department of Oral and Maxillofacialsurgery for their fellowship and support during this work. This success will always be theirs as well. Moreover, this study would not be accomplished without the well co-operative of all attended participants.

Finally, I would like to express my appreciation to my mother for her endless love, caring and support. All of these are expressed from the bottom of my heart.

This study was financially supported by The 90th Anniversary of Chulalongkorn University Fund (Ratchadaphiseksomphot Endowment Fund).

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Page

Page

LIST OF ABBREVIATIONS

COBRA	combined bisulfite restriction analysis
LINE-1	long intersperse nuclear element-1
OSCC	oral squamous cell carcinoma
HNSCC	head and neck squamous cell carcinoma
DNA	deoxyribonucleic acid
CpG	dinucleotide containing cytosine and guanine, respectively
DNMT	DNA methyltransferase
NS	non-smoker
FS	former smoker
CS	current smoker
CU-DREAM-X	connection up- or down-regulation expression analysis of microarrays extension program
ANOVA	analysis of variance
CI	confidence interval
OR	odds ratio

CHAPTER I

INTRODUCTION

Background and Rationale

Tobacco smoking is a predisposing factor of many malignancies *[1-4].* The risk of upper aerodigestive cancers increases with the higher pack-years cigarette smoking *[3, 5, 6]*. However, this risk decreases after discontinuation of smoking and reverts to the non-smoker risk level if smoking is ceased for more than 15 years *[3, 6]*. Additionally, smoking increases the number of keratinised cells in the epithelium of the tongue and hard palate *[7]*. This effect varied in different regions, depending on the extent of direct exposure to smoke *[8].* Interestingly, oral mucosal lesions resolved after cessation of smoking for a period of time *[9, 10].* The mechanism of smoking causing these change is still not well declared.

Mutation, promoter methylation and global hypomethylation are three crucial DNA modification events that lead to cancer development *[11-13]*. Smoking promotes mutation and alteration of gene promoter methylation *[11, 14, 15].* Moreover, the evidence suggesting the association between the degree of global hypomethylation and smoking history of HNSCC patients was shown *[16]*.

Long interspersed nuclear element-1s (LINE-1s) are repetitive transposable elements which are widely distributed in the genome *[17].* There are 500,000 copies of LINE-1 in the human genome *[18].* More than 10,000 LINE-1s contain a 5'UTR *[19].* The reduction of methylation levels of LINE-1, can reflect global hypomethylation *[20].* In most cancers, LINE-1 methylation levels diminish early and progressively which correlate significantly with tumour phenotype, including tumour progression and prognosis *[12, 16, 21-25].* Hypomethylation of LINE-1 significantly increases the risk for head and neck squamous cell carcinoma (HNSCC) *[26]*. Paradoxically, events associating LINE-1

hypermethylation with carcinogenesis have also been found in malignant peripheral nerve sheath tumour, myelodysplastic syndrome and partial hydatidiform moles *[27-29].* In blood samples of HNSCC patients, LINE-1 methylation levels slightly increased with higher pack-years of smoking *[26].*

The effects of smoking on LINE-1 methylation levels in non-cancerous cells have also been reported. No changes were observed in blood cells or in the colonic epithelium of smokers *in vivo [30-32]*. However, an *in vitro* study revealed minimal reduction of LINE-1 methylation levels in the respiratory epithelium under high dosage cigarette smoke condensate treatment *[33].* Oral mucosa is directly exposed to tobacco smoke and its chemical agents. Therefore, it is interesting to clarify whether this epigenetic change occurs before malignant transformation.

Currently, most LINE-1 methylation studies have measured the genome-wide methylation levels of LINE-1s. However, methylation of LINE-1s can be influenced by multiple mechanisms. The measurement of the methylation level alone may not be able to detect LINE-1 methylation changes in certain events, even if such changes can promote cancer development. In normal cells, some functions of LINE-1 methylation are to maintain genomic integrity and regulate gene expression in *cis [20, 34-36].* Consequently, genomic instability and repression of gene expression can be observed on chromosomes in which LINE-1s are hypomethylated. Therefore, in theory, certain conditions that stochastically alter LINE-1 methylation levels will promote carcinogenesis on chromosomes with LINE-1 hypomethylation, but these hypomethylated LINE-1s will be encrypted by other hypermethylated LINE-1 loci. Locus-specific mechanisms causing variations in methylation levels among LINE-1s in different loci has also been reported *[37].*

Recently, a wide range of approaches to obtain quantitative information of genomic DNA methylation have been developed *[38].* Most standard techniques measure several CpGs in each LINE-1. Pyrosequencing often measures 4 CpG dinucleotides *[39]*, whereas combined bisulfite restriction analysis (COBRA) polymerase

chain reaction often measures 2 CpGs *[12].* Compared to previously reported LINE-1 sequences *[37],* the methylation state of 2 of the CpG dinucleotides detected by COBRALINE-1 correlated directly with other CpG dinucleotides on 5'LINE-1s. LINE-1alleles can be classified into four groups depending on the methylation status of 2 CpG dinucleotides on each strand from 5' to 3' detected by COBRALINE-1. The first class contains 2 unmethylated CpGs ($^{u}C^{u}C$) and the second class contains 2 methylated CpGs ($^{m}C^{m}C$), representing hypomethylated and hypermethylated LINE-1 loci, respectively. The third and fourth classes are partially methylated LINE-1s including 5'methylated with 3'unmethylated CpGs ($^{m}C^{u}C$) and 5'unmethylated with 3'methylated CpGs ($^{u}C^{m}C$) (Figure 26A). Recently, our group found that $\%^{u}C^{u}C$ is more effective in determining cancer risk than overall methylation levels *[40, 41].*

However, studies that evaluate the association between smoking and repetitive sequence methylation changes *in vivo* have not yet been conclusive. Herein, we evaluated the possibility that smoking may promote cancer development via genomic hypomethylation by evaluating the LINE-1 methylation pattern found in the oral mucosa of smokers.

Research Question

1. Do the LINE-1 methylation patterns in oral epithelia differ between non-smokers and current smokers?

2. Do the LINE-1 methylation patterns in oral epithelia differ between the high and low pack-year groups in current smokers?

Objective

1. To investigate the LINE-1 methylation patterns in oral epithelia of non-smokers and current smokers.

2. To investigate the LINE-1 methylation patterns in oral epithelia between the high and low pack-year groups.

Hypothesis

- **Hypothesis I**

Ho: LINE-1 methylation patterns in oral epithelia of non-smokers are not significantly different from current smokers.

Ha: LINE-1 methylation patterns in oral epithelia of non-smokers are significantly different from current smokers.

- **Hypothesis II**

Ho: LINE-1 methylation patterns in oral epithelia of current smokers are not significantly different between the high and low pack-year groups.

Ha: LINE-1 methylation patterns in oral epithelia of current smokers are significantly different between the high and low pack-year groups.

Research Design

Analytical cross-sectional research

Expected Benefit

Investigating the alteration of LINE-1 methylation patterns in oral epithelial cells of smokers may benefit the prevention of smoking-associated oral cancer.

Research Methodology Framework

Investigate LINE-1 methylation patterns of oral epithelia collected from oral rinse of non-smokers and current smokers by using COBRA LINE-1 technique

Investigate LINE-1 methylation patterns of oral epithelia between the high and low pack-year groups

CHAPTER II

REVIEWS AND RELATED LITERATURES

Tobacco

The harmful effect of tobacco results in pathology of many organs, including oral cavity (Table 1), cardiovascular system, respiratory system and gastrointestinal system *[2, 42, 43]*. The overall risk of oral cancer among smokers is 7–10 times higher than non-smokers. In addition, the strong dose-response relationship between smoking rates and risk of these cancers are reported (Table 2) *[44]*. Furthermore, the risk of primary, recurrent and secondary oral cancer is related to continuing smoking after treatment *[45]*. This finding indicates that smoking induces permanent change of the biological process of oral epithelium.

Oral leukoplakia was found more frequently in smokers than non-smokers *[46]*. Moreover, oral premalignant lesions such as leukoplakia and erythroplakia found in smokers have an annual cancer transformation rate of about 5% *[42]*. The dose response relationship remains significant between tobacco smoking-oral leukoplakia and tobacco smoking-oral epithelial dysplasia *[44]*. Proliferation of oral epithelium increases in current smokers and former smokers, both HNSCC and healthy person *[47]*.

The relationship between conventional smoking and the anatomical site of oral cancer is less clear. However, carcinogens in tobacco smoke can dissolve in saliva and collect in the gutter areas where saliva is pooled. These situations increase the risk of oral cancer developing in the floor of the mouth and ventral or lateral tongue and the soft palate *[48]*. However, the risk for OSCC and squamous cell carcinoma of the upper aerodigestive tract decreases in smoking cessation patients *[3, 5, 6]*. Moreover, oral precancerous lesion may be also regress or turn to normal epithelium *[44, 46, 49]*.

Tobacco can be consumed through the mouth in various forms, including smokeless tobacco chewing on itself or combined with areca nut and tobacco smoking *[42]*. The manufactured cigarettes is the most prevalent form of tobacco smokers *[44]*. Tobacco use, including smoking, reverse smoking and smokeless tobacco increases the risk of cancers of the upper aerodigestive tract *[42]* and oral cavity which a significant number develop at the site of tobacco placement in smokeless tobacco users *[50]*. However, the risk of oral precancerous lesion or oral cancer varies with smokeless tobacco habit. Smokeless tobacco use may provoke a very slow process for cancer developing *[51]*. People consuming high smokeless tobacco did not reveal high rate either verrucous or invasive squamous cell carcinoma. However, the specific association between tobacco smoking and OSCC was found *[52]*.

Alcohol drinking, a habit commonly goes with smoking, has been established as a common risk factor for oral carcinogenesis *[53-55]*. The odds ratio (OR) of oral cancer for consumption of 6-20 cigarettes/day and for more than 20 cigarettes/day are 3.1 and 7.96, respectively. When more than 50 g of alcohol/day is consumed, the risk results in an OR of 5.3 *[54]*. Although drinking and smoking are independent risk factors, they have a synergistic effect and greatly increase risk together *[48, 53]*. Heavy smokers (40 cigarettes/ day) and heavy drinkers (30 drinks per week) have 38 times the risk of developing oral cancer than abstainers from both products *[44, 51]*. Even some investigators had attempted to differentiate the combination effects of these two agents, the nature of the biological interaction between them has not been definitely established *[54]*. However, Welbourne JP suggested that only cigarette smoking represented the true-causing agent *[52]*. Additionally, epidemiologic studies revealed that up to 80% of oral cancer cases were smokers *[45, 56, 57]*. Therefore, tobacco smoking is recognized as a major risk factor of oral cancer.

Table 1 Oral lesions and conditions associated with tobacco use *[42]*.

Oral precancerous lesions:

Leukoplakia, erythroplakia, smokeless tobacco keratosis

Oral cancers:

Squamous cell carcinomas of the tongue, floor of the mouth, lip and gingiva

Verrucous carcinomas of the buccal mucosa, gingival and alveolar ridge

Periodontal diseases:

Increased plaque and calculus depositions, ischemia, gingival inflammation, periodontal pockets, gingival recession and alveolar bone loss

Root caries

Peri-implantitis

Halitosis

Taste derangement

Stained teeth and restorations

Table 2 The relationship of smoking and oral/oropharyngeal cancer.

Author	Year	Country	Habit	RR/OR	(CI)
Brugere et al.[9] (mouth)	1986	France	Tobacco 10–19 g	3.9	2.5–6.3
			Tobacco 20–29 g	8.6	5.6–13.3
			Tobacco >30 g	15.4	9.7–24.4
Blot et al.[8] (oropharyngeal)	1988	USA	Smoking 1–19/20 + year	1.6	(0.9–2.7)
			Smoking 20–39/20 + year	2.8	(1.8–4.3)
			Smoking 40+/20 + year	4.4	(2.7–7.2)
Talamini et al.[10*] (oral and pharyngeal)	1990	Italy	Cigarette <15/day	3.8	0.2–58.2
			>15/day	12.9	2.3–106.3
Franceschi et al.[11] (oral cavity)	1990	Italy	Cigarette smoker	11.1	3.4–34.8
			Cigar and pipe	20.7	5.6–76.3
Boffetta et al.[12] (oral tongue)	1992	USA	Cigarettes 16–25/day	1.8	0.8–4.2
			26–35/day	1.9	0.7–5.3
			>35/day	2.1	0.9–5.1
			Cigar and pipe	0.2	0.0–2.0
Negri et al.[13] (oral and pharyngeal)	1993	Italy	Moderate smokers	3.6	–
			Heavy smokers	9.4	–
Merletti et al.[14] (oral and oropharyngeal)	1989	Italy (M)	Tobacco 8–15 g/day	4.4	1.0–18.3
			16–25	5.1	1.2–21.0
			>25	6.2	1.4–28.3
		(F)	Tobacco 8–25+	0.6	0.1–2.4
De Stefani et al.[15] (oral)	1998	Uruguay	Non-smoker	1.0	–
			Past smoker	2.2	1.2–3.9
			Current smoker	5.7	3.4–9.5
			All smokers	4.2	2.6–6.8
Schlect et al.[16] (UADT)	1999	Brazil	Smoking	8.3	5.3–13.0
Moreno-Lopez et al.[17] (oral)	2000	Spain	Cigarettes/day 0	1	–
			1–20	3.15	1.53–6.48
			>20	12.5	5.68–27.84
Zavras et al.[18] (oral)	2001	Greece (M)	Current smoker	3.0	1.2–79
		(F)	Current smoker	0.7	0.7–3.7

* Not adjusted because one habit only; UADT-upper aerodigestive tract cancers.

Warnakulasuriya S et al. collected the relative risks (RR) and odds ratio (OR) for smoking in oral and oropharyngeal cancer from many studies. The evidence suggests that most studies revealed OR/RR > 1 and the risk of oral and oropharyngeal cancer are dose-response relationship *[44]*.

- **The Mechanisms of Tobacco Smoking Carcinogenesis**

Tobacco contains many carcinogenic products. The main agents present in tobacco, regardless of how it is used, are nitrosamines derived from nicotine *[48]*. Smokeless tobacco is directly contact to buccal mucosa. They act locally on keratinocyte stem cells, then absorbed and act in many other tissues in the body *[58]*. In addition, chewing of tobacco results in a local exposure of the oral mucosa to tobacco-specific nitrosamines (TSNA) which are usually present high levels of carcinogenic. TSNAs such as N-Nitrosonornicotine (NNN) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) were found in saliva *[44]* and suggested to be the most important factor for smokeless tobacco contribute to oral disease *[59]*.

Tobacco smoke also contains many carcinogenic products which TSNAs derivatives are generated primarily during pyrolysis *[58]*. However, the evidences suggested that polynuclear aromatic hydrocarbons (PAH) predominate as carcinogenic combustion products *[44, 60]*. Principally PAH, including benzopyrene and benzanthracene do not cause carcinogenesis themselves. However, the burning tobacco transforms them into carcinogens which are primarily contact to tissues. These epoxides of tobacco tars are the actual DNA damaging *[60]* (Figure 1).

In addition, there are evidences indicate that smoking are associated with some of the genetic and epigenetic changes. In genetic changes, p53 mutations are frequent in tobacco-related cancers and the mutation load is often higher in cancers from smokers than from non-smokers *[11]*. Epigenetic mechanisms, such as change in gene expression via genomic instability and hypermethylation of tumour suppressor gene promoter, relate to smoking consumption *[61]*.

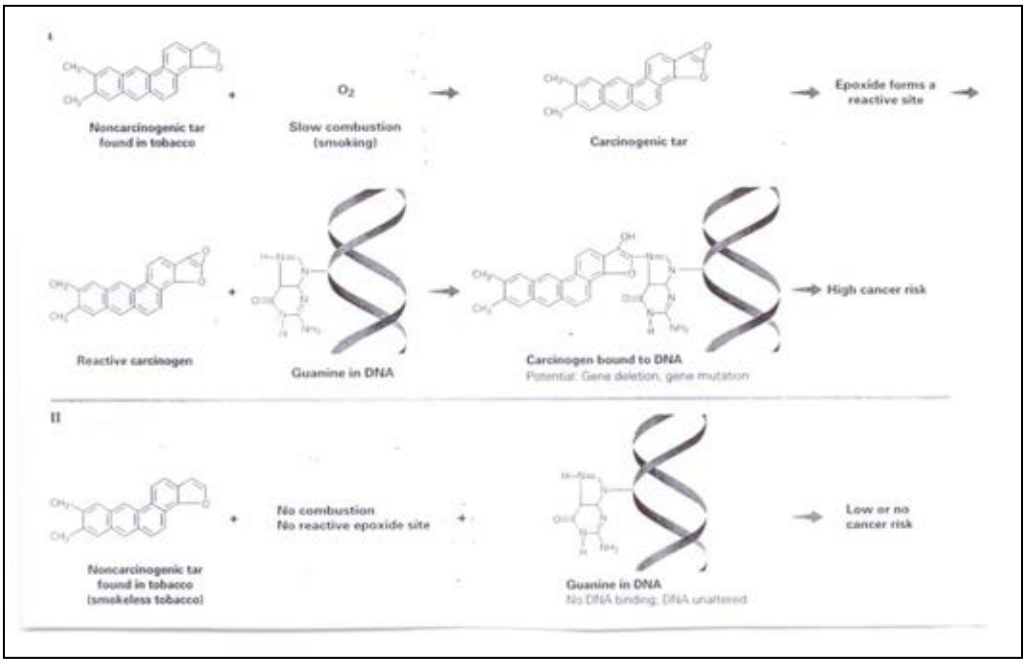


Figure 1 Possibility of carcinogenesis mechanism of smoking. This figure is from Marx RE and Stern D, 20003.

The carcinogens in tobacco are actually carcinogenic when they are partially oxidise to an epoxide by smoking burning process. Conversion of tobacco into their free radicals by the combustion of smoking allow them to bind to DNA, then DNA damage will be processed *[60].*

Genetics and Epigenetics in Cancers

Malignancy is a multistep process, which arises through an accumulation of genetic and epigenetic alterations that disrupt the normal function of human genome *[62]* (Figure 2). The notion that cancer is a genetic disease has shaped the cancer research field for decades from Knudson hypothesis. Following this logic, efforts to improve the detection and treatment of cancer have focused at the genetic and expression aspects of cancer cells in the human genome project of cancer. However, there have been debates as to whether the goals of this program is scientifically sound and practically feasible, given the mounting evidence that perturbation of epigenetic regulation *[63]*. This is mainly due to the accumulation of evidence indicating that

epigenetic deregulation of cells contributes and cooperates with genetic alterations in all stages of cancer development and progression *[64].* In addition, data from sporadic cancers which comprise 90–95% of all cancers, almost uniformly exhibit both genetic and epigenetic defects genome-wide *[65]*.

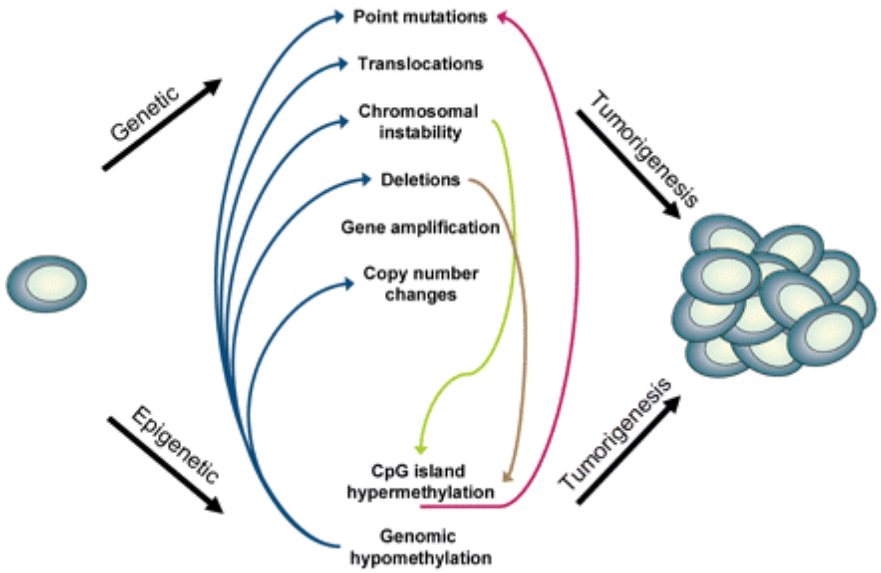


Figure 2 Illustration of carcinogenesis of human cancer.

Hereditary and sporadic human cancers suggest genetic and epigenetic processes are complementary for cancer development and these mechanisms show substantial interaction *[65].*

Genetic aberrations change expression by altering the sequence of adenine (A), thyamine (T), cytosine (C) and guanine (G) bases *[66].* These kinds of aberration; known as mutation, deletion, insertion and rearrangement may change gene and chromosome structure *[67].* Epigenetic mechanism controls gene expression without altering the DNA sequence. This mechanism is a reversible and heritable modification. Three forms of epigenetic alteration appearing in host cells are known as: small-interfering RNAs, histone modification and DNA methylation (Figure 3). These modifications effect

genomic stability, DNA conformation, position of nucleosome, chromatin structure and sequenttiallly nuclear organization. Consequently, these processes determine whether a gene is activated or silenced *[38]*.

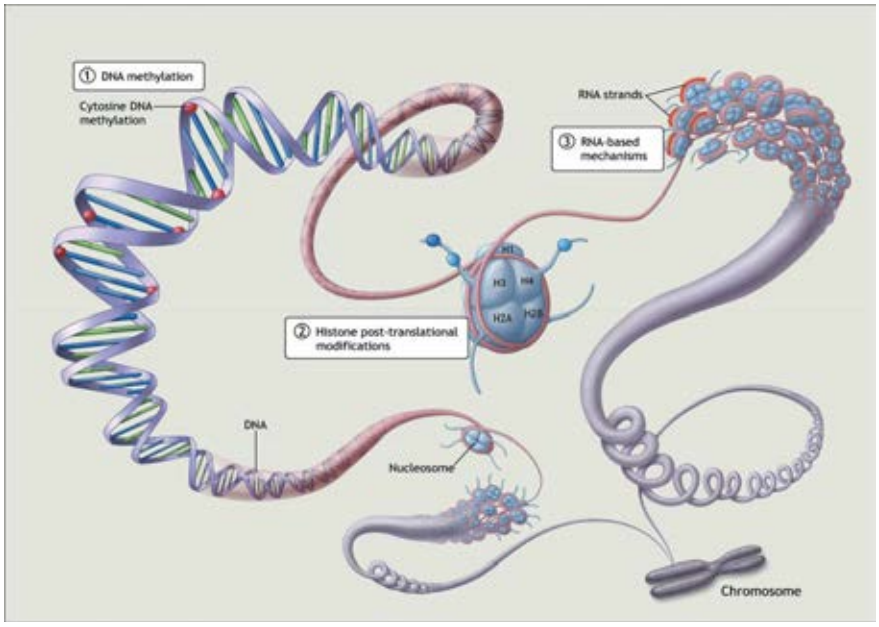


Figure 3 Different types of epigenetic regulation *[68]*.

1. Histone modifications refer to covalent post-translational modifications of N-terminal tails of four core histones (H3, H4, H2A and H2B).
2. RNAs interference, which either in the form microRNA can alter gene expression states in a heritable manner.
3. DNA methylation is a covalent modification of the cytosine that is located 5' to a guanine in a CpG dinucleotide.

Epigenetics in Cancers

Epigenetic changes have associated with cancer-specific expression differences in human malignancies, including HNSCC. These alterations are known to occur the earliest and most common events in tumourigenesis *[38, 69]*. DNA methylation is one of the most commonly existing epigenetic evidence engaging in the human genome *[70]* and also contributes to the carcinogenesis and cancer progression *[63]*.

- **Environmental Epigenetics**

Many scientists are beginning to appreciate the influencing of environment exposures to progression into malignancy from the alteration of genetic and epigenetic changes. The two complementary mechanisms are related at every step of tumourigenesis *[71]* (Figure 4). Therefore, a complex interaction between genetic and epigenetic modifications induced by environmental factors may carry to cancer development.

Environmental stimuli such as endocrine dietary and chemical substance, affects the epigenetic event. Although being inheritable, epigenetic modifications are reversible and can be led to significant cellular dysfunction that impinge on human genomes *[72, 73]*. These epigenetic aberrations can be distinguished from developmental epigenetics and is called environmental epigenetics *[73]*.

A mechanistic model suggested HNSCC arised by the contribution of genetic and epigenetic alterations in oncogenes, tumour suppressor genes and/or DNA stability genes. Global DNA methylation may precede genetic alterations and molecular changes associated with exposure to environment carcinogens such as smoking and drinking in HNSCC *[74]* (Figure 5). Recently, chemical agents such as benzene causing global hypomethylation was reported both *in vitro [75]* and *in vivo [76]*

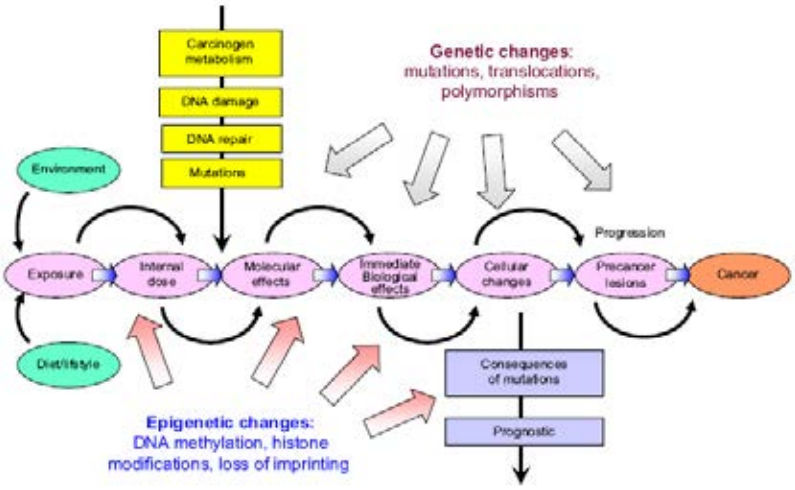


Figure 4 The model suggested cancer progression by the combined of genetic and epigenetic.

Environmental/Nutritional factors induce unsuitable activation or inactivation of specific genes leading to tumour development *[70]*.

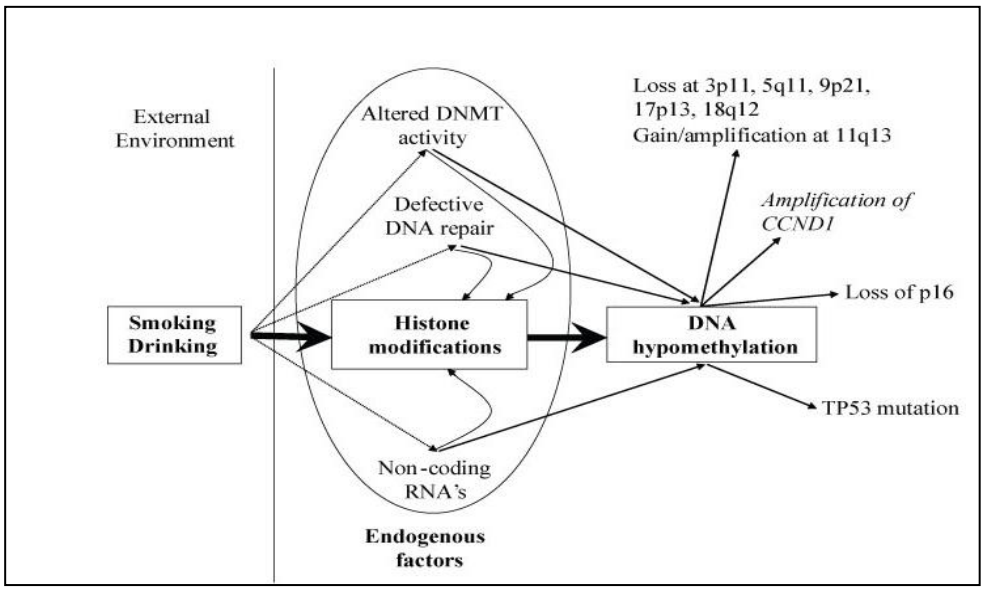


Figure 5 Proposed model of genetic and epigenetic alterations are associated with environmental factors.

Smoking and drinking are associated at the early carcinogenic stage with allelic loss at 3p11, 5q11, 9p21, 17p13, 18p12 gain at 11q13 and amplification of cyclin D1 (CCND1) gene, loss of p16 and TP53 mutation *[74]*.

DNA Methylation

The important epigenetic modification found in mammalian is DNA methylation *[77, 78]* in which 5mC (5-methyl cytosine) is created in situ by DNA methyltransferase (DNMT) enzymes. A methyl group is transfered from the universal methyl donor *S*–adenosyl methionine (SAM) to position 5' of the C ring *[38, 77, 79, 80]* (Figure 6). There are four methylation processes that can occur within the nucleus: the first is de novo methylation, where previously unmethylated Cs, usually in the symmetrical sequence

context CpG, become methylated. The second is maintenance methylation, where the strand symmetry of hemimethylated DNA is maintained after replication by the methylation of the newly synthesised strand. The third is passive demethylation, where the maintenance methylation activity is suppressed, resulting in a 50% decrease in methylation during each round of DNA replication. The fourth is active demethylation, where methylation levels are decreased, in the absence of DNA replication, via an enzymatic process *[77]*.

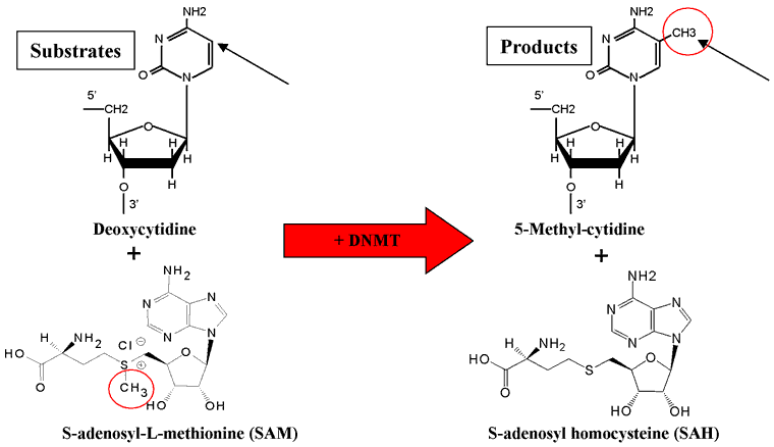


Figure 6 DNA Methylation Processes.

DNA methylation is a covalent modification formed by addition of a methyl group at the 5'carbon of cytosine in the sequence context 5'-CG-3' of the DNA molecule. The reaction is catalyzed by enzymes DNMT which catalyze the transfer a methyl group from the methyl donor *S*-adenosyl methionine (SAM) *[79, 81]*.

Methylation occurs predominantly at C residues in CpG dinucleotides (CpGs) *[38, 82]*, although methylation also takes place in non-CpGs such as CpNpG and nonsymmetrical CpA and CpT at a lower frequency *[38]*. CpG dinucleotides are nonrandom *[42]* and normally under-represented in the genome but are found concentrated at the expected levels in C+G rich regions termed CpG islands (CGIs) in which the frequency of the CG sequence is higher than other regions, where "p" simply indicates that "C" and "G" are connected by a phosphodiester bond, that frequently coincide with promoter or gene regulatory regions which are essential for general cell

functions or other genes frequently expressed in a cell *[79].* However, the majority of CpG dinucleotides are found within the intergenic and intronic regions of DNA particularly within repeat sequences and transposable elements *[82]*.

In normal somatic cells of human tissues, between 70 and 90% of CpG dinucleotides are methylated which constitutes approximately 0.75–1% of the total number of bases in the genome *[82]*. In the healthy genome, most CpG islands are not susceptible to methylation and are often associated with the promoters of both house keeping genes and genes with tissue specific patterns of expression *[79].* Because of the high susceptibility of 5mC to undergo spontaneous deamination to yield T (Figure 7), the mammalian genome has become progressively depleted of CpGs through the course of evolution to protect spontaneous deamination *[38]*. While the CpGs found dispersed throughout the rest of the DNA are mostly methylated *[82],* exceptions to the unmethylated status of CGIs include those that are associated with imprinted genes, genes subject to X-chromosome inactivation, transposable elements *[83]* (Figure 8). However, in specific instances gene promoter regions are methylated as part of normal developmental processes. Conversely, abnormal gene-specific demethylation and global hypomethylation (involving repeat sequences throughout the genome) potentially can lead to overexpression of genes and activation of transposable elements contributing to disease *[82].*

Methylation at CpG dinucletode is catalyzed by three major DNMTs namely DNMT1, DNMT3a and DNMT3b (Table 3). DNMT1 is responsible for maintaining genomic DNA methylation patterns and employs hemi-methylated-CpG dinucleotides, produced after DNA replication or repair, as substrate and fully methylates them. DNMT3a and 3b methylate previously unmodified CpG residues and hence are known as de novo methylases *[82, 84]*. Active *de novo* methylation is known to occur in germ cells, early embryonic stages and helps to maintain tissue specific gene expression patterns and they are also reported to aid in maintenance methylation *[79]* (Figure 9).

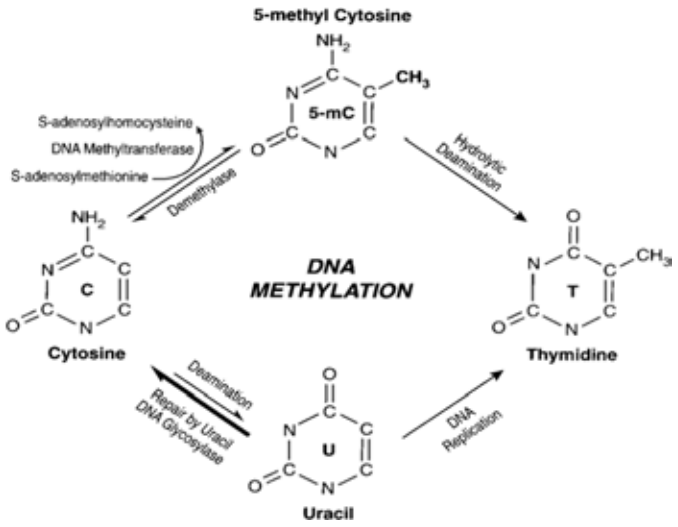


Figure 7 The mechanism of DNA methylation and transcriptional silencing.

The machinary for cytosine methylation, demethylation and mutagenesis of cytosine and 5-methyl cytosine (5mC) DNA methylation by DNMTs and demethylation is catalyzed by demethylase. 5mC undergoes hydrolytic deamination to thyamine. Mutation at CpG occurs because 5mC is more susceptible than cytosine to deamination and because some of the thyamine-guanine mismatches produced by deamination are poorly repaired *[38]*.

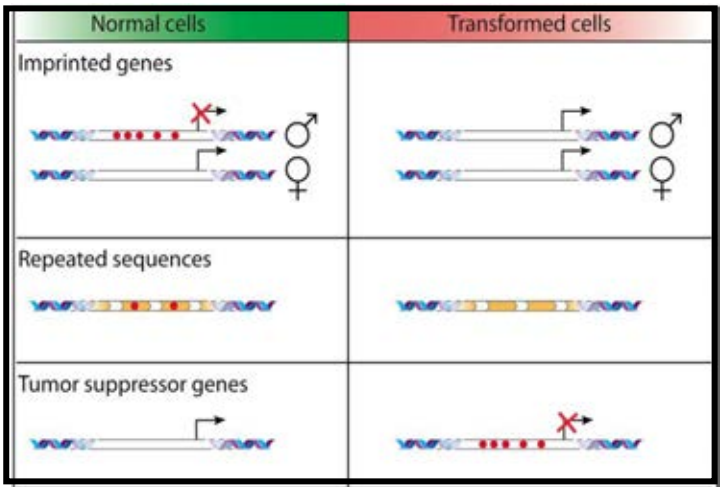


Figure 8 The loci of DNA methylation in mammals.

In normal and cancer cells, the target sites of DNA methylation are difference. In normal cells, DNA methylation inactivation three main types of targets. Cancer

cells often have an aberrant pattern of DNA methylation where some tumour suppressor genes are methylated and repressed. Contrarily, some normally methylated sequences, such as repetitive DNA and imprinted genes can reform unmethylated *[85]*.

Table 3 A family of enzyme DNMTs *[86]*.

Methyl-transferase	*Methylase function*	*Main expression pattern*	*Disruption in mice*	*Mutations in humans*
Dnmt1	Maintenance	Adult/embryo	Die *in utero*	–
Dnmt3a	*De novo*	Embryo	Die at 4 weeks	–
Dnmt3b	*De novo*	Embryo	Die *in utero*	ICF (Box 1)

Abbreviations: ICF, immunodeficiency, centromeric instability and facial anomalies syndrome.

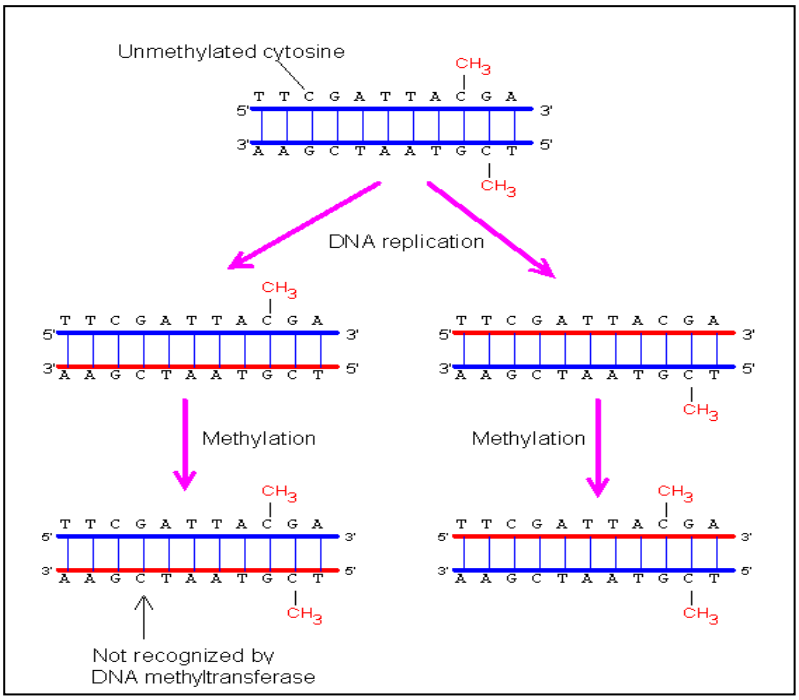


Figure 9 *De novo* methylation pattern.

The DNMT can methylate only the CpG sequence paired with methylated CpG. The CpG sequence not paired with methylated CpG will not be methylated. Hence, the original pattern can be maintained after DNA replication *[87]*.

Roles of DNA Methylation

DNA methylation is the main epigenetic condition that correlated with normal functions in humans such as X chromosome repression *[88-90],* genomic imprinting *[69, 88, 90]*, embryogenesis, gametogenesis *[88, 90]*, environmental exposure, aging *[89]*, silencing of repetitive DNA elements *[90]* and disease processes including mental retardation, autoimmune disease *[89]* and cancer *[69, 88, 89]*.

This heritable epigenetic modification is also associated with transcriptional repression by three mechanisms. First, the methyl group of the 5mC extends into the major groove of DNA and inhibits binding of transcription factors (TFs) to their CpG containing recognition sites. Second, a class of proteins known as methyl-binding proteins (MBDs) specifically bind methylated CGIs and create steric hindrance to access by TFs to their regulatory elements. Both mechanisms will suppress gene transcription. Furthermore, the last upon binding to methylated CGIs, MBDs recruit histone deacetylases (HDACs) and histone methyltransferases (HMTs). These enzymes mediate complex histone modifications and result in the establishment of repressive chromatin structures that permanently silent gene transcription *[83]* (Figure 10).

DNA methylation modifications in cancer were first described by Feinberg et al. 1983. They found the overall pattern of hypomethylation in human cancer; specifically, a 10% reduction in genomic 5mC content in pre-cancerous and cancerous colonic polyps *[66].* These can reveal either in the pattern of hypomethylated CpG or hypermethylated CpG, have been observed in human tumours. Aberration in methylation forms are expressed to either inactivate via hypermethylation or activate via hypomethylation *[25]*. DNA methylation can inactivate gene transcription of the target gene and occur as one of the multi-hits in the Knudson hypothesis by silencing one or both alleles of the tumour suppressor genes in sporadic cancers. It can efficiently occur as a second hit at the time of hereditary cancers transformation *[91]* (Figure 11).

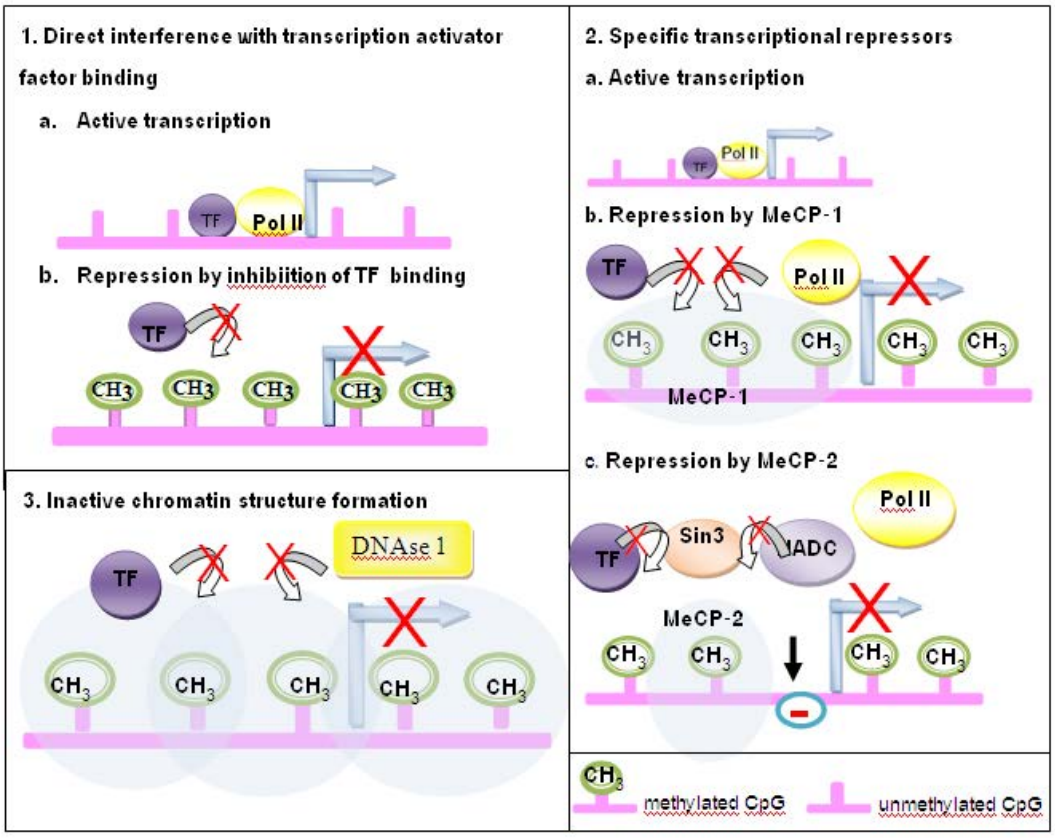


Figure 10 Suggested mechanisms of transcriptional inactivation induced by DNA methylation (modified from Singal R and Ginder GD, 1999) *[83]*.

1. Directly inhibition with the binding of specific TFs to their recognition sites in their respective promoters.

2. Potentially mechanism for methylation imediated silencing is through the direct binding of specific transcriptional repressors to methylated DNA.

3. Methylation induce transcriptional inactivation is by modifying the structure formation of chromatin.

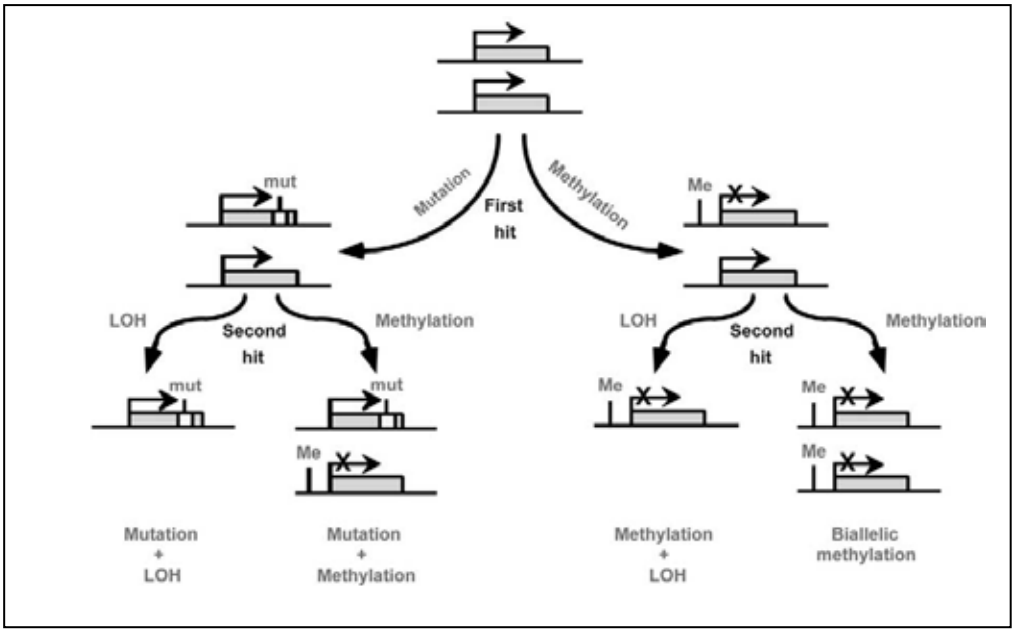


Figure 11 Diagram demonstrated methylation as a second hit.

Epigenetic and genetic mechanisms may act in combination to cause inactivation of gene during the development of hereditary cancers *[91].*

DNA Methylation and Cancer

A role for DNA methylation in carcinogenesis has been proposed for a long time. Many studies have suggested abnormal DNA methylation in DNMTs activity in cancer cells. Abnormal cells often have raised total DNMTs activity, widespread loss of methylation from normally methylated sites and more regional areas of hypermethylated DNA. The potential contribution of DNA methylation to oncogenesis appears to be mediated by one or more of the following mechanisms *[83]* (Figure 12).

1. Signature C → T mutation in cancer cells. The high mutation rate of C residues within the dinucleotide CpG, the target site of mammalian DNMTs, can be accounted for by an increased rate of C to T transitions, which are, in turn a consequence of hydrolytic deamination of 5mC. Unmethylated C can also undergo deamination to yield U, but the well-characterized U-DNA glycosylase efficiently repairs G:U but not G:T mismatch *[83]* (Figure 7, 12).

2. Induction of chromosomal instability. It was proposed that in mismatch repair proficient colon cells a methylation defect directly facilitates the gain and loss of whole chromosomes, leading to the genomic instability necessary for the development and progression of cancer *[83].*

3. Hypermethylation of tumour-suppressor genes. Transcriptional repression by hypermethylation of promoter sequences means for the inactivation of tumour-suppressor genes in cancer. This may result from the increased DNMTs levels that have been demonstrated in various cancers or it could occur as a result of some other transient event that silences tumour-suppressor gene transcription *[83].* This notion is further exemplified by the recent finding that the expression of 20 or more genes involved in cell immortalization and transformation is affected by methylation of CpG island of their promoters. The biological roles of these genes can be categorized as cell cycle regulation (p15 and p16), DNA repair and protection (BRCA1, GSTP1, hMLH1 and MGMT), cell adherence and metastasis (DAPK, E-cadherin and TIMP3) and the APC/ b-catenin route. Indeed, aberrant promoter methylation of these genes has been linked to carcinogenesis in many human cancers, including cancers of the lung, liver, breast, stomach and head-neck *[80].*

4. DNA hypomethylation in cancer. DNA methylation of the entire genome is generally reduced and this condition is known as “global hypomethylation” *[37].* Global methylation reveals to begin early and progressively before cancer formation. From the overall genomic hypomethylation, specific oncogenes have been observed to be hypomethylated in human cancers *[83].* Global hypomethylation occurs not only in transcription control regions, such as promoters, but also in repeated sequences, such as heterochromatic regions and retrotransposons *[26]* such as LINE-1s *[90].* Interestingly, these process have been shown associate with exogenous factors *[26]* (Figure 13).

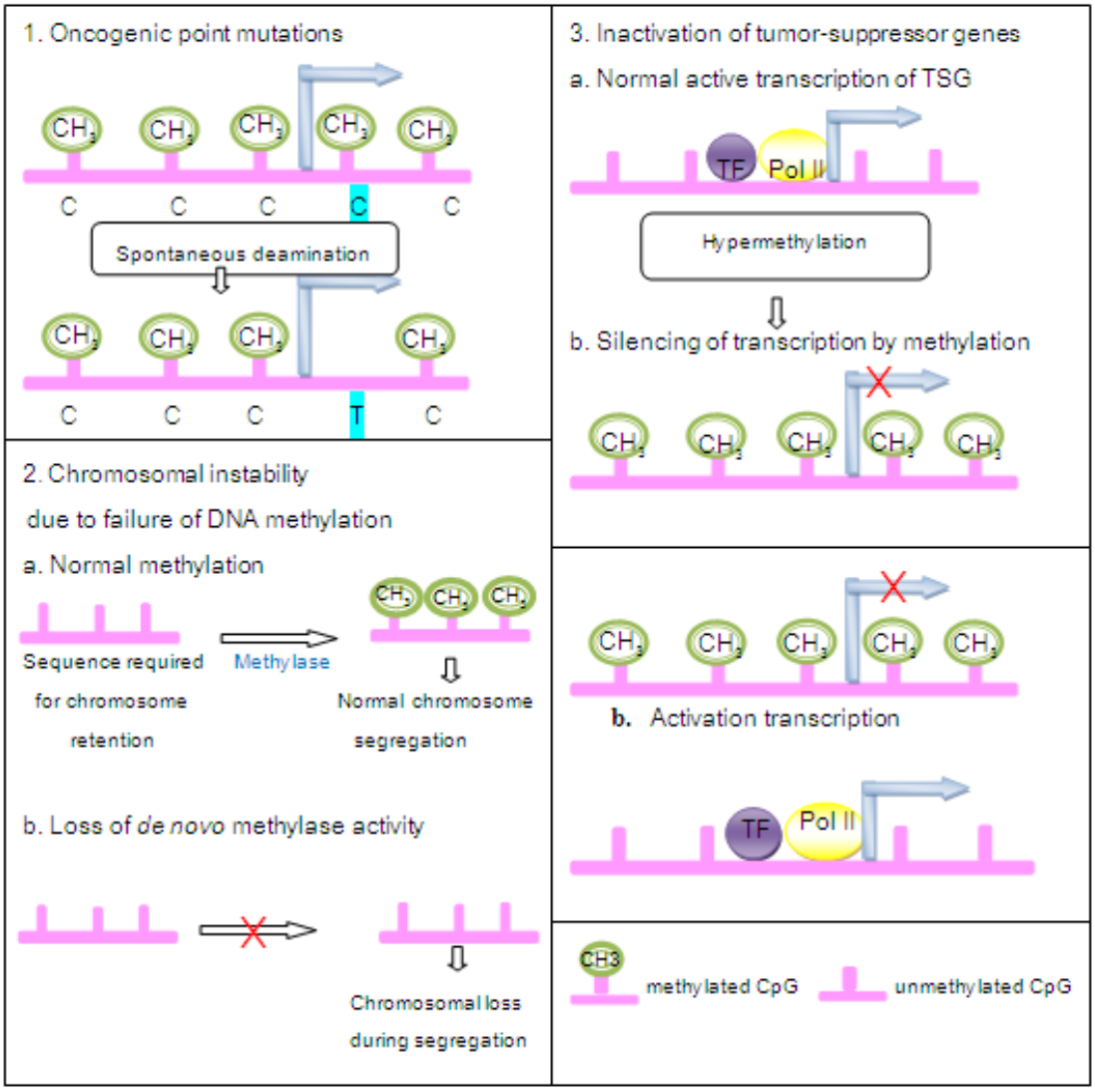


Figure 12 Propose mechanisms for the different processes through which DNA methylation can induce carcinogenesis (modified from Singal R and Ginder GD, 1999) *[83].*

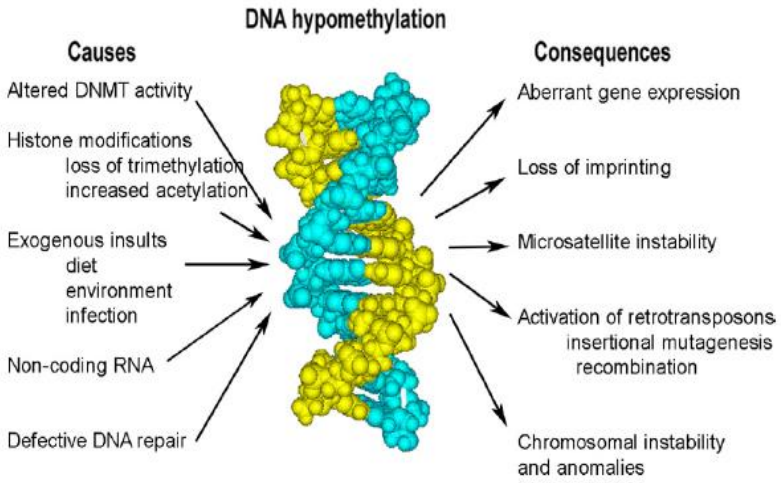


Figure 13 Potential causes and consequences of DNA hypomethylation *[82].*

Methylation Analysis

Currently, there is a wide range of approaches to obtaining quantitative and qualitative information on changes in genomic DNA methylation *[38]*. The method should be clinically ideal diagnostic tests are those which can be carried out on readily accessible body fluids (e.g., serum, urine, saliva, etc.). Such tests should be sensitive, specific, reproducible, cost-effective and minimal number of steps *[84].* Several methods exist to detect changes in the DNA methylation pattern such as methylation-specific Polymerase Chain Reaction (MS-PCR), COBRA, methylation-specific single-strand conformation polymorphism (MS-SSCP) and MethyLight (Figure 14). Most of these techniques use bisulfite treatment which unmethylated DNA is distinguished from methylated and used as a standard procedure prior to validation assays. The underlying principle is based on the ability of sodium bisulfite to deaminate C residues into U in genomic DNA, whereas the methylation C residues are resistant to this modification. After PCR amplication, the U residues are amplified as Ts. Cloning and subsequent sequencing of the DNA fragments containing the CpGs then provide information on the methylation status of each C within the CpGs *[38]*.

1. **Bisulfite sequencing** is a “gold-standard” method used to determine the methylation status of each C over an amplified region of a given gene. The method is used routinely in analysing the methylation status of any target or candidate DNA sequence containing CpGs. It has the advantage of revealing the methylation status of each CpG dinucleotide within the sequence and also the interrelationship between the methylation status of multiple CpG sites. DNA integrity that is less than optimal, as is the case of DNA isolated from microdissected samples or paraffin-embedded tissues, presents significant challenges for this application *[38]*.

2. **MS-PCR**, although MSP can determine DNA methylation with exquisite sensitivity, it is neither quantitative nor can be performed in a high-throughput manner. Another disadvantage is that it allows screening of only a very small sequence (<40 bp) of DNA within a CpG island for methylated residues *[84].*

3. **COBRA**, sequence changes in DNA that result as a consequence of bisulfite-conversion are analysed through restriction enzyme digestion of the resulting PCR product. The main limitation of this approach is that it is only useful for probing DNA methylation status of those CpGs that are harbored within a restriction enzyme site *[84].*

4. **MS-SSCP** provides a fairly quantitative method to access the methylated and unmethylated allele populations. It uses high-resolution gel electrophoresis to generate a specific methylation pattern for determining the percentages of methylation in a targeted sequence *[38]*.

5. **Methylight** is another method of investigating methyated and unmmethylated DNA with impressive sensitivity. This amethod may be performed either in a semi-quantitative or quantitative application and report nucleic acid amplification in real-time without requiring gel electrophoresis *[84].*

Epigenetic alterations in DNA methylation change in cancer demonstrate an interesting treatment point. Because of they are reversible than genetic events. However, in clinical application, the great advantage of DNA methylations tend to be

molecular diagnostics and early detection *[92].* These assays are particularly interesting because of the stability of the DNA *[92, 93]* that present a more clinically and biological source of molecular diagnostic information than RNA or most protein, can be compare with absolute reference points, tend to be universal for individual markers *[92]* and the potential sensitivity of the assays *[93].* However, many methods and technologies are approachable to investigate changes in methylation of C in DNA sequence; each has advantages, disadvantages and areas of applicability. Because of various sample size, the property of the samples, the experience of the examiners and the devices of the laboratory or the institution, there is no "standard procedure" for determining a DNA methylation analysis *[38]*.

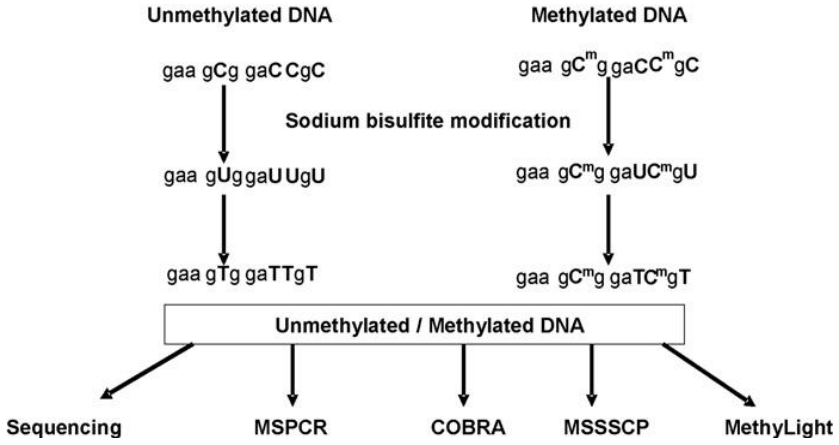


Figure 14 Method for determining methylation analysis.

Different methods can be selected to get the information on the overall interpretation of genes revealing methylation status of DNA in different conditions. These methods used bisulfite treatment principle before uncover the methylation status *[38]*.

<u>Repetitive Sequence</u>

Transposable elements were discovered in maize by Barbara McClintock over 50 years ago now they are known to be to the most abundant component of probably all eukaryotic genomes. These are fragments of DNA that can insert into new

chromosomal locationsn and often make duplicate copies of themselves in the process. They account for at least about 45% of the human genome *[94]* (Figure 15).

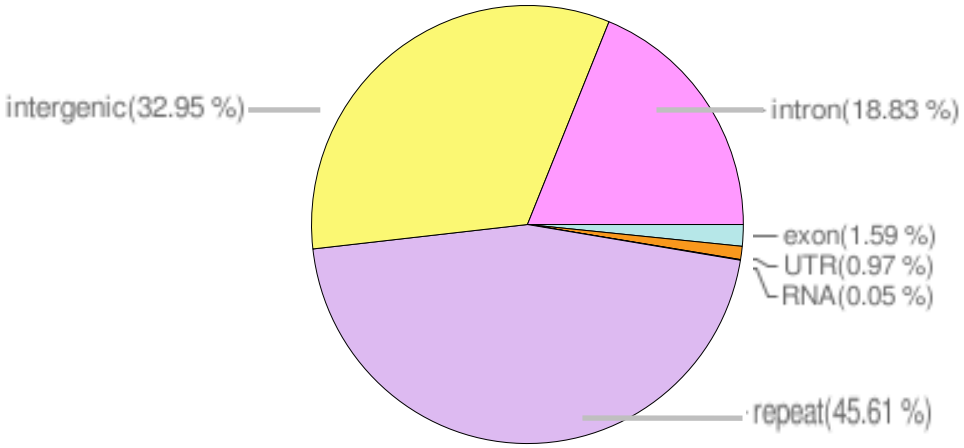


Figure 15 The organization of human genome.

The human genome contains large portion of repeating sequences that constitute about 45% of the genome *[95].*

Almost half of the mammalian genome is derived from ancient transposable elements. The two general types, DNA-transposons and retroelements, often regarded as selfish DNA parasites or junk DNA, encompass 2.8% and 42.2% of the human genome, respectively. Whereas DNA-transposons amplify without an RNA intermediate, retroelements are able to move and integrate into other parts of the genome via a copy and paste mechanism involving an RNA intermediate. Classification of retroelements constitute 90% of the 3 million transposable elements present in the human genome. They are split into two large groups, the non-LTR (long terminal repeat) and LTR elements (Figure 16). There are three non-LTR members which present in extremely high copy numbers in the mammalian germ line including processed pseudogenes, SINE (short interspersed nucleotide elements) and LINE. Processed pseudogenes and SINEs have no protein coding capacity and depend on LINE elements for their amplification. LINE containing the autonomous LINE-1 and LINE-2 sequences *[96].* Distribution of these elements within mammalian genomes is heterogeneous and non-random with densities varying across chromosomes *[82].*

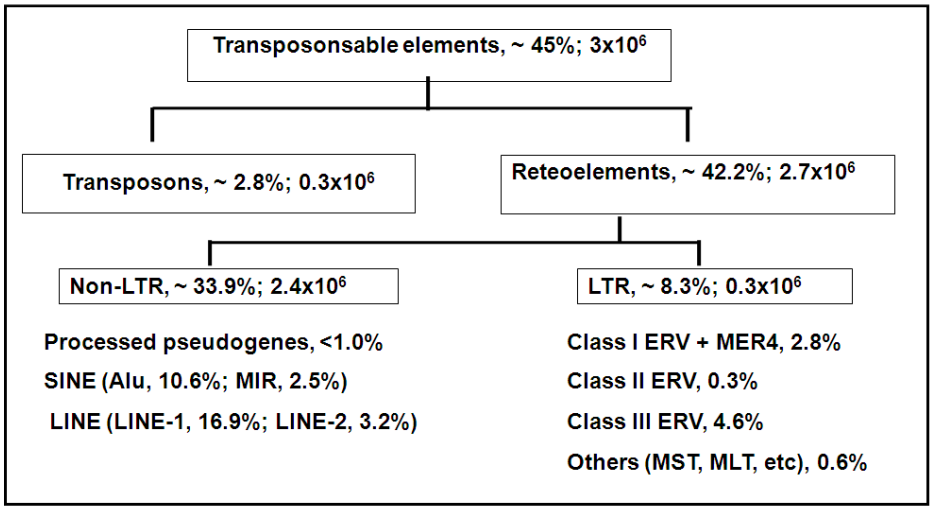


Figure 16 Classification of transposable elements.

The percentage of each element in the genome and the estimated number of the elements of the main groups *[96]*.

- LINE-1

Non-LTR retrotransposons are typified by LINE-1 elements of mammals *[97].* LINE-1 sequences are a large family of transposable elements found in the human genomes, constitute about 16.9% of the human genome with up to 600,000 copies present. Approximately 3000–4000 copies remain in a full-length form and some may be retrotranspositionally active *[37]*. Complete elements are 6.0 kb long *[25, 82, 98]* that are dispersed throughout their host genome as a result of long-term evolution *[99].* They usually have two open reading frames. The ORF2 contains en (endonuclease), rvt (reverse transcriptase) domain as well as a C-rich domain. 5´UTR (5´ untranslated region) contains also internal promoter for RNA polymerase II. 3´UTR contains canonical polyadenylation signal (AATAAA) and a polyA tail (that is also normally absent from the ordinary genes and is only added to mRNA by action of polyA polymerase). LINE-1 is flanked by TSD (target site duplication) that arises during the target primed reverse transcription *[97]* (Figure 17). Because these elements possess strong internal promoters and encode enzymes that enable integration anywhere in the genome *[82]* and encode activities necessary for their retrotransposition, they are called autonomous

even though they probably also require host proteins to complete retrotransposition *[97].* These DNA elements were the most studies example in human retrotransposon, which is the only known LINE active in the human genome *[100].*

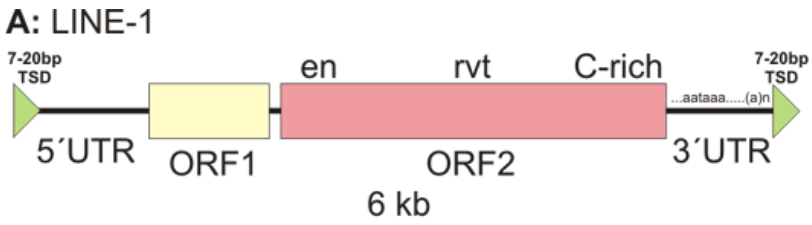


Figure 17 Illustration full-length of LINE-1 element.

LINE-1 is 6 kb consist of a 5'UTR containing an internal promoter, two ORFs, a 3'UTR and a poly (A) signal followed by a poly (A) tail (a)n. LINE-1s are usually flanked by 7- to 20-bp TSDs. The rvt activities, en and a conserved C-rich domain are shown *[97].*

LINE -1 Methyation and Cancer

A number of LINE subfamilies exist ranging in age. The oldest of these are mostly degenerate. However, the younger subfamilies of human specific LINE-1 can still transcribe when activated *[82].* The majority of LINEs are defective due to truncation at the 5' end or internal mutations. Therefore, much smaller than their reported in full length of LINE-1 *[25, 82].*

Disruption of genes by insertion of LINE-1 elements has been found in human cancer and genetic disease. In addition, LINE-1 sequences have been identified at or near chromosomal translocation sites. Presumably to prevent such accidents, most elements are highly methylated in normal adult tissues. Methylation of the LINE-1 promoter sequence has been shown to repress its activity *[25].*

Though it is not completely cleared, it is generally accepted that insertional mutagenesis events leading to alleles associated with diseases in humans are mostly caused by transpositionally active LINE examples include hemophilia A caused by the

disruption of the factor VIII gene on the X-chromosome, insertion into the dystrophin gene in some muscular dystrophy patients. In addition, insertion into somatic cells is mostly irrelevant unless proto oncogenes, tumour suppressor genes, or cancer promoting genes are the targets. Insertion of an LINE-1 element into c-myc was shown to be implicated in a breast carcinoma case or to cause colon cancer if inserted into the adenomatous polyposis coli (APC) gene *[96]* (Table 4).

Table 4 Cases of insertional mutagenesis and recombinations caused by retroelements *[96]*.

Element	Gene	Functional role
LINE-1	Factor VIII	Hemophilia A
LINE-1	Dystrophin	Muscular dystrophy
SINE	Fukutin	Muscular dystrophy
Alu	NF1	Neurofibromatosis
LINE-1	myc	Breast carcinoma
LINE-1	APC*	Colon cancer
LINE-1	Attractin	Soluble protein form
HERV-E	Amylase	Activation of a promoter
HERV-K	FGFR1 kinase	Myeloproliferative disorder
HERVs	AZFa region	Male infertility

From previous study, repetitive DNA elements suffer abnormal hypomethylation, with potential loss of silencing in cancer cells. When used a genome-wide microarray approach to measure DNA methylation changes in HNSCCs and to compare these

changes to alterations found in adjacent non-tumour tissues. Loss of DNA methylation is most pronounced for certain members of the SINE-variable-number tandem repeat-Alu (SVA), human endogenous retrovirus (HERV), LINE-1P, AluY and mammalian apparent LTR retrotransposons (MaLR) families. The methylation levels of retrotransposons are discretely stratified, with younger elements being highly methylated in healthy tissues, while in tumours, these young elements suffer the most dramatic loss of methylation. This study suggested that, in non-tumour adjacent tissues, there is generalized and highly variable disruption of epigenetic control across the repetitive DNA compartment, while in tumour cells, a specific subset of LINE-1 retrotransposons that arose during primate evolution suffers the most dramatic DNA methylation alterations *[101]* (Figure 18).

LINE-1 hypomethylation can arise early in pre-cancerous lesion and has been affected in many cancers when compare to their normal tissues or unaffected adjacent tissues including cancers of the colon, prostate, liver, lung, breast, oesophagus, stomach, urothelial, ovarian, leukemias and head and neck. Moreover, in most cancers studies such as leukemias, urothelial, ovarian and breast cancers, LINE hypomethylation raises with the progression of cancer and has been shown to associate with clinical measurement *[82].*

LINE-1 methylation levels reflect global methylation status in the whole genome *[26]* and in cancerous cells. The methylation levels of most LINE-1 loci demonstrated a positive correlation with each other and with the genome-wide levels. Therefore, the loss of genome-wide methylation in cancerous cells occurs as a generalized process *[37]*. In previous study demonstrated that COBRA LINE-1 could efficiently evaluate the genome-wide methylation status of LINE-1s in genomic DNA and it represents the whole genome methylation status *[12, 102]*. It also found in HNSCCs and OSCCs which revealed a promising trend toward hypomethylation than normal oral epithelia by using COBRA LINE-1 *[12, 37, 102]*. By measuring the quantity of LINE-1 methylation found these levels are varied among type of normal tissues from different organ and was independent of

age and gender *[12, 102-104]*. While methylation levels in most tissues were narrow range distribution, some tissues such as thyroid and esophagus revealed widely distributed. Generally, cancerous cells revealed a lower percentage of methylation comparing with their normal tissue counterparts, except cancers of kidney, thyroid and lymph mode *[12]* (Figure 19).

In some events, due to the combination of multiple mechanisms, the measurement of level alone may not be able to detect LINE-1 methylation changes even if the changes can promote cancer development. In normal cells, some mechanisms of LINE-1 methylation are to control genomic integrity and regulate gene expression in *cis*, locating on the same chromosomes *[20, 34-36]*. Consequently, genomic instability and repression of gene expression can be observed on chromosomes in which LINE-1s are hypomethylated.

Recently, there is a wide range of approaches to obtain quantitative information of genomic DNA methylation *[38].* Most standard techniques measured several CpGs of each LINE-1. Pyrosequencing often measures four CpG dinucleotides *[39]*, whereas COBRA often measures two CpGs *[12]*. However, only COBRA is able to demonstrate the pattern information by distinguishing LINE-1 loci depending on their methylation statuses. For this reason, we classified the methylation statuses ofLINE-1 loci using COBRA to determine the methylation pattern of the 2 CpG dinucleotides in each LINE-1 sequence *[40, 41].* This technique differentiated LINE-1 sequences into 4 methylation-status categories: hypermethylated, hypomethylated and 2 forms of partially methylated loci. We also show that, unlike the case with conventional LINE-1-methylation levels, the percentage of hypomethylated loci (%$^{U}C^{U}C$) can be used to significantly distinguish between normal cells and cancer cells from OSCC and cancers of the nasopharynx, lung, liver and colon *[40, 41].*

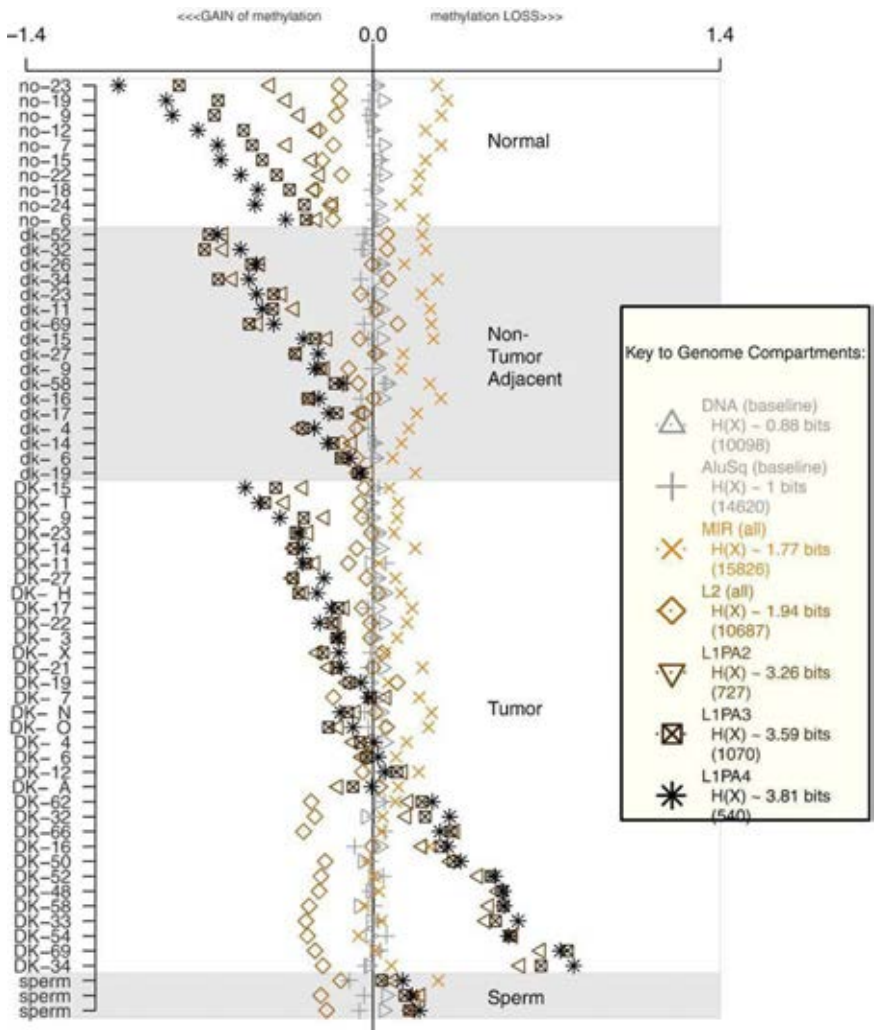


Figure 18 Methylation levels of repetitive element categories per experiment.

Numbers in parenthesis indicate how many probes were averaged per experiment. Comparing the values across the classes of repetitive elements, that younger, primate-specific classes of LINE-1 elements; (LINE-1PA3 (L1PA3), LINE-1PA4 (L1PA4) and LINE-1PA5 (L1PA5) are more strongly methylated in normal tissue and suffer more dramatic losses in DNA methylation in tumours and sperm *[101].*

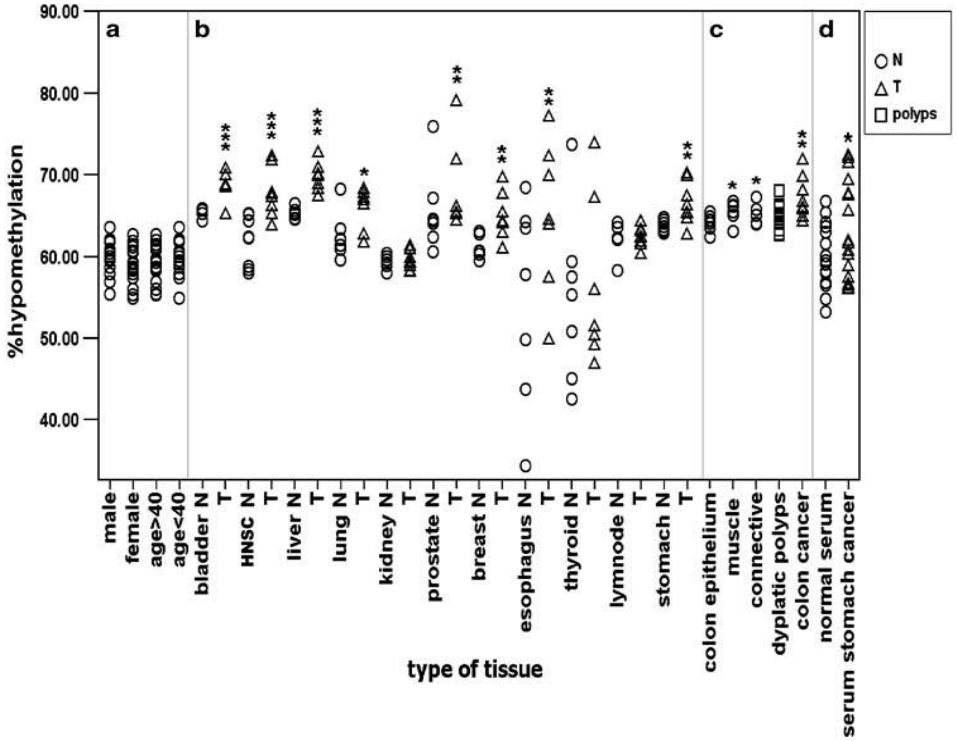


Figure 19 LINE-1 hypomethylation levels in several tissue types.

Circles, triangles and squares are levels of COBRA LINE-1 from normal, malignant of LINE-1 methylation. Sample types are labeled. (a–d) are the hypomethylation levels of leukocytes, cancers, microdissected colonic tissues and sera, respectively. Single, double and triple asterisks indicate significant differences in hypomethylation levels between normal tissues and the tested samples at $p<0.05$, <0.01 and <0.001, respectively. HNSC stands for head and neck squamous cell. N and T are normal and malignant tissues, respectively *[12].*

Intragenic LINE-1 Methylation

Gene containing LINE-1 (intragenic LINE-1) is controlled by DNA methylation status and the transcription activity of a LINE- 1 element is directly correlated with its hypomethylation level. Previously, Aporntewan et al., 2011 proved that lower methylation of intragenic LINE-1s in cancer resulted in enhancing LINE-1 transcription and repressing host genes by double stranded RNA and AGO2 complex (Figure 20) *[36]*.

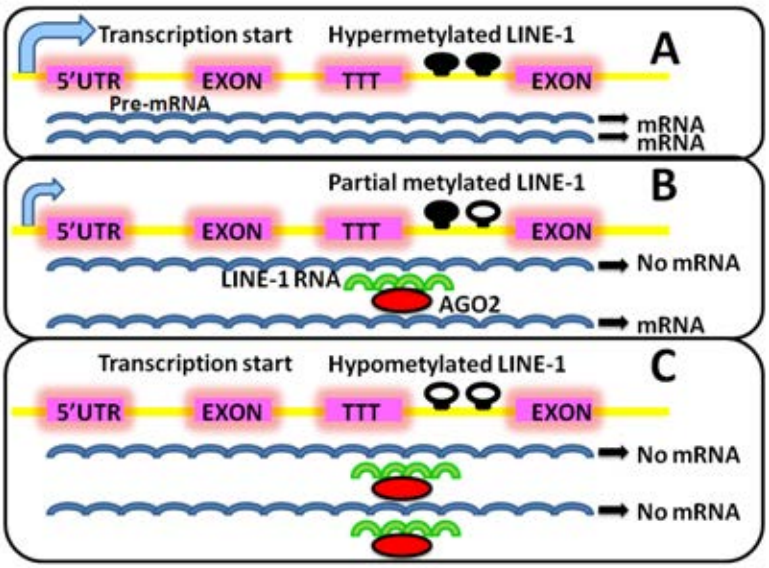


Figure 20 The mechanism of hypomethylation in Intragenic LINE-1s could repress host gene expression via AGO2 complex (modified from Kitkumthorn N and Mutirangura A, 2011).

The diagram reveals that the same gene from three different cells has different levels of intragenic LINE-1 methylation. When intragenic LINE-1 hypermethylation, normal host gene transcription and expression are processed (A). However, in condition of LINE-1 partially methylation, LINE-1 RNA is produced when the methylation of the intragenic LINE-1 is reduced. The LINE-1 RNA–pre-mRNA complex is bound by AGO2 and mRNA production is prevented. Therefore, host gene transcription and expression are partially inhibited (B). In the last condition when intragenic LINE-1 hypomethylation. The situations as same as (b) were found. However, host gene transcription and expression are completely inhibited (C) *[36]*.

LINE-1 Methylation and Tobacco

One of the main characteristic of oncogenesis is epigenetic aberrations. Hypermethylation of the promoters of specific tumour suppressor genes [13] and

genome-wide DNA hypomethylation have been studied in many cancer types *[12].* Interestingly, global DNA hypomethylation is influenced by tobacco smoking *[26, 31]*.

HNSCC patients who smoked had lower methylation level of LINE-1s than HNSCC who did not smoke *[16]*. In addition, the hypomethylation of LINE-1 significantly increases the risk for HNSCC. However, LINE-1 methylation levels in blood samples from patients with HNSCC slightly increased with higher pack-years of smoking *[26]*. The effects of smoking on LINE-1 methylation levels to non-cancerous cells both in vivo and in vitro were also reported. No change was observed in blood cells and colonic epithelium, *in vivo [30-32].* However, a study *in vitro* revealed a minimal change of respiratory epithelium under high dosage cigarette smoke condensate treatment *[33].*

The variation of methylation level of LINE-1s in normal tissues and the hypomethylation in cancerous cells have been confirmed by several studies. These data indicate that the epigenetics plays an important function for cellular reactions, not only in cancer cells, but also in normal cells. In addition, studies that evaluated the association between smoking which is the high risk to oral cancer and repetitive sequence methylation changes *in vivo* have not yet been conclusive. Since oral epithelia directly contact with chemical gradients of tobacco. Herein, we evaluated the possibility that smoking may promote cancer development via genomic hypomethylation by evaluating the LINE-1 methylation pattern in non-cancerous persons of smokers in comparison with non-smokers.

CHAPTER III

MATERIALS AND METHODS

Sample

- **Sample Selection**

Inclusion criteria for non-smokers (NS)

1. No cancer of oral or other organs.
2. No oral mucosal ulcer or lesion.
3. History of never smoking may be with or without alcohol consumption and betel chewing.

Inclusion criteria for current smokers (CS)

1. No cancer of oral or other organs.
2. No oral mucosal ulcer or lesion.
3. History of smoking may be with or without alcohol consumption and betel chewing.

Every subject was provided with research information, benefits and protocols approved by the ethics committee of the Faculty of Dentistry, Chulalongkorn University, Bangkok (7/2010). After receiving the written consent forms, which had been completed by each participant, sample collection process was begun.

All participants were given a self-administered questionnaire to collect medical history including having tumour/cancer of other organs. The demographic information, as well as information of tobacco, alcohol and betel consumption were also recorded. Smoking history was ascertained with an instrument that assesses the number of years smoked, the number of cigarettes smoked daily, age at which an individual started smoking and the number of years since quitting. Similar information was obtained about consumption of alcohol and betel chewing habit. The thoroughly oral examination was

performed in all participants. The volunteers who have no oral mucosal pathology were asked for sample collection by oral rinses.

These subjects were classified into 2 groups.

Group I: NS, subjects who have no previous history of smoking.

Group II: CS, subjects who are currently smoking

Non-smokers (60 volunteers) were included in control group (35 males and 25 females) and current smokers (96 volunteers) were included in study group (80 males and 16 females). By coincidence, 17 smoker volunteers had stopped smoking for more than 1 year, thus they were classified as former smoker (FS) group. Consequently, all of the subjects were finally classified into 3 groups.

- **Sample Size**

Calculation of the sample size for hypothesis testing of two populations from CS and NS were obtained from pilot study.

$$n = \frac{2\sigma^2 (Z_{1-\alpha/2} + Z_{1-\beta})^2}{(\mu_1 - \mu_2)^2}$$

If $n_1 = n_2$

$$\sigma^2 = \frac{(S_1^2 + S_2^2)}{2}$$

n=sample size by group

σ=standard deviation

α=0.05 (at 95% confidence interval)

β=0.1 (power of test 80%)

μ_1=means of methylation levels of LINE-1s in oral epithelium of smoking subjects

μ_2=means of methylation levels of LINE-1s in oral epithelium of non-smoking subjects

The data was obtained from a pilot of 10 NS (5 males and 5 females) and 10 CS (5 males and 5 females). Then, determining LINE-1 methylation levels by independent *t*-test for hypothesis testing of two population means after the two population means were normally distribution were calculated as follow.

μ_1	S_1	μ_2	S_2	$Z_{1-\alpha_2}$	$Z_{1-\beta}$	S_1^2	S_2^2	s^2	$(Z_{1-\alpha/2}+Z_{1-\beta})^2$	$(\mu_1-\mu_2)^2$
34.209	2.76071	36.8164	1.17486	1.96	1.282	7.62	1.38	4.50	10.51	6.80

Calculation from sample size formula

$$n = \frac{2\sigma^2(Z_{1-\alpha/2} + Z_{1-\beta})^2}{(\mu_1 - \mu_2)^2}$$

=13.92

Therefore more than 14 subjects should be included in each group.

Oral Mucosal Cell Collection

NS, FS and CS who had no oral lesion were included in this study. Oral epithelia were collected from oral rinse. Ten millilitres of sterile 0.9 % normal saline solution was gargled for 15 seconds. This solution was kept in a sterile tube and stored at 4°C until the DNA extraction process.

Genomic DNA Extraction

After oral rinses were centrifuged at 4°C, 2500 g for 15 minutes, the supernatant was discarded. The cell pellets were washed twice in sterile PBS. One millilitre of the DNA extraction buffer with 10% SDS and proteinase K (0.5 mg/ml) was added to the cell pellets. The mixtures were then incubated at 50°C for two nights. A phenol-chloroform extraction was used to purify and desalt the digested cell pellets. After centrifuging at 4°C, 14000 g for 15 minutes, 10 M ammonium acetate and cold absolute ethanol were added to the upper aqueous phase for DNA precipitation. The precipitated DNA was washed with 70% ethanol. The air-dried DNA was then resuspended in Tris-EDTA-treated water.

COBRA LINE-1

COBRA for LINE-1 was performed as previously described, the 5'UTR of LINE-1.2 sequence from NCBI Accession Number M80343 was used *[12]*. COBRA LINE-1 consists of three main steps; sodium bisulfate treatment, LINE-1 amplification by PCR and specific restriction enzyme digestion. This quantitative technique is able to determine LINE-1 methylation levels from small amounts of DNA. The standard approachs of COBRA technique are

1. Perform a bisulfite reaction on the DNA
2. Amplify the target site by strand-specific PCR with the design primers
3. Perform digestion by specific restriction enzymes
4. Identify the cut products by gel electrophoresis
5. Measuring the band density (Figure 21)

1. **Sodium Bisulfite Treatment**

- **Principle**

The conversion process of DNA sample is performed in the bisulfite reaction (Figure 22). The DNA samples were converted by a bisulfite reaction such that unmethylated cytosine (^{u}C) would be converted to uracil (U), whereas methylated cytosine (^{m}C) would remain as cytosine (C) (Figure 23).

- **Technique**

Genomic DNA 500 ng in 20 µl water were denatured in 0.2 M NaOH at 37^{o}C for 10 minutes and then incubated with 30 µl of 10 mM hydroquinone and 520 µl of 3 M sodium bisulfite at 50^{o}C, 16-20 hours. After that, bisulfite-treated DNA was desalted with DNA Clean-Up system. Subsequently, it was desulfonated by 0.3 M NaOH and precipitated with ethanol. Finally the DNA was then resuspended in 20 µl of water.

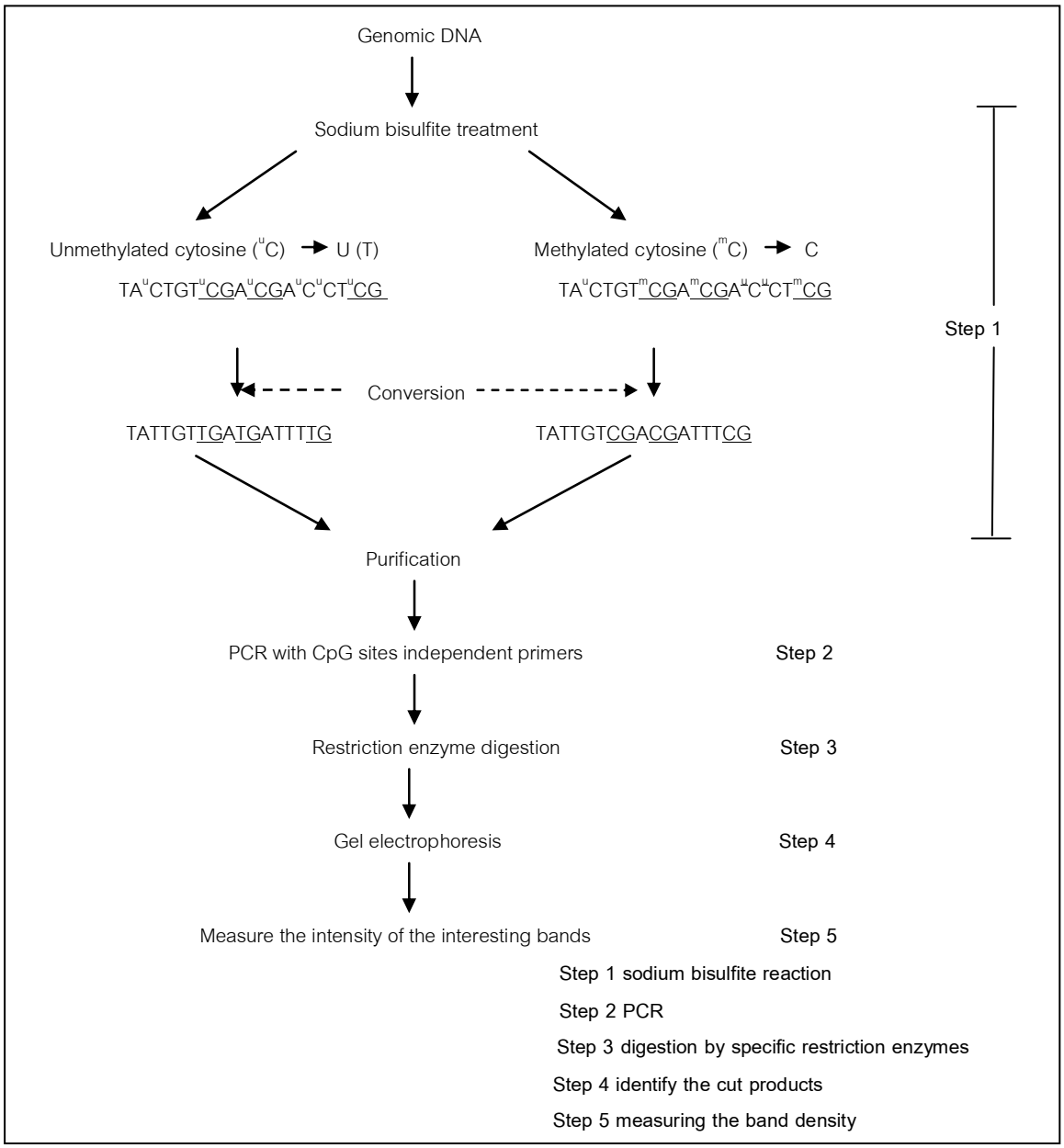


Figure 21 Concept of COBRA.

DNA samples were reacted with bisulfate. Briefly, unmethylated cytosines (^{u}C) are converted to uracils, whereas methylated cytosines (^{m}C) are retained as Cytosines. The sequence of interest is then amplified by PCR and subjected to digestion with enzymes. After the products are identified by gel electrophoresis, the intensity of the interesting size were measured *[105].*

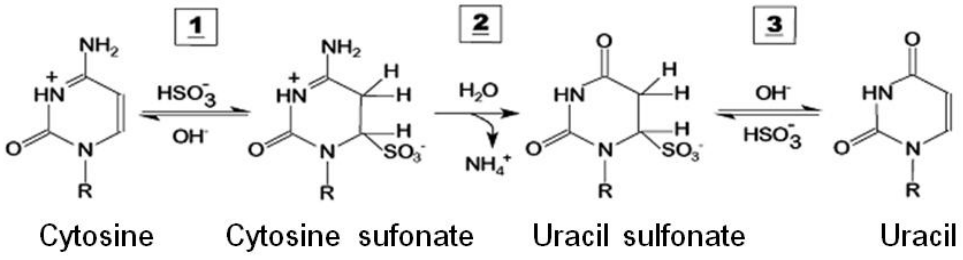


Figure 22 Concept of the bisulfite conversion reaction.

The deamination of cytosine by sodium bisulfite treatment from cytosine to give uracil *[106].*

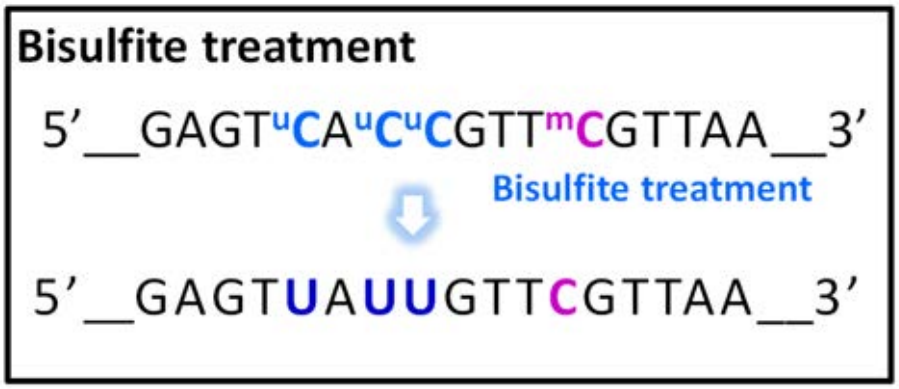


Figure 23 DNA sequences after bisulfite modification

After treatment with the process of bisulfite, unmethylated cytosine (^{u}C) would be converted to uracil (U), whereas methylated cytosine (^{m}C) would remain as cytosine (C).

2. PCR

- Principle

PCR is the method of choice for DNA amplification, both as analytical and as diagnostic technique, because it rapidly generates a large number of copies of the target DNA sequence. Refinement of the PCR technique has allowed the detection of DNA fragments in samples where the quantity and/or the quality of DNA present is too

low to permit other types of molecular analysis. In particular, it is now possible to detect the presence of a target DNA sequence in small samples in which the DNA has been heavily degraded by aging and/or processing treatment *[107].*

After DNA is treated with sodium bisulfite, PCR is performed to amplify LINE-1 sequences. The principle is based on the bisulfite converted Us in genomic DNA will be amplified as Ts, whereas unconverted Cs are resistant to this modification. Thus after PCR, the DNA sequences which contain unmethylated Cs will be changed from their original sequences; while the ones that contain methylated Cs will retain their original sequences (Figure 24).

- **Technique**

One microlitre of bisulfite DNA was then subjected to 35 cycles of PCR , at a 50°C annealing temperature using the following primer sets: LINE-1-F (5'-CCGTAA GGGGTTAGGGAGTTTTT-3') and LINE-1-R (5'-RTAAAACCCTCCRAACCAAATATAAA-3').These DNA were denatured at 95°C, 15 minute and 95°C, 1 minute, annealed at of 50°C, 1 minute and extended at 72°C, 1 minute. The PCR amplicon sizes were 160 bp.

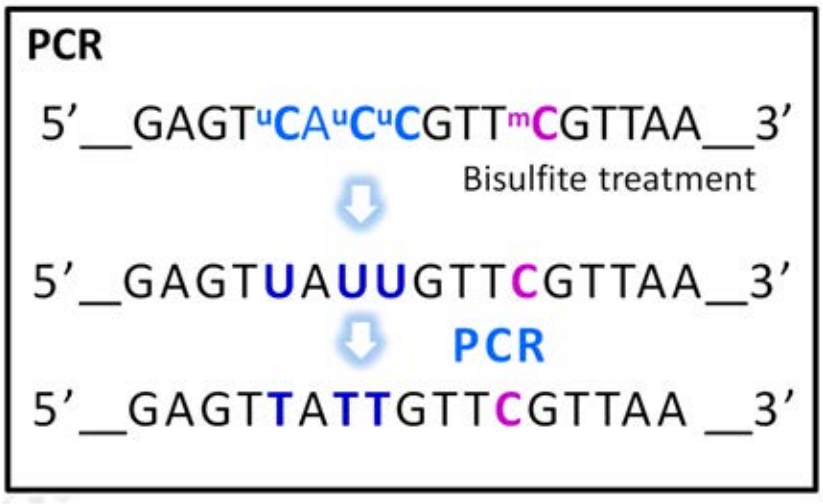


Figure 24 The amplified sequence after bisulfite treatment and PCR.

In PCR, after bisulfite treatment, the biulfite converted uracils (U) will be amplifies as thymines (T), where as unconverted cytosines will be amplified as cytosines.

.3. Specific Restriction Enzyme Digestion

The LINE-1 amplicons (160 bp) were digested with 2 U of *Taq*I and 2 U of *Tas*I (Figure 25) in NEB3 buffer (New England Biolabs, Ontario, Canada) at 65°C overnight. The products were identified by polyacrylamide gel electrophoresis in 8% nondenaturing and stained with SYBR green nucleic acid gel stain (Sigma-Aldrich, St. Louis, Missouri).

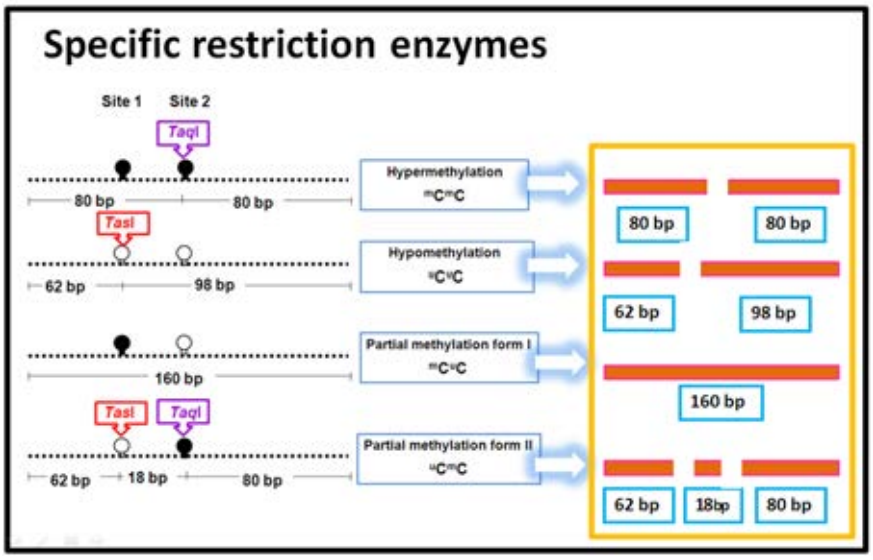


Figure 25 COBRA LINE-1 PCR amplicon with *Taq*I and *Tas*I recognition site.

After bisulfite treatment and PCR, methylated cytosines are demonstrated by black oval marks and unmethylated cytosines are demonstrated by white oval marks. *Taq*I restriction enzyme recognizes methylated cytosine site and *Tas*I enzyme recognizes unmethylated cytosine site 1. Hypermethylation pattern yielded two fragments of 80 bp. Hypomethylation pattern yielded 62 and 98 bp. Partial methylation pattern form I could not be cleaved by any enzyme, providing

160 bp. Partial methylation pattern form II could be cut by both enzymes, giving 62, 18 and 80 bp.

Distilled water was used as a negative control. The same preparation of DNA from 3 cell lines, HeLa (cervical cancer), Daudi (Human Burkitt's lymphoma) and Jurkat (acute T cell leukemia) (ATCC, Manassas, VA, USA) were used as positive controls in all experiments and for inter-assay variation adjustment.

COBRALINE-1 Product Analysis

Here, we classified LINE-1s into four groups depending on the methylation status of 2 CpG dinucleotides on each strand from 5' to 3' detected by COBRALINE-1 as described previously *[41]*. These COBRA-detected LINE-1s were categorised into the following four classes: 2 unmethylated CpGs ($^{u}C^{u}C$), 2 methylated CpGs ($^{m}C^{m}C$), 5'methylated and 3'unmethylated CpGs ($^{m}C^{u}C$), or 5'unmethylated and 3'methylated CpGs ($^{u}C^{m}C$) (Figure 26A). LINE-1 methylation levels and the percentage of loci of each class were calculated from COBRALINE-1 digested products. Intensities of COBRALINE-1 bands were measured by a phosphoimager using ImageQuant Software (Molecular Dynamics, GE Healthcare, Slough, UK). After enzymatic digestion, the COBRALINE-1 amplicons were separated into 5 DNA strands depending on their length, 160, 98, 80, 62 and 18 bp (Figure 26B). The 18 bp band was not used in the following calculation. The 160 bp band contains 2 CpGs, in which the 5'CpG is methylated and the other 3'CpG is unmethylated. The 98 bp band contains 2 unmethylated CpGs. The 80 bp and 62 bp bands each contain 1 methylated and 1 unmethylated CpG. The CpGs of the 160 bp and 98 bp bands were derived from $^{m}C^{u}C$ and $^{u}C^{u}C$, respectively. The CpGs of the 80 bp band were derived from 3'methylated CpGs of $^{m}C^{m}C$ and $^{u}C^{m}C$, respectively, while the CpGs of the 62 bp band were derived from 5'unmethylated CpGs of $^{u}C^{u}C$ and $^{u}C^{m}C$ (Figure 25). To normalise each band to represent the total number of CpG dinucleotides present, the intensity of each band was divided by the number of basepairs of double stranded DNA as follows: %160/160=A, %98/94=B, %80/78=C and

% 62/62=D. Then, the LINE-1 methylation levels were computed with the following formula: percentage of LINE-1 methylation level (%^{m}C)=100×(C+A)/(C+A+A+B+D), percentage number of ^{m}C^{u}C loci (%^{m}C^{u}C)=100×(A)/(((C-D+B)/2)+A+D), %^{u}C^{m}C= 100×(D-B)/(C-D+B)/2)+A+D, %^{u}C^{u}C=100×B/(((C-D+B)/2)+A+D) and %^{m}C^{m}C=100×((C-D+B)/2)/(((C-D+B)/2)+D+A).

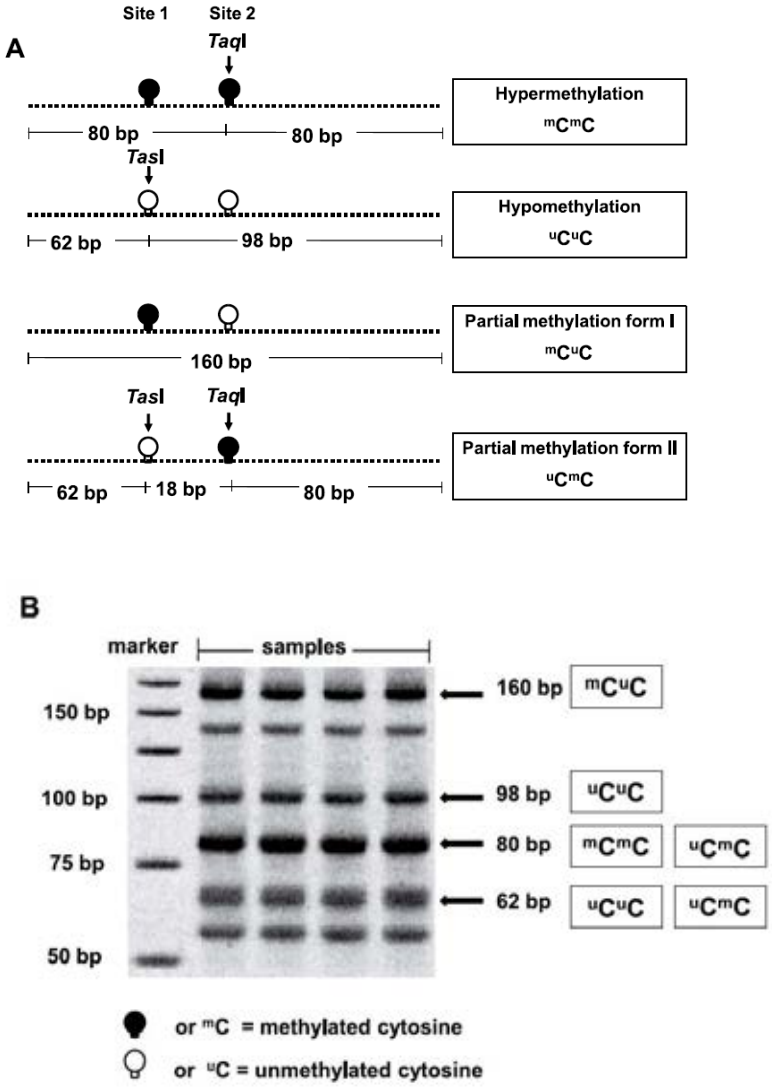


Figure 26 Methylation patterns of amplified LINE-1s.

(A) The LINE-1 amplicons were 160 bp and had 2 CpG dinucleotides. Four patterns of methylated CpGs were detected, including hypermethylation (^{m}C^{m}C), hypomethylation (^{u}C^{u}C) and two forms of partial methylation (^{m}C^{u}C and ^{u}C^{m}C). The

*Tas*I enzyme targets unmethylated cytosine site 1 and *Taq*I targets methylated cytosine site 2. (B) After restriction digestion with *Tas*I and *Taq*I, four sizes of products (160, 98, 80 and 62 bp) were identified, depending on the methylation status.

Connection Up- or Down-Regulation Expression Analysis of Microarrays Extension Program (CU-DREAM-X) for LINE-1s

Recently, the CU-DREAM-X program has been used to observe the association between the up- or down-regulation of genes containing LINE-1 *[108]*. Briefly, "intragenic" and "intergenic" LINE-1s, identified using the NCBI Reference Sequence (RefSeq) annotation *[108]*, were cross-referenced with the L1base (http://l1base.molgen.mpg.de) *[19]*.

System Requirements of CU-DREAM-X

CU-DREAM-X requires a computer with the following settings.

1) Windows operating system.

2) Microsoft .NET framework 3.5 or higher (download from the link below) http://www.microsoft.com/downloads/details.aspx?FamilyId=333325fd-ae52-4e35-b531-508d977d32a6&displaylang=en

3) Microsoft Excel 2007 or higher.

4) Microsoft Office system Primary Interop Assemblies or PIA (download from this link http://www.microsoft.com/downloads/details.aspx?FamilyID= 59daebaa-bed4-4282-a28cb864d8bfa513&displaylang=en)

An Example of CU-DREAM-X

In this section illustrates how to intersect microarray data sets from the Gene Expression Omnibus (GEO). Go to NCBI website (http://www.ncbi.nlm.nih.gov) and search for datasets as below pictures.

1. Create a working directory, for example, C:\1

2. Go to NCBI website (http://www.ncbi.nlm.nih.gov), select at GEO Datasets and search for interesting datasets (Figure 27 and Figure 28).

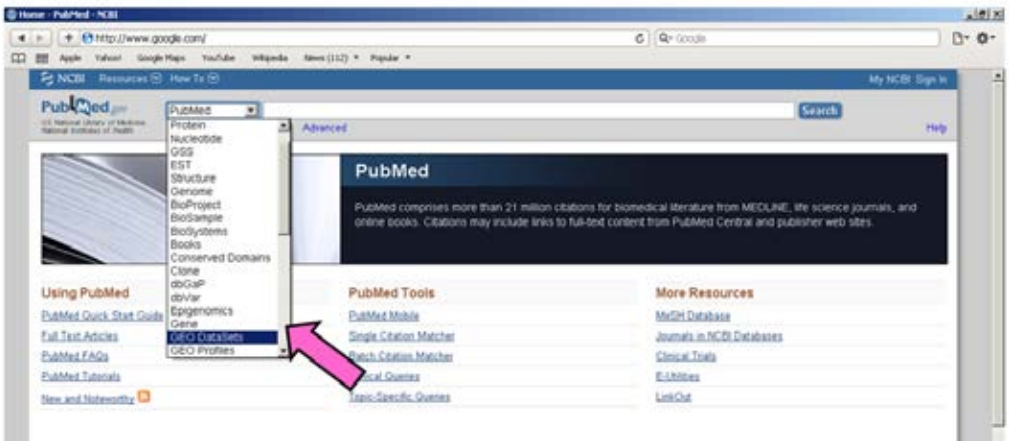


Figure 27 NCBI website: http://www.ncbi.nlm.nih.gov , select at GEO Datasets.

Figure 28 Search for interesting of datasets.

3. Click at the GSE title (Figure 29).

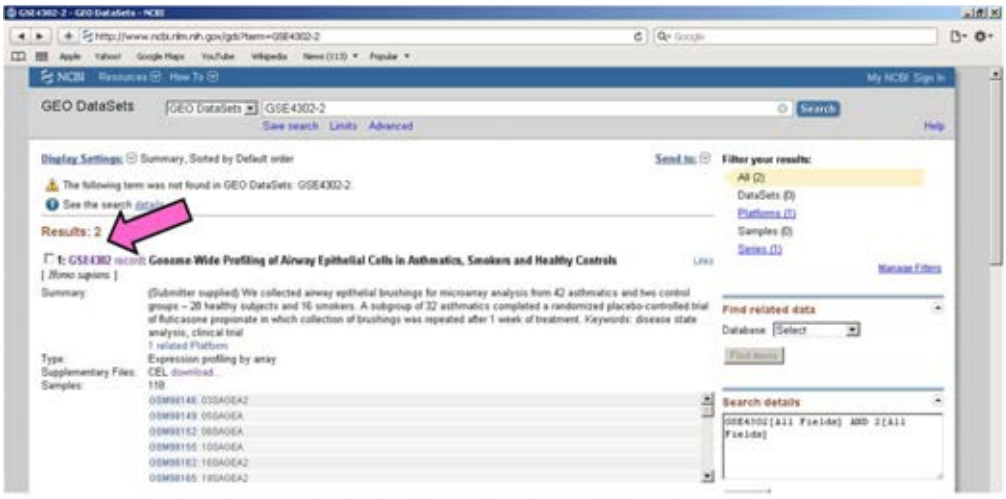


Figure 29 Click at the GSE title.

4. Scroll down the page, download the series matrix file, save it to the working directory and uncompress it (Figure 30).

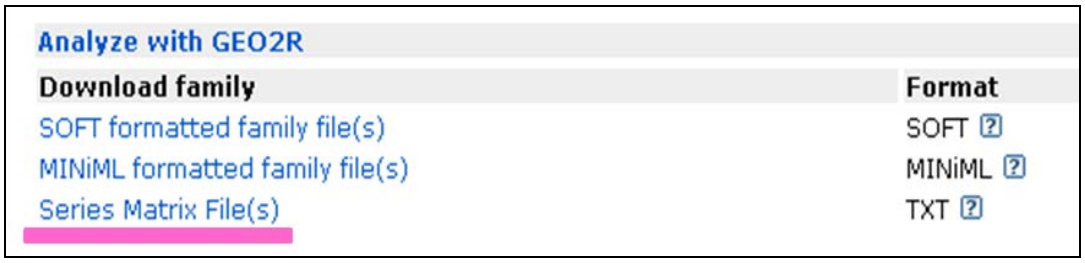


Figure 30 Download the series matrix file.

5. On the same page of the previous step, click at the platform title (Figure 31), then download full table and save it to the working directory (Figure 32)

Figure 31 Click at the platform title.

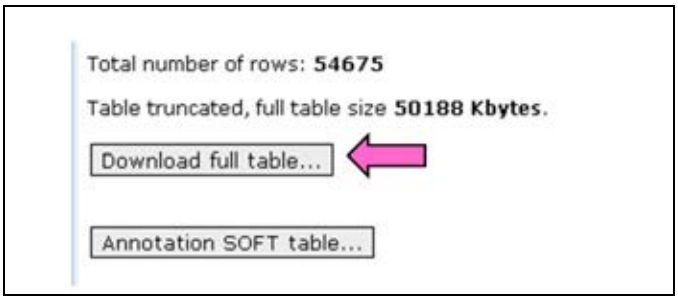


Figure 32 Download the full table.

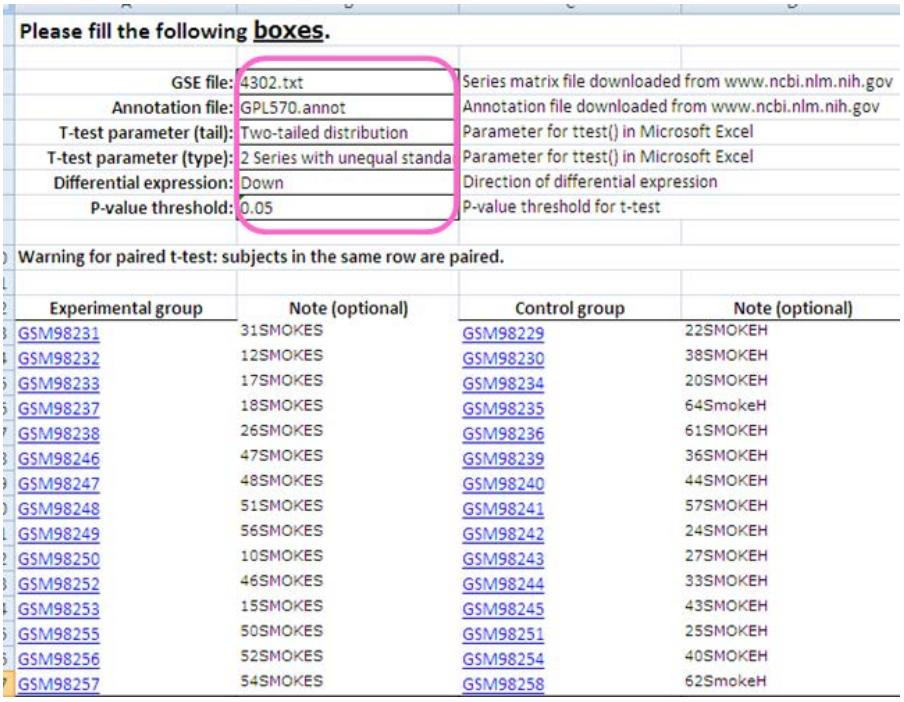
6. Download “template.xls” from website (the URL below), make the copies, name GSE4302.xls and save them to the working directory.

URL: http://pioneer.netserv.chula.ac.th/~achatcha/cu-dream/template.xls

7. In the template.xls, edit GSE4302.xls to set the parameters (Figure 33).

8. Download the executable file from the URL below and save it to the working directory.

URL: http://pioneer.netserv.chula.ac.th/~achatcha/cu-dream/cu-dream.exe



Please fill the following boxes.

GSE file:	4302.txt	Series matrix file downloaded from www.ncbi.nlm.nih.gov
Annotation file:	GPL570.annot	Annotation file downloaded from www.ncbi.nlm.nih.gov
T-test parameter (tail):	Two-tailed distribution	Parameter for ttest() in Microsoft Excel
T-test parameter (type):	2 Series with unequal standa	Parameter for ttest() in Microsoft Excel
Differential expression:	Down	Direction of differential expression
P-value threshold:	0.05	P-value threshold for t-test

Warning for paired t-test: subjects in the same row are paired.

Experimental group	Note (optional)	Control group	Note (optional)
GSM98231	31SMOKES	GSM98229	22SMOKEH
GSM98232	12SMOKES	GSM98230	38SMOKEH
GSM98233	17SMOKES	GSM98234	20SMOKEH
GSM98237	18SMOKES	GSM98235	64SmokeH
GSM98238	26SMOKES	GSM98236	61SMOKEH
GSM98246	47SMOKES	GSM98239	36SMOKEH
GSM98247	48SMOKES	GSM98240	44SMOKEH
GSM98248	51SMOKES	GSM98241	57SMOKEH
GSM98249	56SMOKES	GSM98242	24SMOKEH
GSM98250	10SMOKES	GSM98243	27SMOKEH
GSM98252	46SMOKES	GSM98244	33SMOKEH
GSM98253	15SMOKES	GSM98245	43SMOKEH
GSM98255	50SMOKES	GSM98251	25SMOKEH
GSM98256	52SMOKES	GSM98254	40SMOKEH
GSM98257	54SMOKES	GSM98258	62SmokeH

Figure 33 The template.xls for setting the parameters.

9. Start the “Command Prompt” in Programs → Accessories (Figure 34).

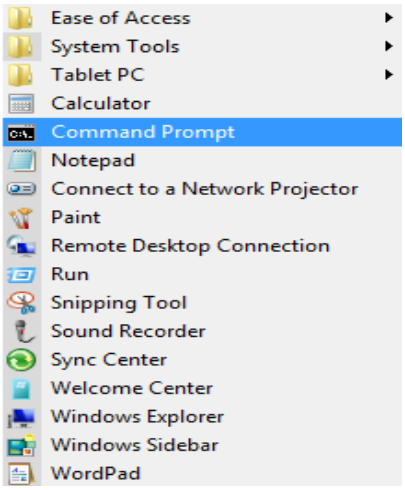


Figure 34 Open the Programs.

10. Change path to the working directory by typing “cd c:\1” and pressing enter (Figure 35).

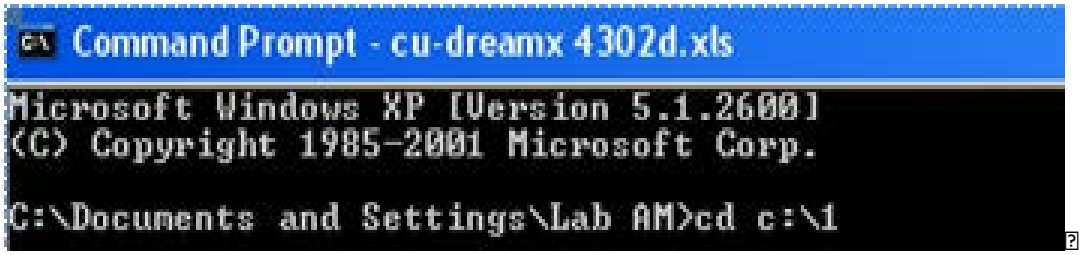


Figure 35 Change path to the working directory.

11. Start the program by typing “cu-dreamx 4302.xls” and pressing enter (Figure 36).

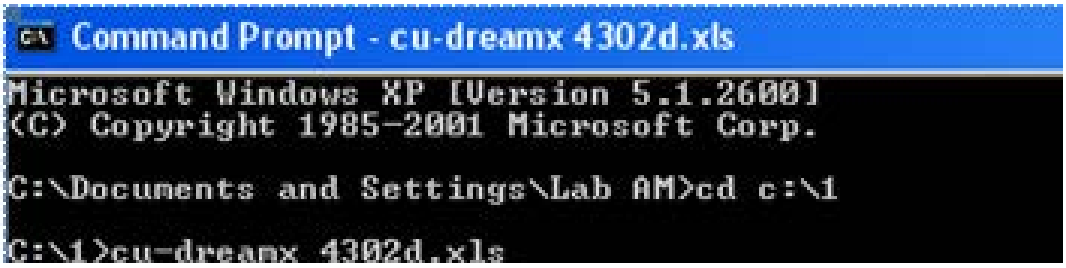


Figure 36 Start the program of CU-DREAM-X.

12. If the program succeeds, you will see the following message as picture (Figure 37).

13. Finally, the file “Intersect_L1_4302d.xls” is obtained in the working directory. The first sheet shows GSE4302 array (Figure 38).

Figure 37 The message after completing the running programs.

	A	B	C	D	E	F	G	H	I
1	Probe ID	Gene Symbol	Mean1	Mean2	Mean1 - Mean2	P-value	GSM98231	GSM98232	GSM98233
2									
3	1007_s_at	DDR1	10.45	10.64	-0.19	3.62E-02	10.56	10.61	10.35
4	1053_at	RFC2	6.05	6.14	-0.09	2.80E-01	5.94	6.21	5.98
5	117_at	HSPA6	6.39	6.24	0.15	1.46E-02	6.29	6.24	6.63
6	121_at	PAX8	7.30	7.33	-0.02	7.10E-01	7.19	7.24	7.30
7	1255_g_at	GUCA1A	4.80	4.75	0.05	7.38E-02	4.89	4.77	4.83
8	1294_at	UBA7	8.03	7.99	0.04	7.88E-01	8.24	8.26	8.27

Figure 38 The experimental array sheet.

The columns, from left to right, are probe id, gene symbol, the mean of experimental group, the mean of control group, differential mean, unadjusted *p*-value (*t*-test), experimental group and control group.

14. The second sheet shows the intersection table, *p*-value, OR, confidence interval and lists of genes in column a, b, c and d (Figure 39)

Simple Count Algorithm in CU-DREAM-X

It is important to address how we count the number of genes for chi-square test in CU-DREAM-X table (Figure 39). The *t*-test is performed on all probes. A gene is up-regulated if "Mean1" (experimental group) is greater than "Mean2" (control group). Otherwise, the gene is down-regulated (Figure 38).

			4302d					
			Down (0.01)	Not down				
		L1	178	1,115	1,293		P-value	9.87E-25
		No L1	1,232	18,198	19,430		Odd Ratio	2.36
			1,410	19,313	20,723		Upper 95% CI	1.99
							Lower 95% CI	2.79
			87.98	1,205.02				
			1,322.02	18,107.98				
a = L1 and (4302d down)								
b = L1 and (4302d not down)								
c = no L1 (4302d down)								
d = no L1 (4302d not down)								
	a	b	c	d				
	KIF6	WFDC2	CYP2E1	DDR1				
	TMEM67	ADAM32	SLC46A1	RFC2				
	SLC44A5	SPATA17	FAM122C	HSPA6				
	RIMS1	C15orf27	PDE7A	PAX8				
	SCAMP1	ADAMTSL1	WFDC6///SPIN	GUCA1A				
	SNX13	LACE1	WDR17	UBA7				

Figure 39 An example of intersection of CU-DREAM-X table.

In this study, we used the methylation arrays from the gene expression omnibus (GEO) data sets of the airway epithelia of smokers (GSE4302-2, GSE19667-1, GSE19667-2, GSE11906-8, GSE11906-5, GSE4498, GSE13933-2, GSE11906-7, GSE3320, GSE8545-2, GSE13933-1, GSE11906-6, GSE27002, GSE7895-1 and GSE994-1) (Appendix I). Independent *t*-tests and chi-squared tests were used to test the significance of the gene symbols for each probe. The effect of intragenic LINE-1s on gene expression or promoter methylation was assessed by this program to verify gene regulation. The resulting models were considered significant if $p<0.05$. The status of the LINE-1 regions was then distinguished from the rest of the genes with a two-way table.

Statistical Analysis

Statistical analyses were performed using SPSS version 17.0 (SPSS Inc., Chicago, IL). All *p*-values less than 0.05 were considered significant. All the variables were normally distributed (Kolmogorov-Smirnov Test). We used a two-way analysis of variance (ANOVA) to determine the effects of two factors, alcohol and smoking, on the methylation levels of LINE-1. An independent sample *t*-test was performed to compare LINE-1 methylation in males and females, NS and CS, and pack-year groups. In addition, the paired *t*-test was used for a matched-case analysis. The chi-squared test and odds ratio (OR) were used to test the association among LINE-1 methylation variables.

CHAPTER IV

RESULT

Part I: Demographic Information, History and Examination

Table 5 Group frequencies in NS CS and FS.

		Non-smokers	Current smokers	Former smokers
Total Subjects		60	96	17
Gender	Male	35	80	15
	Female	25	16	2
Age	Mean ± SD	44.63 ± 14.19	41.60 ± 4.60	46.59 ± 16.39
History of smoking	Currently smoke	-	96	-
	Previously smoke	-	-	17
	Never smoke	60	-	-
History of alcohol drinking	Currently drink	36	86	14
	Previously drink	-	-	3
	Never drink	24	10	-
History of betel chewing	Currently chew	-	-	-
	Previously chew	-	-	-
	Never chew	-	-	-

Part II: The Percentage of each LINE-1 Methylation Pattern in Males and Females

Oral rinses were collected from 60 NS volunteers (35 males and 25 females), 96 CS volunteers (80 males and 16 females) and 17 FS volunteers (15 males and 2 females). No significant differences in the percentage of LINE-1 products were detected between the males and females in either group (Table 6).

Table 6 Percentage of LINE-1 products in Males and Females.

All cases									
	Non-smokers			Current smokers			Former smokers		
	Male	Male	Male	Male	Female	Total	Male	Female	Total
Number of subjects	35 (58.37%)	80 (83.33%)	80 (83.33%)	80 (83.33%)	16 (16.67%)	96 (100%)	15 (88.24%)	2 (11.76%)	17 (100%)
Age (year ± SD)	44.29 ± 17.66	40.78 ± 16.16	40.78 ± 16.16	40.78 ± 16.16	42.29 ± 7.68	41.60 ± 4.60	48.88 ± 15.91	30.00 ± 2.83	46.59 ± 16.39
% ^{m}C (mean ± SD)	42.55 ± 1.82	42.34 ± 2.62	42.34 ± 2.62	42.34 ± 2.62	41.29 ± 2.49	42.16 ± 2.62	41.01 ± 2.44	42.22 ± 2.28	41.16 ± 2.38
% $^{m}C^{m}C$ (mean ± SD)	16.94 ± 4.03	18.21 ± 4.68	18.21 ± 4.68	18.21 ± 4.68	15.93 ± 5.03	17.82 ± 4.79	15.47 ± 4.95	15.49 ± 0.38	15.47 ± 4.61
% $^{u}C^{u}C$ (mean ± SD)	31.84 ± 2.32	33.53 ± 3.83	33.53 ± 3.83	33.53 ± 3.83	33.35 ± 3.38	33.50 ± 3.74	33.46 ± 1.35	31.07 ± 4.95	33.16 ± 1.97
% $^{m}C^{u}C$ (mean ± SD)	26.73 ± 2.13	23.83 ± 3.61	23.83 ± 3.61	23.83 ± 3.61	25.03 ± 3.74	24.03 ± 3.64	24.72 ± 4.14	23.07 ± 2.89	24.51 ± 3.97
% $^{u}C^{m}C$ (mean ± SD)	24.49 ± 4.50	24.43 ± 8.02	24.43 ± 8.02	24.43 ± 8.02	25.69 ± 6.80	24.64 ± 7.81	26.35 ± 5.46	30.37 ± 8.22	26.85 ± 5.67
% $^{m}C^{u}C+^{u}C^{m}C$ (mean ± SD)	51.22 ± 5.48	48.26 ± 6.76	48.26 ± 6.76	48.26 ± 6.76	50.72 ± 6.98	48.68 ± 6.82	51.07 ± 5.37	53.44 ± 5.33	51.36 ± 5.25

Part III: The Percentage of Loci of each LINE-1 Methylation Pattern in NS and CS

Smoking behaviour is closely related to alcohol consumption. However, the association between smoking and alcohol consumption and its contribution to malignant potency has not been completely elucidated. To determine the interaction between alcohol and smoking on the LINE-1 methylation pattern, we used two-way ANOVA. No interactions between alcohol and smoking consumption were found for any of the patterns, including (p>0.05) (Table 7). Therefore, only the possible impact of smoking on LINE-1 was analysed. The percentages of all the patterns are presented in Table 6. The CS had significantly higher %$^{m}C^{m}C$ and %$^{u}C^{u}C$ and lower %$^{m}C^{u}C$ and %$^{m}C^{u}C+^{u}C^{m}C$ than the NS (p=0.002, 0.015, <0.0001 and <0.0001, respectively). However, no significant difference was found in %^{m}C and %$^{u}C^{m}C$ (p=0.327 and 0.835, respectively) (Figure 40).

Table 7 The interaction effect of smoking and alcohol in each methylated CpG pattern

DNA methylation pattern	Type III Sum of Squares	df	Mean Square	F	*p*-value
% ^{m}C	13.457	1	13.457	2.227	0.138
% $^{m}C^{m}C$	44.444	1	44.444	2.317	0.130
% $^{u}C^{u}C$	0.446	1	0.446	0.038	0.845
% $^{m}C^{u}C$	23.407	1	23.407	2.015	0.158
% $^{u}C^{m}C$	117.445	1	117.445	2.351	0.127
% $^{m}C^{u}C+^{u}C^{m}C$	35.998	1	35.998	.960	0.329

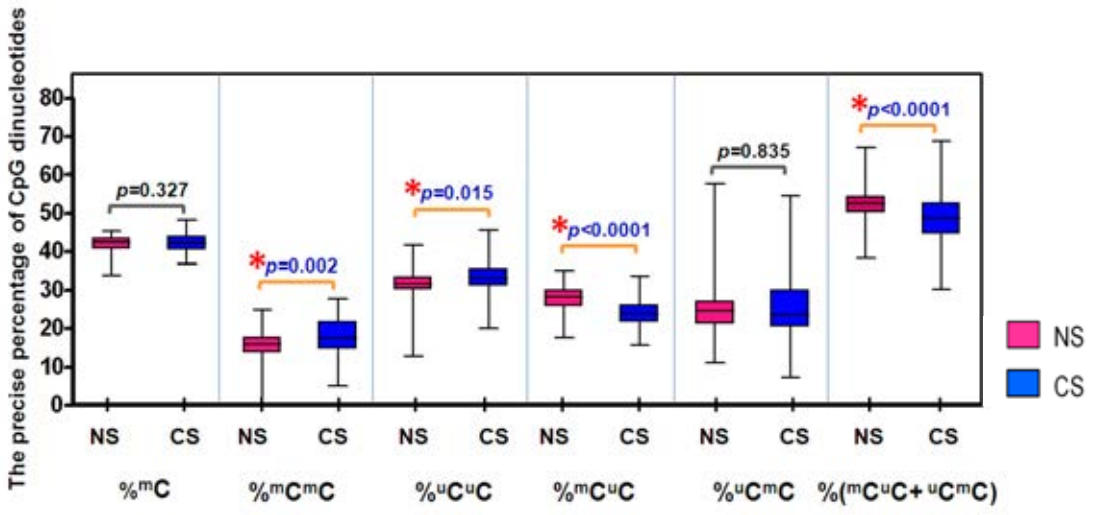


Figure 40 Percentage of each methylated CpG pattern in the subjects.

^{m}C represents the overall methylation level of the amplified LINE-1s, %^{m}C. $^{m}C^{m}C$ and $^{u}C^{u}C$ represent %$^{m}C^{m}C$ and %$^{u}C^{u}C$, respectively. $^{m}C^{u}C$ and $^{u}C^{m}C$ represent %$^{m}C^{u}C$ and %$^{u}C^{m}C$. $^{m}C^{u}C+ ^{u}C^{m}C$ is the sum of partially methylated loci of both forms. The horizontal line within each box indicates the mean of the percentage. Stars indicate statistical significance at $p<0.05$. The results demonstrated that the CS had a significantly higher %$^{m}C^{m}C$ and %$^{u}C^{u}C$ and a lower %$^{m}C^{u}C$ and %$^{m}C^{u}C$+%$^{u}C^{m}C$ than the NS.

Part IV: The Percentage of Loci of each LINE-1 Methylation Pattern in matched cases

Although gender and the interaction of alcohol consumption and smoking did not influence the methylation levels of LINE-1s, no prior study has indicated the impact of age on this analytical method. Therefore, the NS were matched to the CS based on age, gender and alcohol drinking behaviour, which produced 29 pairs (males 14 pairs and females 15 pairs). The same tendency of differences in the LINE-1 methylation patterns as those found in the total sample was found. However, only $\%^{m}C^{u}C$ resulted in a significant difference at $p<0.0001$ (Figure 41, Figure 44A and Table 8). Other patterns of LINE-1 methylation showed no significant differences (Figure 41 and Table 8).

Table 8 Demographic characteristic of subjects and percentage of LINE-1 products in matched cases.

	Matched cases					
	Non-smokers			Current smokers		
	Male	Female	Total	Male	Female	Total
Number of subjects	14	15	29	14	15	29
Age (year ± SD)	44.29 ± 17.66	46.47 ± 3.34	45.41 ± 5.33	44.07 ± 6.25	45.07 ± 12.39	44.59 ± 4.13
$\%\ ^{m}C$ (mean ± SD)	42.55 ± 1.82	42.02 ± 2.59	42.28 ± 2.23	43.44 ± 2.79	41.29 ± 3.14	42.32 ± 3.12
$\%\ ^{m}C^{m}C$ (mean ± SD)	16.94 ± 4.03	17.78 ± 5.07	17.38 ± 4.54	18.30 ± 4.69	15.81 ± 5.54	17.49 ± 5.35
$\%^{u}C^{u}C$ (mean ± SD)	31.84 ± 2.32	33.74 ± 4.03	32.82 ± 3.40	32.43 ± 4.88	33.23 ± 3.84	32.84 ± 4.31
$\%\ ^{m}C^{u}C$ (mean ± SD)	26.73 ± 2.13	27.19 ± 3.59	26.97 ± 2.93	22.59 ± 3.55	25.09 ± 3.66	23.89 ± 3.77
$\%\ ^{u}C^{m}C$ (mean ± SD)	24.48 ± 4.50	21.29 ± 6.78	22.83 ± 5.92	25.68 ± 9.18	25.86 ± 7.04	25.78 ± 7.99
$\%\ ^{m}C^{u}C+^{u}C^{m}C$ (mean ± SD)	51.22 ± 5.48	48.49 ± 7.55	49.80 ± 6.66	48.28 ± 7.77	50.96 ± 7.16	49.66 ± 7.45

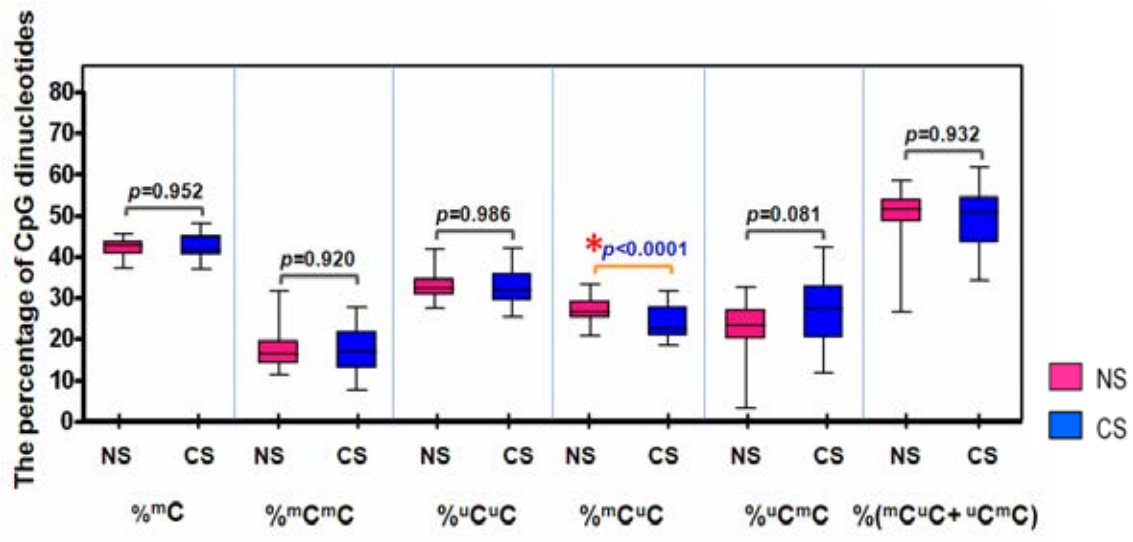


Figure 41 Percentage of each methylated CpG pattern in the matched cases.

By comparing the methylation patterns of the NS to the patterns of the CS. The NS were matched to the CS based on age, gender, smoking and alcohol drinking behaviour. The $\%^{m}C^{u}C$ provided significance, while the other patterns were not.

<u>Part V: Additional studies</u>

<u>The Percentage of Loci of each LINE-1 Methylation Pattern in NS and FS</u>

An additional investigation was performed on 17 FS, 15 males and 2 females, who had quit smoking for no less than 1 year and who had no mucosal lesions. Demographic characteristics were shown in Table 6. We found that the FS had a lower level of $\%^{m}C^{u}C$ than the NS; this difference was significant with p=0.001 (Figure 42 and Figure 44B). Other patterns of LINE-1 methylation did not reveal any significant differences (Figure 42).

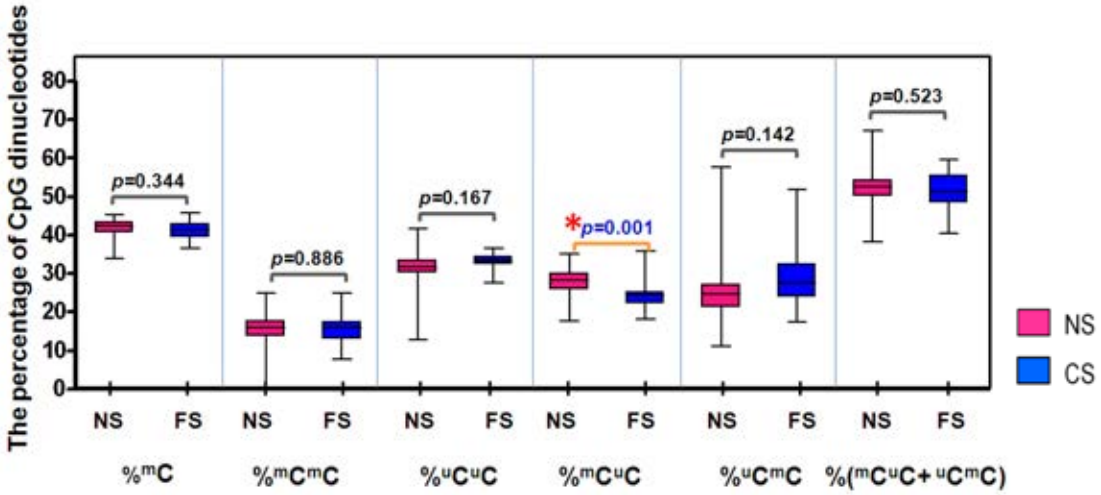


Figure 42 Percentage of each methylated CpG pattern between NS and FS.

A comparison of the percentages of the LINE-1-methylation patterns in NS and FS. Only the % $^{m}C^{u}C$ of NS significantly differed from that of FS.

<u>Part VI: The Percentage of Loci of each LINE-1 Methylation Pattern in pack year groups</u>

Based on the intensity of smoking, all of the smoking subjects were categorised into 2 groups based on the average pack-year (group I≤13.23 and group II>13.23 pack-years). The % $^{m}C^{u}C$ pattern was observed to be significantly different between the groups; % $^{m}C^{u}C$ was significantly lower in group II, p=0.028 (Figure 43, Figure 44C and Table 9), while the other pattern were not (Figure 43 and Table 9).

Table 9 Demographic characteristic of subjects and percentage of LINE-1 products in and pack-year smoking groups.

	Pack-year smoking	
	≤13.23 (group I)	>13.23 (group II)
Number of subjects Toatal= 96 (Current smokes)	54	42
Gender Male Female	 44 10	 36 6
Age (year ± SD)	33.34 ± 10.26	52.31 ± 10.63
% ^{m}C (mean ± SD)	42.32 ± 2.83	42.09 ± 2.44
% $^{m}C^{m}C$ (mean ± SD)	17.69 ± 5.10	18.39 ± 4.65
% $^{u}C^{u}C$ (mean ± SD)	33.05 ± 4.00	34.22 ± 3.44
% $^{m}C^{u}C$ (mean ± SD)	24.30 ± 3.43	22.71 ± 3.05
% $^{u}C^{m}C$ (mean ± SD)	24.97 ± 8.42	24.69 ± 7.36
% $^{m}C^{u}C+^{u}C^{m}C$ (mean ± SD)	49.27 ± 7.22	47.39 ± 6.57

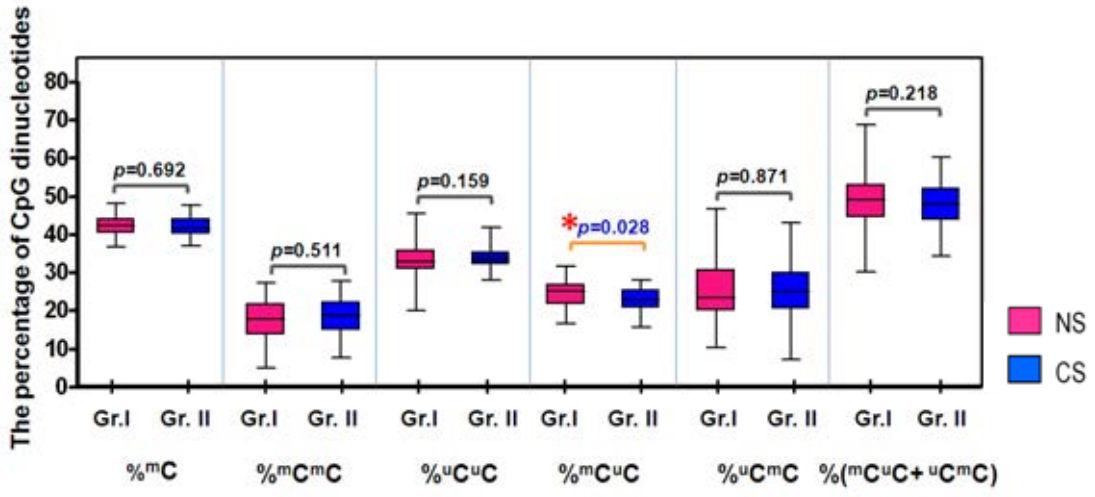


Figure 43 Percentage of LINE-1 products in pack-year smoking groups.

The LINE-1 methylation patterns were compared between the pack-year smoking groups. Only the $^{m}C^{u}C$ % in the higher pack-year smoking (group II) was significantly lower than that from the lower pack-year smoking (group I).

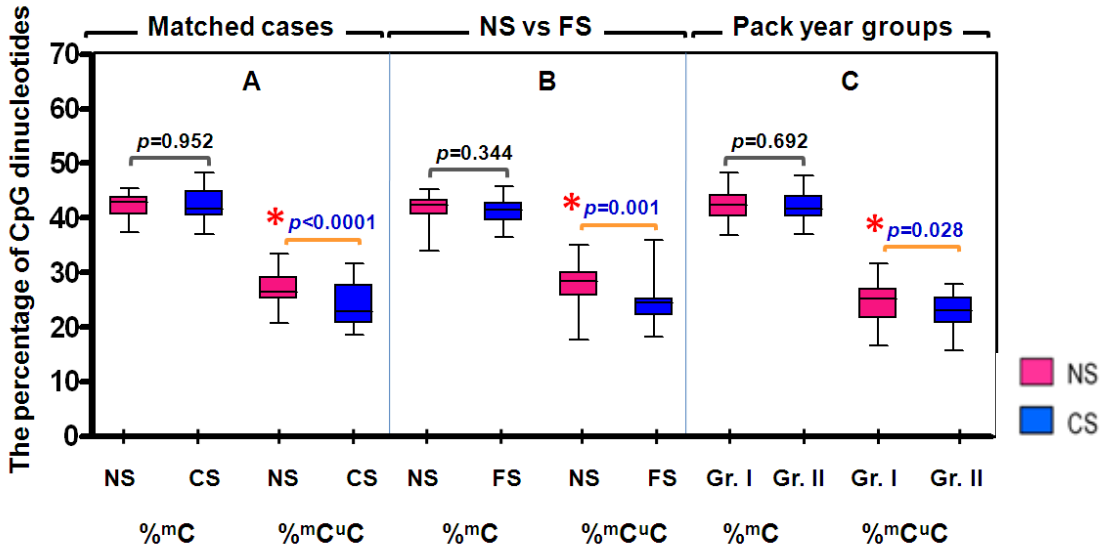


Figure 44 Percentage of $^{m}C^{u}C$ in the matched cases, FS and pack-year groups.

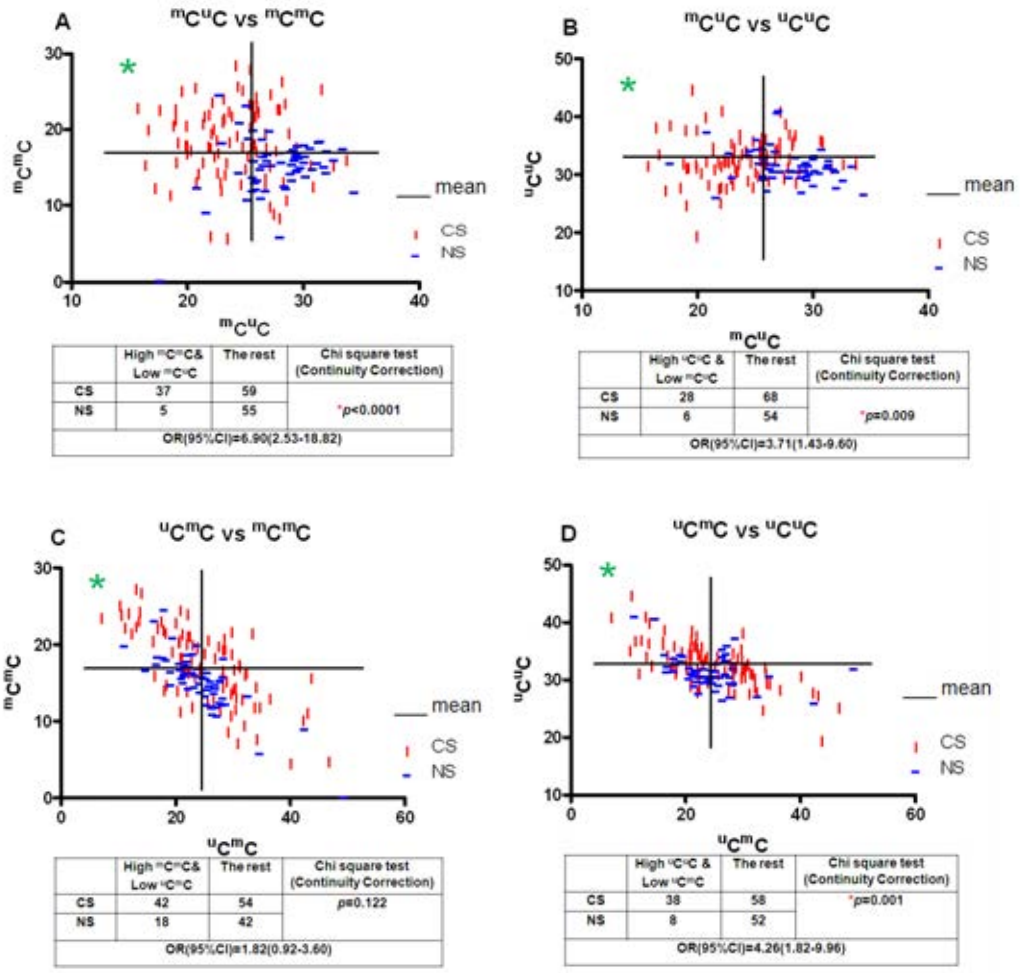
$^{m}C^{u}C$ represented $\%^{m}C^{u}C$. The LINE-1 methylation level was depicted as ^{m}C. (A) To reduce the effect of other confounding factors, we matched the NS to the CS based on age, gender and alcohol drinking. The CS showed a significantly lower $\%^{m}C^{u}C$. (B) The FS also had a significantly lower $\%^{m}C^{u}C$. (C) $\%^{m}C^{u}C$ was significantly lower in the higher pack-year smoking (group II) than in lower pack-year group (group I). While the alteration of the overall methylation level in these 3 measurements were not found.

Part VII: Additional studies

The Pattern of LINE-1 Methylation is Interchangeable

Encouraged by the information that the change in $^{m}C^{u}C$ was opposite that of $^{m}C^{m}C$ and $^{u}C^{u}C$, we further analysed the possibility of methylation switching between these forms. The number of NS and CS who had a lower $\%^{m}C^{u}C$ and a higher $\%^{m}C^{m}C$ than

the group means were counted and compared with the remainder of the group using the chi-squared test. A two-way table was created to calculate the continuity correction and the OR. We found that the OR was 6.90 and the 95% confidential interval (CI) was 2.53-18.82 with $p<0.0001$ (Figure 45A). We performed the same test for the low $\%^{m}C^{u}C$ and the high $\%^{u}C^{u}C$ set and the OR and 95% CI were 3.71 and 1.43-9.60, respectively with $p=0.009$ (Figure 45B). These results implied that a reduction of $^{m}C^{u}C$ in the CS was associated with an increase in $^{m}C^{m}C$ or $^{u}C^{u}C$. However, to clarify the possibility that $^{u}C^{m}C$ is converted to $^{m}C^{m}C$ or $^{u}C^{u}C$, the same analysis was performed in the group with low $\%^{u}C^{m}C$ and either high $\%^{m}C^{m}C$ or $\%^{u}C^{u}C$. We found that while $^{u}C^{m}C$ could not change to $^{m}C^{m}C$ (OR =1.82, 95% CI=0.92-3.60, $p=0.122$) (Figure 45C), it could be converted to the $^{u}C^{u}C$ form (OR=4.26, 95% CI=1.82-9.96, $p=0.001$) (Figure 45D).



	High mCmC& Low mCmC	The rest	Chi square test (Continuity Correction)
CS	37	59	
NS	5	55	*p<0.0001
OR(95%CI)=6.90(2.53-18.82)			

	High uCuC & Low mCmC	The rest	Chi square test (Continuity Correction)
CS	28	68	
NS	6	54	*p=0.009
OR(95%CI)=3.71(1.43-9.60)			

	High mCmC& Low uCmC	The rest	Chi square test (Continuity Correction)
CS	42	54	p=0.122
NS	18	42	
OR(95%CI)=1.82(0.92-3.60)			

	High uCuC & Low uCmC	The rest	Chi square test (Continuity Correction)
CS	38	58	*p=0.001
NS	8	52	
OR(95%CI)=4.26(1.82-9.96)			

Figure 45 The interchangeable methylation patterns of the LINE-1s.

$^{m}C^{u}C$, $^{u}C^{m}C$, $^{m}C^{m}C$ and $^{u}C^{u}C$, represented %$^{m}C^{u}C$, $^{u}C^{m}C$, $^{m}C^{m}C$ and $^{u}C^{u}C$, respectively. The graphs were plotted for the percentages of either $^{m}C^{u}C$ or $^{u}C^{m}C$ on the X-axis and either $^{m}C^{m}C$ or $^{u}C^{u}C$ on the Y-axis. The vertical and horizontal lines indicate the mean percentages of each axis. The graph is divided into 4 quadrants. The numbers of NS and CS who fell in the upper left quadrant were counted and compared to the remainder of the group using the chi-squared test. The results are shown in the tables below the graphs. (A) The numbers of NS and CS who had high $^{m}C^{m}C$ and low $^{m}C^{u}C$ were analysed with the remainder of the group. The correlation of these two patterns showed an OR=6.90 (significantly at $p<0.0001$). (B) A significant correlation of high $^{u}C^{u}C$ and low $^{m}C^{u}C$ was also discovered at OR=3.71, $p=0.009$. (C) There was no correlation between high $^{m}C^{m}C$ and low $^{u}C^{m}C$ (OR=1.82, $p=0.122$). D) However, high $^{u}C^{u}C$ was significantly correlated with low $^{u}C^{m}C$ (OR=4.26, $p=0.001$).

Part VIII: Additional studies

The Influence of Smoking on the CU-DREAM-X Analysis in the LINE-1s

Previously, we compared the expression of genes with intragenic LINE-1s in cancer and demethylated cells. We found that in most cancers, genes with intragenic LINE-1s were down-regulated (OR>1) or prevented from up-regulation (OR<1) *[36]*. In this study, we tested 15 microarray expression experiments from the airway epithelia of smokers. Eleven experiments demonstrated evidence of LINE-1 regulation ($p<0.05$) (Table 10). Five experiments showed down-regulation of gene expression similar to cancer (OR>1 for the down-regulated genes and OR<1 for the up-regulated group). Two experiments showed gene up-regulation only (OR<1 for the down-regulated group and OR>1 for the up-regulated group). Interestingly, four cases showed ORs>1 for both the down- and up-regulated genes, implying the consequences of hypomethylated and hypermethylated intragenic LINE-1s. This result confirmed the epigenetic changes of LINE-1s and, consequently, controlled for the smoking-induced changes in gene expression.

Table 10 Connection Up- or Down-Regulation Expression Analysis of Microarrays Extension program (CU-DREAM-X) for the LINE-1s.

Experiment	Cell	Down		UP	
		OR	*p*-value	OR	*p*-value
GSE4302-2	Airway epithelial brushing	1.9	6.48E-24	0.8	8.99E-04
GSE19667-1	Small airways	1.68	5.53E-16		2.27E-01
GSE19667-2	Small airways	1.74	2.97E-14	0.74	1.82E-02
GSE11906-8	Trachea	1.63	1.05E-11		1.56E-01
GSE11906-5	Small airways	2.37	3.93E-06	1.32	1.79E-0
GSE4498	Small airways	0.74	9.96E-04	1.54	4.85E-05
GSE13933-2	Trachea	1.24	1.00E-03	1.52	2.58E-09
GSE11906-7	Trachea	1.26	5.50E-03	1.21	2.94E-02
GSE3320	Small airway	1.51	6.11E-03		2.15E-01
GSE8545-2	Small airway	0.68	3.84E-02	1.33	3.89E-06
GSE13933-1	Small airway	1.19	4.47E-02	1.63	2.60E-11
GSE11906-6	Small airway		9.70E-02		7.87E-01
GSE27002	Small airway		4.10E-01		9.90E-02
GSE7895-1	Bronchial Epithelium		6.28E-01		5.23E-01
GSE994-1	Bronchial Epithelium		7.37E-01		8.12E-01

CHAPTER V

DISCUSSION

Similar to blood cells and colonic epithelium *[30-32],* cigarette smoke does not change LINE-1 methylation level of oral mucosa. However, there are alterations in patterns of LINE-1 methylation that we found both $^{u}C^{u}C$ and $^{m}C^{m}C$ loci increased. The unchanged methylation level can be explained by the fact that the LINE-1 methylation level is a sum of the methylation from all the LINE-1s. Therefore, the increases in both the $^{u}C^{u}C$ and $^{m}C^{m}C$ loci counterbalance each other, neutralising their effect on the LINE-1 methylation levels. This evidence supported by the study of Kitkumthorn and Mutirangura, which revealed that the LINE-1 methylation level measurement was not sufficiently sensitive or accurate to determine the LINE-1 methylation changes in pathological conditions *[20].* If possible, the re-evaluating LINE-1 methylation pattern of the previous studies reported unaltered overall methylation level may let some benefits.

Contrary to the reduction of genome-wide methylation levels caused by some chemical agents *[75, 76]*, smoking could paradoxically promote both an increase and decrease in methylation in certain LINE-1s. Interestingly, smoking-induced hypomethylation originated from both forms of partially methylated LINE-1, $^{m}C^{u}C$ and $^{u}C^{m}C$. While the hypermethylated LINE-1s derived from only one form, $^{m}C^{u}C$. These observations suggest that the mechanisms that increase or decrease methylation are different. The hypomethylation mechanism seems to be a generalised process that affects many LINE-1s regardless of the original methylation patterns and is similar to the global hypomethylation found in cancer *[12, 20].* Accordingly, cancer and smoking may reduce genome-wide methylation by the same mechanism.

Even though the methylation differences between smokers and non-smokers are just a few percentage points difference, the alteration should be significant. Global hypomethylation can cause cancer by promoting genomic instability and by altering

gene expression in *cis [20].* There are evidences suggesting that DNA methylation maintains genomic integrity in *cis.* First, a close correlation between the site of the chromosome translocation and the loss of the methylation of satellite DNA has been reported *[109, 110]*. Recently, the study revealed that the repair of the replication of independent DNA double-strand breaks occurring within hypomethylated regions was more error prone *[35].* For gene expression, the repression of mRNA production by hypomethylated intragenic LINE-1swas reported *[36].* This study also found that epigenetic changes of intragenic LINE-1s by smoking regulated gene expression. Therefore, the increasing number of hypomethylated LINE-1s induced by smoking should promote cancer at certain loci in *cis* (Figure 46).

Previously, the lower methylation of intragenic LINE-1s in cancer enhanced LINE-1 transcription and repressed the host genes by double-stranded RNA and the AGO2 complex were reported *[36].* Interestingly, this study revealed three categories of smoking-induced regulation of genes containing LINE-1. This implied an increase in LINE-1 methylation in cases of gene up-regulation and LINE-1 hypomethylation when genes were repressed. However, the lack of changes in gene expression may be due to a lower dosage of smoke exposure.

This study revealed the deterioration of smoking occurring before detectable change of oral mucosa. Additionally, this harm is the-dose response relationship. Thus, it is crucial for educating people about smoking hazard. Also, the encouragement and counseling in abstinence from smoking provides some global benefits including reduction of cancer incidence and expense.

Interestingly, second-hand smoke is also known as environmental tobacco smoke and it is classified as a known human carcinogen by World Health Organization International Agency for Research and Cancer *[111]*. In addition, tobacco smoking are not only affected to the smoker but the stream of carcinogenic matter could have negative influence to those found to the second-hand smoke exposure *[112]*. Prenatal tobacco smoke exposure was associated with detectable changes in global DNA

methylation levels which obtained the DNA samples from buccal scrape of children. However, exposed children had a significantly lower level of methylation for AluYb8 but not for LINE-1 *[113]*. However, additional studies are required to get better information for supporting about the disadvantage of second-hand smokes to cellular response by investigating DNA methylation patterns.

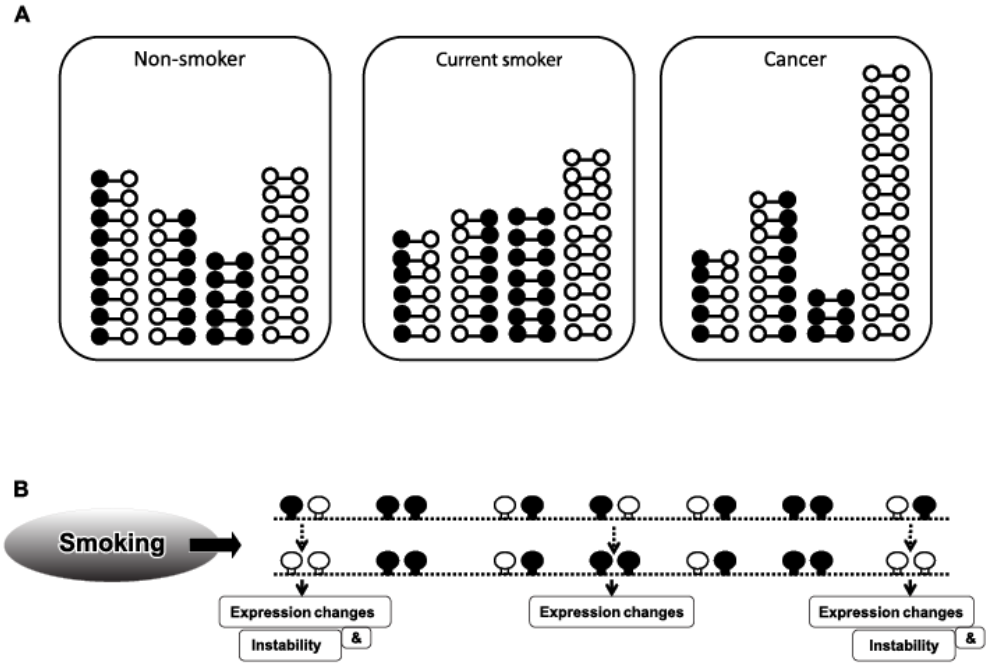


Figure 46 Influence of smoking on the epigenetic progression of multistep carcinogenesis.

(A) Models of LINE-1 methylation patterns in oral mucosal cells of a NS, the oral mucosal cells of a CS and in cancer cells (HeLa) are shown. Although the overall methylation level did not change in the CS, some alterations in the methylation patterns were detected. While the numbers of $^{m}C^{m}C$ and $^{u}C^{u}C$ were increased, only one form of partial methylation, $^{m}C^{u}C$, was decreased. Moreover, the addition of $^{m}C^{m}C$ and $^{u}C^{u}C$ correlated with the depletion of $^{m}C^{u}C$. In contrast, a reduction in the overall methylation level was found in cancer cells. The numbers of $^{m}C^{m}C$ and $^{m}C^{u}C$ were significantly decreased, while the numbers of $^{u}C^{u}C$ were significantly increased. The numbers of $^{u}C^{m}C$ were slightly increased. (B) The

smoking-induced hypomethylated loci could be derived from both classes of the partial methylation patterns and could result in genome instability and gene expression changes. However, the smoking-induced hypermethylated loci were from $^{m}C^{u}C$ only and, consequently, effected gene expression.

CHAPTER VI

CONCLUSION

In conclusion, smoking paradoxically alters LINE-1 methylation by increasing or decreasing methylation of certain loci. The mechanisms causing LINE-1 hypomethylation and hypermethylation are different. Nevertheless, the biological consequences of LINE-1 hypomethylation in smoking and cancer are similar. In addition, the dose-response relationship between the intensity of smoking and methylation change was found. Further exploration of methylation pattern changes of other intersperse repetitive sequences and gene promoters whether they are related to other smoking-associated malignancies, as well as other carcinogens, is necessary. Future studies should focus for a long term follow up on progression from normal to potentially malignant disorders of oral mucosa in smokers and other environmental risk factors. Moreover, the cumulative effect of lifetime smoking exposure on oral epithelia is crucial for better understanding mechanisms of hypomethylation in smokers. Finally, a better understanding of the causes and mechanisms of genome-wide methylation changes will be crucial for cancer prevention.

.

References

[1]. Olshan, A.F.; Weissler, M.C.; Watson, M.A.; and Bell, D.A. GSTM1, GSTT1, GSTP1, CYP1A1, and NAT1 polymorphisms, tobacco use, and the risk of head and neck cancer. Cancer Epidemiol Biomarkers Prev. 9 (2000): 185-191.

[2]. DeMarini, D.M. Genotoxicity of tobacco smoke and tobacco smoke condensate: a review. Mutat Res. 567 (2004): 447-474.

[3]. Schlecht, N.F.; Franco, E.L.; Pintos, J.; and Kowalski, L.P. Effect of smoking cessation and tobacco type on the risk of cancers of the upper aero-digestive tract in Brazil. Epidemiology. 10 (1999): 412-418.

[4]. Spitz, M.R.; Wei, Q.; Li, G.; and Wu, X. Genetic susceptibility to tobacco carcinogenesis. Cancer Invest. 17 (1999): 645-659.

[5]. Macfarlane, G.J., et al. Alcohol, tobacco, diet and the risk of oral cancer: a pooled analysis of three case-control studies. Eur J Cancer B Oral Oncol. 31B (1995): 181-187.

[6]. Spitz, M.R.; Fueger, J.J.; Goepfert, H.; Hong, W.K.; and Newell, G.R. Squamous cell carcinoma of the upper aerodigestive tract. A case comparison analysis. Cancer. 61 (1988): 203-208.

[7]. Banoczy, J. Observations sur l'alteration de la cornification de la muqueuse buccale sous l'influence du tabac. Bull Group Int Rech Sci Stomatol. 5 (1962): 543-553.

[8]. Meyer, J.; Rubinstein, A.S.; and Medak, H. Early effects of smoking on surface cytology of the oral mucosa: II. Cell changes in smokers. Oral Surgery, Oral Medicine, Oral Pathology. 30 (1970): 700-710.

[9]. Martin, G.C.; Brown, J.P.; Eifler, C.W.; and Houston, G.D. Oral leukoplakia status six weeks after cessation of smokeless tobacco use. J Am Dent Assoc. 130 (1999): 945-954.

[10]. Shibly, O.; Cummings, K.M.; and Zambon, J.J. Resolution of oral lesions after tobacco cessation. J Periodontol. 79 (2008): 1797-1801.

[11]. Greenblatt, M.S.; Bennett, W.P.; Hollstein, M.; and Harris, C.C. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res. 54 (1994): 4855-4878.

[12]. Chalitchagorn, K., et al. Distinctive pattern of LINE-1 methylation level in normal tissues and the association with carcinogenesis. Oncogene. 23 (2004): 8841-8846.

[13]. Dammann, R., et al. CpG island methylation and expression of tumour-associated genes in lung carcinoma. *Eur J Cancer.* 41 (2005): 1223-1236.

[14]. Toyooka, S., et al. Mutational and epigenetic evidence for independent pathways for lung adenocarcinomas arising in smokers and never smokers. Cancer Res. 66 (2006): 1371-1375.

[15]. Salskov, A., et al. Hypermethylation of CCND2 May Reflect a Smoking-Induced Precancerous Change in the Lung. J Oncol. (2011).

[16]. Smith, I.M.; Mydlarz, W.K.; Mithani, S.K.; and Califano, J.A. DNA global hypomethylation in squamous cell head and neck cancer associated with smoking, alcohol consumption and stage. Int J Cancer. 121 (2007): 1724-1728.

[17]. Graham, T.; and Boissinot, S. The genomic distribution of L1 elements: the role of insertion bias and natural selection. J Biomed Biotechnol. 2006 (2006): 75327.

[18]. Lander, E.S., et al. Initial sequencing and analysis of the human genome. Nature. 409 (2001): 860-921.

[19]. Penzkofer, T.; Dandekar, T.; and Zemojtel, T. L1Base: from functional annotation to prediction of active LINE-1 elements. Nucleic Acids Res. 33 (2005): D498-500.

[20]. Kitkumthorn, N.; and Mutirangura, A. Long interspersed nuclear element-1 hypomethylation in cancer: biology and clinical applications. Clin Epigenet. 2 (2011): 315-330.

[21]. Shuangshoti, S.; Hourpai, N.; Pumsuk, U.; and Mutirangura, A. Line-1 hypomethylation in multistage carcinogenesis of the uterine cervix. Asian Pac J Cancer Prev. 8 (2007): 307-309.

[22]. Tangkijvanich, P., et al. Serum LINE-1 hypomethylation as a potential prognostic marker for hepatocellular carcinoma. Clin Chim Acta. 379 (2007): 127-133.

[23]. Iacopetta, B., et al. Methylation levels of LINE-1 repeats and CpG island loci are inversely related in normal colonic mucosa. Cancer Sci. 98 (2007): 1454-1460.

[24]. Iramaneerat, K.; Rattanatunyong, P.; Khemapech, N.; Triratanachat, S.; and Mutirangura, A. HERV-K hypomethylation in ovarian clear cell carcinoma is associated with a poor prognosis and platinum resistance. Int J Gynecol Cancer. 21 (2011): 51-57.

[25]. Florl, A.R.; Lower, R.; Schmitz-Drager, B.J.; and Schulz, W.A. DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. Br J Cancer. 80 (1999): 1312-1321.

[26]. Hsiung, D.T., et al. Global DNA methylation level in whole blood as a biomarker in head and neck squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev. 16 (2007): 108-114.

[27]. Feber, A., et al. Comparative methylome analysis of benign and malignant peripheral nerve sheath tumors. Genome Res. 21 (2011): 515-524.

[28]. Romermann, D., et al. Methylation status of LINE-1 sequences in patients with MDS or secondary AML. Verh Dtsch Ges Pathol. 91 (2007): 338-342.

[29]. Perrin, D., et al. Specific hypermethylation of LINE-1 elements during abnormal overgrowth and differentiation of human placenta. Oncogene. 26 (2007): 2518-2524.

[30]. Figueiredo, J.C., et al. Global DNA hypomethylation (LINE-1) in the normal colon and lifestyle characteristics and dietary and genetic factors. Cancer Epidemiol Biomarkers Prev. 18 (2009): 1041-1049.

[31]. Hillemacher, T., et al. Global DNA methylation is influenced by smoking behaviour. Eur Neuropsychopharmacol. 18 (2008): 295-298.

[32]. Madrigano, J., et al. Prolonged exposure to particulate pollution, genes associated with glutathione pathways, and DNA methylation in a cohort of older men. Environ Health Perspect. 119 (2011): 977-982.

[33]. Liu, F., et al. Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. Oncogene. 29 (2010): 3650-3664.

[34]. Kongruttanachok, N., et al. Replication independent DNA double-strand break retention may prevent genomic instability. Mol Cancer. 9 (2010): doi: 10.1186/1476-4598-1189-1170.

[35]. Pornthanakasem, W., et al. LINE-1 methylation status of endogenous DNA double-strand breaks. Nucleic Acids Res. 36 (2008): 3667-3675.

[36]. Aporntewan, C., et al. Hypomethylation of intragenic LINE-1 represses transcription in cancer cells through AGO2. PLoS One. 6 (2011): e17934.

[37]. Phokaew, C.; Kowudtitham, S.; Subbalekha, K.; Shuangshoti, S.; and Mutirangura, A. LINE-1 methylation patterns of different loci in normal and cancerous cells. Nucleic Acids Res. 36 (2008): 5704-5712.

[38]. Ho, S.M.; and Tang, W.Y. Techniques used in studies of epigenome dysregulation due to aberrant DNA methylation: an emphasis on fetal-based adult diseases. Reprod Toxicol. 23 (2007): 267-282.

[39]. Baba, Y., et al. Epigenomic diversity of colorectal cancer indicated by LINE-1 methylation in a database of 869 tumors. Mol Cancer. 9 (2010).

[40]. Pobsook, T.; Subbalekha, K.; Sannikorn, P.; and Mutirangura, A. Improved measurement of LINE-1 sequence methylation for cancer detection. Clin Chim Acta. 412 (2010): 314-321.

[41]. Kitkumthorn, N.; Tuangsintanakul, T.; Rattanatanyong, P.; Tiwawech, D.; and Mutirangura, A. LINE-1 methylation in the peripheral blood mononuclear cells of cancer patients. Clin Chim Acta. 413 (2012): 869-874.

[42]. Sham, A.S.; Cheung, L.K.; Jin, L.J.; and Corbet, E.F. The effects of tobacco use on oral health. Hong Kong Med J. 9 (2003): 271-277.

[43]. Wald, N.J.; and Hackshaw, A.K. Cigarette smoking: an epidemiological overview. Br Med Bull. 52 (1996): 3-11.

[44]. Warnakulasuriya, S.; Sutherland, G.; and Scully, C. Tobacco, oral cancer, and treatment of dependence. Oral Oncol. 41 (2005): 244-260.

[45]. Greenberg, M.S.; Glick, M.; and Ship, J.A., Burket's Oral Medicine. 11th ed., Hamilton: BC Decker Inc. 2008.

[46]. Banoczy, J.; Gintner, Z.; and Dombi, C. Tobacco use and oral leukoplakia. J Dent Educ. 65 (2001): 322-327.

[47]. van Oijen, M.G.; Gilsing, M.M.; Rijksen, G.; Hordijk, G.J.; and Slootweg, P.J. Increased number of proliferating cells in oral epithelium from smokers and ex-smokers. Oral Oncol. 34 (1998): 297-303.

[48]. Soames, J.V.; and Southam, J.C., Oral Pathlogy. 4th ed., Oxford; Newyork: Oxford University Press. 2005.

[49]. Gupta, P.C.; Murti, P.R.; Bhonsle, R.B.; Mehta, F.S.; and Pindborg, J.J. Effect of cessation of tobacco use on the incidence of oral mucosal lesions in a 10-yr follow-up study of 12,212 users. Oral Dis. 1 (1995): 54-58.

[50]. Neville, B.W.; and Terry, D.A. Oral cancer and precancerous lesions. CA cancer J Clin. 52 (2002): 195-215.

[51]. Cawson, R.A.; and Odell, E.W., Cawson's essentials of oral pathology and oral medicine. 8th ed.

. Edinburgh London: Churchill Livingstone. 2008.

[52]. Welbourne, J.P., Premalignant and malignant epithelial tumors of mucosa and skin In R.E. Marx and D. Stern, Oral and maxillofacial pathology: a rationale for diagnosis and treatment, pp. 283-373. Illinois: Quintessence Inc., 2003.

[53]. Tsantoulis, P.K.; Kastrinakis, N.G.; Tourvas, A.D.; Laskaris, G.; and Gorgoulis, V.G. Advances in the biology of oral cancer. Oral Oncol. 43 (2007): 523-534.

[54]. Moreno-Lopez, L.A., et al. Risk of oral cancer associated with tobacco smoking, alcohol consumption and oral hygiene: a case-control study in Madrid, Spain. Oral Oncol. 36 (2000): 170-174.

[55]. Warnakulasuriya, S. Global epidemiology of oral and oropharyngeal cancer. Oral Oncol. 45 (2009): 309-316.

[56]. Sapp, J.P.; Eversole, L.R.; and Wysocki, G.P., Epithelial disorders. In J.P. Sapp; L.R. Eversole; G.P. Wysocki and editors, Contemporary Oral and Maxillofacial Pathology, pp. 164-206. St. Louis: Mosby, 2004.

[57]. Neville, B.W.; Damm, D.D.; Allen, C.M.; and Bouquot, J.E., Oral and maxillofacial pathology. 3rd ed., St. Louis: Saunders Elsevir. 2009.

[58]. Johnson, N. Tobacco use and oral cancer: a global perspective. J Dent Educ. 65 (2001): 328-339.

[59]. Winn, D.M. Tobacco use and oral disease. J Dent Educ. 65 (2001): 306-312.

[60]. Marx, R.E.; and Stern, D., Oral and maxiloofacial pathology: a rationale for diagnosis and treatment. Chicago: Quintessence Publishing. 2003.

[61]. International Agency for Reseach on Cancer. Tobacco Smoke and Involuntary Smoking. In IARC Monograph on the Evaluation of the Carcinogenic risks to humans. Lyon, France, 2004.

[62]. Supic, G.; Kozomara, R.; Brankovic-Magic, M.; Jovic, N.; and Magic, Z. Gene hypermethylation in tumor tissue of advanced oral squamous cell carcinoma patients. Oral Oncol. 45 (2009): 1051-1057.

[63]. Zhu, J.; and Yao, X. Use of DNA methylation for cancer detection: promises and challenges. Int J Biochem Cell Biol. 41 (2009): 147-154.

[64]. Ballestar, E.; and Esteller, M. Advances in genetics. Chapter 9 Epigenetic gene regulation in cancer. 61 (2008): 247-267.

[65]. Brena, R.M.; and Costello, J.F. Genome-epigenome interactions in cancer. Hum Mol Genet. 16 Spec No 1 (2007): R96-105.

[66]. Shaw, R. The epigenetics of oral cancer. Int J Oral Maxillofac Surg. 35 (2006): 101-108.

[67]. Mehrotra, R.; Gupta, A.; Singh, M.; and Ibrahim, R. Application of cytology and molecular biology in diagnosing premalignant or malignant oral lesions. Mol Cancer. 5 (2006): doi: 10.1186/1476-4598-1185-1111.

[68]. Matouk, C.C.; and Marsden, P.A. Epigenetic regulation of vascular endothelial gene expression. Circ Res. 102 (2008): 873-887.

[69]. Mund, C., et al. Array-based analysis of genomic DNA methylation patterns of the tumour suppressor gene p16INK4A promoter in colon carcinoma cell lines. Nucleic Acids Res. 33 (2005): doi: 10.1093/nar/gni1072.

[70]. Das, P.M.; and Singal, R. DNA methylation and cancer. J Clin Oncol. 22 (2004): 4632-4642.

[71]. Herceg, Z.; and Hainaut, P. Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. Mol Oncol. 1 (2007): 26-41.

[72]. Altun, G.; Laurent, L.C.; and Loring, J.F. Epigenetic remodeling and stem cells. Drug Discovery Today: Technologies. 5 (2008): e139-e142.

[73]. Sakurada, K. Environmental epigenetic modifications and reprogramming-recalcitrant genes. Stem Cell Res. 4 (2010): 157-164.

[74]. Guerrero-Preston, R., et al. Global DNA methylation: a common early event in oral cancer cases with exposure to environmental carcinogens or viral agents. P R Health Sci J. 28 (2009): 24-29.

[75]. Teneng, I.; Montoya-Durango, D.E.; Quertermous, J.L.; Lacy, M.E.; and Ramos, K.S. Reactivation of L1 retrotransposon by benzo(a)pyrene involves complex genetic and epigenetic regulation. Epigenetics. 6 (2011): 355-367.

[76]. Bollati, V., et al. Changes in DNA methylation patterns in subjects exposed to low-dose benzene. Cancer Res. 67 (2007): 876-880.

[77]. Oakeley, E.J. DNA methylation analysis: a review of current methodologies. Pharmacol Ther. 84 (1999): 389-400.

[78]. Glazer, C.A.; Chang, S.S.; Ha, P.K.; and Califano, J.A. Applying the molecular biology and epigenetics of head and neck cancer in everyday clinical practice. Oral Oncol. 45 (2009): 440-446.

[79]. Gopisetty, G.; Ramachandran, K.; and Singal, R. DNA methylation and apoptosis. Mol Immunol. 43 (2006): 1729-1740.

[80]. Yang, W.H.; Wu, J.R.; Wang, T.H.; and Au, L.C. Methylation profiling using degenerated oligonucleotide primer-PCR specific for genome-wide amplification of bisulfite-modified DNA. Anal Biochem. 369 (2007): 120-127.

[81]. Robertson Laboratory. Department of Biochemistry and molecular Biology. DNA methylation in mamalian cell [Online]. Available from : http:/www.med.ufl.edu/biochem/keithr/fig1pt1.html.

[82]. Wilson, A.S.; Power, B.E.; and Molloy, P.L. DNA hypomethylation and human diseases. Biochim Biophys Acta. 1775 (2007): 138-162.

[83]. Singal, R.; and Ginder, G.D. DNA methylation. Blood. 93 (1999): 4059-4070.

[84]. Qureshi, S.A.; Bashir, M.U.; and Yaqinuddin, A. Utility of DNA methylation markers for diagnosing cancer. Int J Surg. 8 (2010): 194-198.

[85]. Prokhortchouk, E.; and Defossez, P.A. The cell biology of DNA methylation in mammals. Biochim Biophys Acta. 1783 (2008): 2167-2173.

[86]. Newell-Price, J.; Clark, A.J.; and King, P. DNA methylation and silencing of gene expression. Trends Endocrinol Metab. 11 (2000): 142-148.

[87]. Molecular Biology Web Book. DNA methylation and CpG Island [Online]. Available from : http://www.web-book.com/MoBio/Free/ch7F2.htm.

[88]. Schilling, E.; and Rehli, M. Global, comparative analysis of tissue-specific promoter CpG methylation. Genomics. 90 (2007): 314-323.

[89]. Chhibber, A.; and Schroeder, B.G. Single-molecule polymerase chain reaction reduces bias: application to DNA methylation analysis by bisulfite sequencing. Anal Biochem. 377 (2008): 46-54.

[90]. Gopalakrishnan, S.; Van Emburgh, B.O.; and Robertson, K.D. DNA methylation in development and human disease. Mutat Res. 647 (2008): 30-38.

[91]. Jones, P.A. Overview of cancer epigenetics. Semin Hematol. 42 (2005): S3-8.

[92]. Laird, P.W. The power and the promise of DNA methylation markers. Nat Rev Cancer. 3 (2003): 253-266.

[93]. Cottrell, S.E. Molecular diagnostic applications of DNA methylation technology. Clin Biochem. 37 (2004): 595-604.

[94]. Wessler, S.R., Eukaryotic transposable elements : teaching old genome new tricks

In L. Caporale, The Implicit on genome, pp. 138-165. New York: Oxford university, 2006.

[95]. The Contents of Human Genome [Online]. Available from : http://www.ncrna.org/stat genome.

[96]. Bannert, N.; and Kurth, R. Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A. 101 Suppl 2 (2004): 14572-14579.

[97]. Kazazian, H.H., Jr. Mobile elements: drivers of genome evolution. Science. 303 (2004): 1626-1632.

[98]. Sheen, F.M., et al. Reading between the LINEs: human genomic variation induced by LINE-1 retrotransposition. Genome Res. 10 (2000): 1496-1508.

[99]. Hong, S.J., et al. DNA methylation and expression patterns of key tissue-specific genes in adult stem cells and stomach tissues. J Korean Med Sci. 24 (2009): 918-929.

[100]. Flavell, A. Retrotransposons rule in Carry-le-Rouet. Trends Genet. 17 (2001): 489-490.

[101]. Szpakowski, S., et al. Loss of epigenetic silencing in tumors preferentially affects primate-specific retroelements. Gene. 448 (2009): 151-167.

[102]. Subbalekha, K., et al. Detection of LINE-1s hypomethylation in oral rinses of oral squamous cell carcinoma patients. Oral Oncol. 45 (2009): 184-191.

[103]. Bollati, V., et al. Decline in genomic DNA methylation through aging in a cohort of elderly subjects. Mech Ageing Dev. 130 (2009): 234-239.

[104]. Jintaridth, P.; and Mutirangura, A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. Physiol Genomics. 41 (2010): 194-200.

[105]. Toyooka, S.; and Shimizu, N. Models for studying DNA methylation in human cancer: a review of current status. Drug Discovery Today: Disease Models. 1 (2004): 37-42.

[106]. Clark, S.J.; Harrison, J.; Paul, C.L.; and Frommer, M. High sensitivity mapping of methylated cytosines. Nucleic Acids Res. 22 (1994): 2990-2997.

[107]. Marmiroli, N.; and Maestri, E., Polymerase chain reaction (PCR) In P. Yolanda, Food toxicants analysis: techniques, strategies and developments pp. 147-187. Amsterdam: Elsevier 2007.

[108]. Aporntewan, C.; and Mutirangura, A. Connection up- and down-regulation expression analysis of microarrays (CU-DREAM): a physiogenomic discovery tool. Asian Biomedicine. 5 (2011): 257-262.

[109]. Maraschio, P.; Zuffardi, O.; Dalla Fior, T.; and Tiepolo, L. Immunodeficiency, centromeric heterochromatin instability of chromosomes 1, 9, and 16, and facial anomalies: the ICF syndrome. J Med Genet. 25 (1988): 173-180.

[110]. Ji, W., et al. DNA demethylation and pericentromeric rearrangements of chromosome 1. Mutat Res. 379 (1997): 33-41.

[111]. Percival, J.; and Queally, B. Second-hand smoke: how damaging is it to health? Nurs Times. 101 (2005): 50-52.

[112]. Guingab-Cagmat, J., et al. Methods in tobacco abuse: proteomic changes following second-hand smoke exposure. Methods Mol Biol. 829 (2012): 329-348.

[113]. Breton, C.V., et al. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit Care Med. 180 (2009): 462-467.

APPENDICES

APPENDIX A

Informed Consent Sheet for All Participants (in Thai)

Consent Form for All Participants (in Thai)

Withdrawal Form in Case Drop-out is Demanded (in Thai)

ข้อมูลและรายละเอียดเกี่ยวกับการทำวิจัยที่ใช้ประกอบการพิจารณาเข้าร่วมโครงการ (Inform Consent)

การวิจัยนี้เป็นการนำวิธีการทางห้องปฏิบัติการมาศึกษาเมทิเลชันของไลน์-1 ในเซลล์เยื่อบุช่องปาก โดยใช้น้ำบ้วนปากที่จะมีส่วนประกอบของเยื่อบุช่องปาก ซึ่งเป็นเซลล์ที่จะได้รับการสัมผัสจากสารเคมีของบุหรี่โดยตรง ดังนั้นการศึกษานี้จำเป็นต้องใช้น้ำบ้วนปากจากมนุษย์เพื่อนำมาศึกษาค่าความแตกต่างของเมทิเลชันของไลน์-1 ระหว่างคนที่สูบบุหรี่กับไม่สูบบุหรี่

วัตถุประสงค์ของการวิจัย

ใช้วิธีการทางห้องปฏิบัติการเพื่อ

1. ศึกษาเมทิเลชันของไลน์-1 ในเซลล์เยื่อบุช่องปากในคนที่สูบบุหรี่กับไม่สูบบุหรี่
2. ศึกษาความสัมพันธ์ระหว่างเมทิเลชันของไลน์-1 ในเซลล์เยื่อบุช่องปากในคนที่สูบบุหรี่กับจำนวนบุหรี่ที่สูบ/ปี

วิธีการทดลอง

เก็บน้ำบ้วนปากอาสาสมัครด้วยวิธีการใช้น้ำเกลือที่สะอาดผ่านการทำให้ปราศจากเชื้อ อมกลั้วไว้ในช่องปาก จากนั้นให้บ้วนน้ำเกลือดังกล่าวผ่านกรวยลงในหลอดที่สะอาดผ่านการทำให้ปราศจากเชื้อ แล้วนำน้ำบ้วนปากไปสกัดสารพันธุกรรมและผ่านกระบวนการต่างๆ เพื่อวัดเมทิเลชันของไลน์-1 ในเชิงปริมาณ

ขั้นตอนการทดลองที่เกี่ยวข้องกับอาสาสมัคร

1. ให้อาสาสมัครทำแบบสอบถามด้วยตนเองโดยเก็บข้อมูลทั่วไป ประวัติโรคทางระบบประวัติทางการแพทย์ ข้อมูลการสูบบุหรี่ การดื่มแอลกอฮอล์ และการเคี้ยวหมาก
2. ตรวจภายในช่องปาก เพื่อประเมินการมีรอยโรคหรือแผลในช่องปาก
3. การเก็บน้ำบ้วนปากจากอาสาสมัครที่เข้าร่วมโครงการ ทำโดยวิธีการใช้น้ำเกลือที่สะอาดผ่านการทำให้ปราศจากเชื้อปริมาณ 10 มิลลิลิตร อมกลั้วไว้ในช่องปากประมาณ 15 วินาที จากนั้นให้บ้วนน้ำเกลือดังกล่าวผ่านกรวยลงในหลอดที่สะอาดผ่านการทำให้ปราศจากเชื้อขนาด 15 มิลลิลิตร แล้วนำไปเก็บในตู้เย็นอุณหภูมิ 4 องศาจนกว่าจนกว่าจะถึงขั้นตอนการดำเนินการต่อไป

คุณสมบัติและความรับผิดชอบของอาสาสมัครในกลุ่มที่สูบบุหรี่

1. อาสาสมัครจะต้องเป็นผู้ที่มีสุขภาพร่างกายแข็งแรงไม่เป็นมะเร็งในช่องปากหรือมะเร็งในอวัยวะอื่นๆ
2. อาสาสมัครจะต้องไม่มีรอยโรคหรือแผลในช่องปาก

3. อาสาสมัครจะต้องมีประวัติการสูบบุหรี่ ร่วมกับอาจมีหรือไม่มีประวัติการดื่มแอลกอฮอล์หรือการเคี้ยวหมาก
4. อาสาสมัครที่ขาดคุณสมบัติข้อใดข้อหนึ่งข้างต้นอาจถูกร้องขอให้ถอนตัวจากการวิจัย
5. อาสาสมัครที่ไม่พร้อมหรือขาดคุณสมบัติข้อใดข้อหนึ่งสามารถถอนตัวจากการวิจัยได้

คุณสมบัติและความรับผิดชอบของอาสาสมัครในกลุ่มที่ไม่สูบบุหรี่

1. อาสาสมัครจะต้องเป็นผู้ที่มีสุขภาพร่างกายแข็งแรงไม่เป็นมะเร็งในช่องปากหรือมะเร็งในอวัยวะอื่นๆ
2. อาสาสมัครจะต้องไม่มีรอยโรคหรือแผลในช่องปาก
3. อาสาสมัครจะต้องไม่มีประวัติการสูบบุหรี่ ร่วมกับอาจมีหรือไม่มีประวัติการดื่มแอลกอฮอล์หรือการเคี้ยวหมาก
4. อาสาสมัครที่ขาดคุณสมบัติข้อใดข้อหนึ่งข้างต้นอาจถูกร้องขอให้ถอนตัวจากการวิจัย
5. อาสาสมัครที่ไม่พร้อมหรือขาดคุณสมบัติข้อใดข้อหนึ่งสามารถถอนตัวจากการวิจัยได้

ค่าตอบแทนที่อาสาสมัครจะได้รับ

ไม่มี

ความเสี่ยงที่อาจจะเกิดขึ้นการเก็บน้ำบ้วนปากด้วยวิธีการใช้น้ำเกลือที่สะอาดผ่านการทำให้ปราศจากเชื้อไม่มีความเสี่ยงจากภาวะที่ไม่พึงประสงค์

ระยะเวลาที่คาดว่าอาสาสมัครจะต้องเกี่ยวข้องกับการวิจัย

1 ครั้งภายในวันที่ทำการเก็บน้ำบ้วนปาก

จำนวนของอาสาสมัครโดยประมาณที่จะใช้ในการวิจัย

อาสาสมัครที่จะใช้ในการวิจัยโดยประมาณ

กลุ่มที่สูบบุหรี่จำนวน 100 คน

กลุ่มที่ไม่สูบบุหรี่จำนวน 50 คน

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัยนี้

เพื่อศึกษาด้านชีววิทยาของเซลล์เยื่อบุช่องปากต่อการเปลี่ยนแปลงเมทิเลชันของไลน์-1 จากการสูบบุหรี่

หมายเหตุ หากอาสาสมัครมีข้อสงสัยใดๆ สามารถสอบถามรายละเอียดเพิ่มเติมได้จากผู้วิจัยทุกเมื่อที่ ทญ. ศิริพร วังศรี ภาควิชาศัลยศาสตร์ คณะทันตแพทยศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย โทรศัพท์ 086-866-3225

เอกสารยินยอมเข้าร่วมการวิจัย (Consent Form)

การวิจัยเรื่อง

การเปรียบเทียบเมทิเลชันของไลน์-1 ในคนที่สูบบุหรี่กับไม่สูบบุหรี่

Comparison of LINE-1 methylation between smokers and non-smokers

ก่อนที่จะลงนามในใบยินยอมให้ทำการวิจัยนี้ ข้าพเจ้าได้รับการอธิบายจากผู้วิจัยถึงวัตถุประสงค์ของการวิจัย วิธีการวิจัย รวมทั้งประโยชน์ที่จะเกิดขึ้นจากการวิจัยอย่างละเอียดและมีความเข้าใจดีแล้ว

ผู้วิจัยรับรองว่าจะตอบคำถามต่างๆ ที่ข้าพเจ้าสงสัยด้วยความเต็มใจไม่ปิดบังซ่อนเร้นจนข้าพเจ้าพอใจ

ข้าพเจ้าเข้าร่วมโครงการวิจัยนี้โดยสมัครใจ ข้าพเจ้ามีสิทธิที่จะบอกเลิกการเข้าร่วมในโครงการวิจัยนี้เมื่อใดก็ได้และการบอกเลิกการเข้าร่วมการวิจัยนี้ จะไม่มีผลใดๆต่อข้าพเจ้า

ผู้วิจัยรับรองว่าจะเก็บข้อมูลเฉพาะเกี่ยวกับตัวข้าพเจ้าเป็นความลับ และจะเปิดเผยได้เฉพาะในรูปที่เป็นสรุปผลการวิจัย การเปิดเผยข้อมูลเกี่ยวกับตัวข้าพเจ้าต่อหน่วยงานต่างๆ ที่เกี่ยวข้องกระทำได้เฉพาะกรณีจำเป็น ด้วยเหตุผลทางวิชาการเท่านั้น

ผู้วิจัยรับรองว่าหากเกิดอันตรายใดๆ จากการวิจัยดังกล่าว ข้าพเจ้าจะได้รับการรักษาพยาบาลโดยไม่คิดมูลค่า

ข้าพเจ้าได้อ่านข้อความข้างต้นแล้ว และมีความเข้าใจดีทุกประการ และได้ลงนามในใบยินยอมนี้ด้วยความเต็มใจ

ลงนาม.....ผู้ยินยอม

(.....)

ลงนาม.....พยาน

(.....)

ลงนาม.....พยาน

(.....)

ลงนาม.....หัวหน้าโครงการวิจัย

(..... .....ทญ. ศิริพร วังศรี.....)

วันให้คำยินยอมเข้าร่วมวิจัย วันที่.....เดือน.....พ.ศ.....

ข้าพเจ้าไม่สามารถอ่านหนังสือได้ แต่ผู้วิจัยได้อ่านข้อความในใบยินยอมนี้ให้แก่ข้าพเจ้าฟังจนเข้าใจดีแล้ว ข้าพเจ้าจึงลงนาม หรือประทับลายนิ้วหัวแม่มือขวาของข้าพเจ้าในใบยินยอมนี้ด้วยความเต็มใจ

ลงนาม.....ผู้ยินยอม

(.....)

ลงนาม.....พยาน

(..... .)

ลงนาม.....พยาน

(..... .)

ลงนาม.....หัวหน้าโครงการวิจัย

(.....ทญ. ศิริพร วังศรี..... .)

วันให้คำยินยอมเข้าร่วมวิจัย วันที่.....เดือน.....พ.ศ.....

ในกรณีที่ผู้ถูกทดลองยังไม่บรรลุนิติภาวะ จะต้องได้รับการยินยอมจากผู้ปกครองหรือผู้อุปการะโดยชอบด้วยกฎหมาย

ลงนาม.....ผู้ยินยอม

(.....)

ลงนาม.....พยาน

(..... .)

ลงนาม.....พยาน

(..... .)

ลงนาม.....หัวหน้าโครงการวิจัย

(.....ทญ. ศิริพร วังศรี..... .)

วันให้คำยินยอมเข้าร่วมวิจัย วันที่.....เดือน.....พ.ศ.....

เอกสารยกเลิกการยินยอมเข้าร่วมวิจัย (Withdrawal Form)

การวิจัยเรื่อง

การเปรียบเทียบเมทิเลชันของไลน์-1 ในคนที่สูบบุหรี่กับไม่สูบบุหรี่

Comparison of LINE-1 methylation between smokers and non-smokers

เหตุผลในการยกเลิกการยินยอมเข้าร่วมวิจัย

- ❑ ย้ายภูมิลำเนา
- ❑ ไม่สะดวกในการเดินทาง
- ❑ เหตุผลอื่น

.....

ลงนาม.....ผู้ยกเลิกการยินยอม

(.....)

ลงนาม.....พยาน

(.....)

ลงนาม.....พยาน

(.....)

ลงนาม.....หัวหน้าโครงการวิจัย

(.....ทญ. ศิริพร วังศรี.....)

วันยกเลิกการยินยอมเข้าร่วมวิจัย วันที่.....เดือน.....พ.ศ. .....

APPENDIX B

Questionaires

แบบสอบถามวิทยานิพนธ์เรื่อง

การเปรียบเทียบเมทิเลชันของไลน์-1 ในคนที่สูบบุหรี่กับไม่สูบบุหรี่

Comparison of LINE-1 methylation between smokers and non-smokers

หมายเลข.....

ส่วนที่ 1 ข้อมูลส่วนบุคคล

1. ชื่อ.....นามสกุล.....ที่อยู่ที่ติดต่อได้.....

..... เบอร์โทร……

2. อายุ…… ปี

3. เพศ O ชาย O หญิง

4. อาชีพหลักของท่าน

O เกษตรกร / ประมง	O ลูกจ้างขายของ ลูกจ้างทั่วไป
O พนักงานขาย มีเงินเดือนประจำ	O พนักงานขาย/เซลล์แมน
O พนักงานบริษัท มีเงินเดือนประจำ	O ผู้บริหาร ผู้จัดการ เจ้าของกิจการ
O งานวิชาชีพ/วิชาการ/รับราชการ	O งานบริการ เช่น ช่างตัดผม ช่างซ่อมรถ
O ขนส่ง ขับยานพาหนะทุกประเภท	O กรรมกรก่อสร้าง /กรรมกรทั่วไป/ช่างฝีมือ
O คนทำงานโรงงาน	O ลูกจ้างรัฐ/ รัฐวิสาหกิจ
O นักเรียน / นักศึกษา	O ไม่มีงานทำ-แม่บ้าน

ส่วนที่ 2 การตรวจช่องปาก

O ไม่พบแผล รอยโรค หรือความผิดปกติในช่องปาก

O พบแผล รอยโรค หรือความผิดปกติในช่องปาก ระบุ.....

ส่วนที่ 3 ข้อมูลพฤติกรรมการสูบบุหรี่

1. พฤติกรรมการสูบบุหรี่ของท่าน	O ปัจจุบันยังสูบอยู่
	O เคยสูบแต่เลิกแล้ว
	O ไม่เคยสูบ (ข้ามไปตอบข้อ 10 ในส่วนที่ 3)
2. ชนิดของบุหรี่ที่ใช้ปัจจุบัน	O บุหรี่สำเร็จรูป (Manufactured cigarettes)
	O บุหรี่ชนิดมวน (Roll-up cigarettes)
	O กล้องยาสูบ (Pipe)
	O ซิการ์ (Cigars)
	O อื่น ๆ
	O ใช้ร่วมกันระหว่าง ระบุ……..
3. ประเภทของบุหรี่ที่สูบ	O สูบบุหรี่ในประเทศเท่านั้น
	O สูบบุหรี่ต่างประเทศเท่านั้น
	O สูบบุหรี่ทั้งสองประเภท

4. ยี่ห้อของบุหรี่ในประเทศที่สูบ
O กรองทิพย์
O สายฝน
O สามิต
O กรุงทอง
O วอนเดอร์
O อื่นๆ ระบุ.....

5. ยี่ห้อของบุหรี่ต่างประเทศที่สูบ
O มาร์โบโร (Marlboro)
O แอลเอ็ม (LM)
O Mind Seven
O Dunhill
O 555
O อื่น ๆ ระบุ.....

6. ความถี่ในการสูบบุหรี่
O ทุกวัน
O วันเว้นวัน
O สัปดาห์ละ 2 – 3 ครั้ง
O สัปดาห์ละครั้ง
O น้อยกว่าสัปดาห์ละครั้ง
O อื่น ๆ ระบุ.....

7. จำนวนที่สูบ/วัน .....ซอง.....มวน

8. อายุที่เริ่มสูบ.....ปี.....เดือน

9. จำนวนปีที่สูบบุหรี่ตั้งแต่เริ่มสูบจนถึงปัจจุบัน .. ..ปี.....เดือน
กรณีที่เลิกสูบแล้ว จำนวนปีที่เลิกสูบบุหรี่ตั้งแต่เลิกจนถึงปัจจุบัน.. .....ปี.....เดือน

10. ในที่พักอาศัยของท่านมีผู้ใดสูบบุหรี่บ้าง O เพื่อน
O ญาติ
O พี่น้อง
O บิดา
O มารดา
O อื่นๆ

ส่วนที่ 3 ข้อมูลพฤติกรรมการดื่มแอลกอฮอล์

1. พฤติกรรมการดื่มแอลกอฮอล์ของท่าน

O ปัจจุบันดื่มอยู่ O เคยดื่มแต่เลิกแล้ว O ไม่เคยดื่ม (ข้ามไปตอบส่วนที่ 4)

2. เครื่องดื่มแอลกอฮอล์ประเภทใดบ้างที่ท่านดื่ม (ตอบได้มากกว่า 1 ข้อ)

O เหล้าสี
O เหล้าขาว
O สุราพื้นบ้าน
O เชี่ยงชุน/ยาดอง
O เบียร์
O ไวน์
O เครื่องดื่มผสมแอลกอฮอล์/ดริ้ง/คอกเทล
O อื่น ๆ ระบุ.....

3. ท่านดื่มแอลกอฮอล์บ่อยแค่ไหน

O ทุกวัน
O 1-2 ครั้ง/สัปดาห์
O 3-4 ครั้ง/สัปดาห์
O 1-2 ครั้ง/เดือน
O นาน ๆ ครั้ง
O ดื่มในช่วงเทศกาล/วันเกิด
O เฉพาะในงานสังคม
O อื่น ๆ ระบุ.....

4. ปริมาณแอลกอฮอล์ที่ดื่ม/ครั้ง.....

5. อายุที่เริ่มดื่มแอลกอฮอล์.....ปี.....เดือน

6. จำนวนปีที่ดื่มแอลกอฮอล์ตั้งแต่เริ่มดื่มจนถึงปัจจุบัน .. ..ปี.....เดือน

กรณีที่เลิกดื่มแล้ว จำนวนปีที่เลิกดื่มตั้งแต่เลิกจนถึงปัจจุบัน.. .....ปี.....เดือน

ส่วนที่ 4 ข้อมูลพฤติกรรมการเคี้ยวหมาก

1. ท่านเคยเคี้ยวหมากหรือไม่

O ปัจจุบันยังเคี้ยวหมากอยู่
O เคยเคี้ยวหมากแต่เลิกแล้ว
O ไม่เคยเคี้ยวหมาก (จบการทำแบบสอบถาม)

2. จำนวนครั้งที่เคี้ยวหมาก / วัน.....ครั้ง

3. อายุที่เริ่มเคี้ยวหมาก.....ปี.....เดือน

4. จำนวนปีที่เคี้ยวหมาก ตั้งแต่เริ่มเคี้ยวจนถึงปัจจุบัน .. ..ปี.....เดือน

กรณีที่เลิกเคี้ยวหมากแล้ว จำนวนปีที่เลิกเคี้ยวตั้งแต่เลิกจนถึงปัจจุบัน.. .....ปี.....เดือน

APPENDIX C

Statistic Output

The Percentage of Each LINE-1 Methylation Pattern in Males and Females

1. Normality test in the value of each methylation pattern in male and female of CS, FS and NS

One-Sample Kolmogorov-Smirnov Test

Smoking	Sex			^{m}C	$^{m}C^{m}C$	$^{u}C^{u}C$	$^{m}C^{u}C$	$^{u}C^{m}C$	$^{m}C^{u}C+^{u}C^{m}C$
current smokers	male	N		80	80	80	80	80	80
		Normal Parameters[a,,b]	Mean	42.3399	18.2094	33.5296	23.8310	24.4294	48.2609
			Std. Deviation	2.62229	4.68389	3.82630	3.61408	8.02108	6.75634
		Most Extreme Differences	Absolute	.057	.049	.090	.084	.059	.073
			Positive	.057	.036	.075	.062	.059	.070
			Negative	-.043	-.049	-.090	-.084	-.058	-.073
		Kolmogorov-Smirnov Z		.509	.436	.798	.743	.528	.647
		Asymp. Sig. (2-tailed)		.958	.991	.547	.639	.943	.797
	female	N		16	16	16	16	16	16
		Normal Parameters[a,,b]	Mean	41.2875	15.9263	33.3500	25.0306	25.6919	50.7237
			Std. Deviation	2.48876	5.02955	3.37988	3.73765	6.80322	6.97719
		Most Extreme Differences	Absolute	.134	.148	.136	.181	.164	.102
			Positive	.134	.094	.108	.124	.164	.088
			Negative	-.075	-.148	-.136	-.181	-.141	-.102
		Kolmogorov-Smirnov Z		.534	.593	.544	.722	.656	.410
		Asymp. Sig. (2-tailed)		.938	.874	.928	.674	.782	.996
former smokers	male	N		15	15	15	15	15	15
		Normal Parameters[a,,b]	Mean	41.0050	15.4707	33.4621	24.7193	26.3457	51.0657
			Std. Deviation	2.43994	4.95096	1.34788	4.14319	5.45561	5.37065
		Most Extreme Differences	Absolute	.123	.116	.123	.271	.135	.136
			Positive	.116	.109	.112	.271	.135	.102
			Negative	-.123	-.116	-.123	-.186	-.134	-.136
		Kolmogorov-Smirnov Z		.460	.433	.460	1.014	.506	.507
		Asymp. Sig. (2-tailed)		.984	.992	.984	.255	.960	.959
	female	N		2	2	2	2	2	2
		Normal Parameters[a,,b]	Mean	42.2150	15.4900	31.0700	23.0700	30.3700	53.4400
			Std. Deviation	2.28395	.38184	4.94975	2.88500	8.21658	5.33159

		Most Extreme Differences	Absolute	.260	.260	.260	.260	.260	.260
			Positive	.260	.260	.260	.260	.260	.260
			Negative	-.260	-.260	-.260	-.260	-.260	-.260
		Kolmogorov-Smirnov Z		.368	.368	.368	.368	.368	.368
		Asymp. Sig. (2-tailed)		.999	.999	.999	.999	.999	.999
never smokers	male	N		35	35	35	35	35	35
		Normal Parameters[a,,b]	Mean	42.5500	16.9407	31.8414	26.7300	24.4886	51.2179
			Std. Deviation	1.81819	4.03109	2.31906	2.13137	4.50196	5.48156
		Most Extreme Differences	Absolute	.112	.201	.140	.186	.145	.241
			Positive	.063	.201	.129	.186	.133	.119
			Negative	-.112	-.124	-.140	-.139	-.145	-.241
		Kolmogorov-Smirnov Z		.418	.752	.525	.695	.541	.902
		Asymp. Sig. (2-tailed)		.995	.623	.946	.719	.931	.390
	female	N		25	25	25	25	25	25
		Normal Parameters[a,,b]	Mean	41.5087	15.2164	32.1980	28.1240	24.4620	52.5858
			Std. Deviation	2.32245	3.61068	3.03521	3.33173	6.32884	4.78752
		Most Extreme Differences	Absolute	.160	.147	.148	.108	.184	.209
			Positive	.097	.112	.148	.054	.184	.209
			Negative	-.160	-.147	-.103	-.108	-.110	-.138
		Kolmogorov-Smirnov Z		1.075	.983	.992	.723	1.237	1.402
		Asymp. Sig. (2-tailed)		.198	.288	.278	.672	.094	.039

a. Test distribution is Normal.

b. Calculated from data.

2. Comparison in the value of each methylation pattern between male and female in CS, FS and NS by independent sample *t*-test

Group Statistics

Smoking		Sex	N	Mean	Std. Deviation	Std. Error Mean
current smokers	^{m}C	male	80	42.3399	2.62229	.29503
		female	16	41.2875	2.48876	.62219
	$^{m}C^{m}C$	male	80	18.2094	4.68389	.52698
		female	16	15.9263	5.02955	1.25739
	$^{u}C^{u}C$	male	80	33.5296	3.82630	.43049
		female	16	33.3500	3.37988	.84497
	$^{m}C^{u}C$	male	80	23.8310	3.61408	.40662
		female	16	25.0306	3.73765	.93441
	$^{u}C^{m}C$	male	80	24.4294	8.02108	.90244
		female	16	25.6919	6.80322	1.70081
	$^{m}C^{u}C+^{u}C^{m}C$	male	80	48.2609	6.75634	.76015
		female	16	50.7237	6.97719	1.74430
former smokers	^{m}C	male	15	41.0050	2.43994	.65210
		female	2	42.2150	2.28395	1.61500
	$^{m}C^{m}C$	male	15	15.4707	4.95096	1.32320
		female	2	15.4900	.38184	.27000
	$^{u}C^{u}C$	male	15	33.4621	1.34788	.36024
		female	2	31.0700	4.94975	3.50000
	$^{m}C^{u}C$	male	15	24.7193	4.14319	1.10731
		female	2	23.0700	2.88500	2.04000
	$^{u}C^{m}C$	male	15	26.3457	5.45561	1.45807
		female	2	30.3700	8.21658	5.81000
	$^{m}C^{u}C+^{u}C^{m}C$	male	15	51.0657	5.37065	1.43537
		female	2	53.4400	5.33159	3.77000
never smokers	^{m}C	male	35	42.5500	1.81819	.48593
		female	25	41.5087	2.32245	.34621
	$^{m}C^{m}C$	male	35	16.9407	4.03109	1.07735
		female	25	15.2164	3.61068	.53825

	$^{u}C^{u}C$	male	35	31.8414	2.31906	.61979
		female	25	32.1980	3.03521	.45246
	$^{m}C^{u}C$	male	35	26.7300	2.13137	.56963
		female	25	28.1240	3.33173	.49666
	$^{u}C^{m}C$	male	35	24.4886	4.50196	1.20320
		female	25	24.4620	6.32884	.94345
	$^{m}C^{u}C+^{u}C^{m}C$	male	35	51.2179	5.48156	1.46501
		female	25	52.5858	4.78752	.71368

Independent Samples Test[a]

			Levene's Test for Equality of Variances		t-test for Equality of Means						
										95% Confidence Interval of the Difference	
Smoking			F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
current smokers	^{m}C	Equal variances assumed	.083	.774	1.476	93	.143	1.05237	.71312	-.36375	2.46849
		Equal variances not assumed			1.528	22.287	.141	1.05237	.68859	-.37462	2.47937
	$^{m}C^{m}C$	Equal variances assumed	.185	.668	1.756	93	.082	2.28312	1.29984	-.29811	4.86434
		Equal variances not assumed			1.675	20.610	.109	2.28312	1.36335	-.55540	5.12164
	$^{u}C^{u}C$	Equal variances assumed	.000	.995	.174	93	.862	.17962	1.03022	-1.86620	2.22544
		Equal variances not assumed			.189	23.493	.851	.17962	.94831	-1.77984	2.13908
	$^{m}C^{u}C$	Equal variances assumed	.268	.606	-1.204	93	.232	-1.19961	.99634	-3.17815	.77893
		Equal variances not assumed			-1.177	21.073	.252	-1.19961	1.01905	-3.31839	.91917
	$^{u}C^{m}C$	Equal variances assumed	.121	.728	-.588	93	.558	-1.26251	2.14864	-5.52928	3.00427

		Equal variances not assumed			-.656	24.265	.518	-1.26251	1.92539	-5.23403	2.70901
	$^{m}C^{u}C$ + $^{u}C^{m}C$	Equal variances assumed	.107	.744	-1.323	93	.189	-2.46286	1.86215	-6.16072	1.23499
		Equal variances not assumed			-1.294	21.092	.210	-2.46286	1.90273	-6.41876	1.49304
former smokers	^{m}C	Equal variances assumed	.112	.742	-.659	14	.521	-1.21000	1.83625	-5.14837	2.72837
		Equal variances not assumed			-.695	1.350	.587	-1.21000	1.74168	-13.48731	11.06731
	$^{m}C^{m}C$	Equal variances assumed	2.657	.125	-.005	14	.996	-.01929	3.60726	-7.75609	7.71752
		Equal variances not assumed			-.014	13.794	.989	-.01929	1.35047	-2.91981	2.88124
	$^{u}C^{u}C$	Equal variances assumed	15.85	.001	1.707	14	.110	2.39214	1.40143	-.61363	5.39791
		Equal variances not assumed			.680	1.021	.618	2.39214	3.51849	-40.17118	44.95546
	$^{m}C^{u}C$	Equal variances assumed	.039	.845	.537	14	.600	1.64929	3.07380	-4.94335	8.24192
		Equal variances not assumed			.711	1.665	.564	1.64929	2.32115	-10.53406	13.83263
	$^{u}C^{m}C$	Equal variances assumed	.460	.509	-.934	14	.366	-4.02429	4.30680	-13.26146	5.21289
		Equal variances not assumed			-.672	1.130	.613	-4.02429	5.99016	-62.30983	54.26126
	$^{m}C^{u}C$ + $^{u}C^{m}C$	Equal variances assumed	.009	.925	-.585	14	.568	-2.37429	4.05773	-11.07724	6.32867
		Equal variances not assumed			-.589	1.309	.641	-2.37429	4.03400	-32.28961	27.54104
never smokers	^{m}C	Equal variances assumed	1.220	.274	1.534	57	.130	1.04133	.67863	-.31759	2.40026
		Equal variances not assumed			1.745	27.457	.092	1.04133	.59665	-.18194	2.26460
	$^{m}C^{m}C$	Equal variances assumed	.799	.375	1.518	57	.134	1.72427	1.13558	-.54969	3.99823
		Equal variances not assumed			1.432	19.933	.168	1.72427	1.20433	-.78845	4.23699
	$^{u}C^{u}C$	Equal variances assumed	.145	.705	-.404	57	.688	-.35657	.88366	-2.12607	1.41293

	Equal variances not assumed			-.465	28.183	.646	-.35657	.76738	-1.92801	1.21487
$^{m}C^{u}C$	Equal variances assumed	1.822	.182	-1.470	57	.147	-1.39400	.94842	-3.29318	.50518
	Equal variances not assumed			-1.845	34.404	.074	-1.39400	.75575	-2.92920	.14120
$^{u}C^{m}C$	Equal variances assumed	.133	.717	.015	57	.988	.02657	1.82442	-3.62676	3.67990
	Equal variances not assumed			.017	30.494	.986	.02657	1.52898	-3.09391	3.14705
$^{m}C^{u}C$ + $^{u}C^{m}C$	Equal variances assumed	1.458	.232	-.902	57	.371	-1.36792	1.51616	-4.40397	1.66813
	Equal variances not assumed			-.839	19.577	.411	-1.36792	1.62960	-4.77192	2.03608

a. No statistics are computed for one or more split files

APPENDIX D

Statistic Output

The interaction effect of smoking and alcohol by two way ANOVA

1. Normality test in the value of each methylation pattern in smoking and alcohol drinking factors.

One-Sample Kolmogorov-Smirnov Test

Smoking	Alcohol			^{m}C	$^{m}C^{m}C$	$^{u}C^{u}C$	$^{m}C^{u}C$	$^{u}C^{m}C$	$^{m}C^{u}C+^{u}C^{m}C$
current smoke	current drink	N		89	89	89	89	89	89
		Normal Parameters[a,,b]	Mean	42.2293	17.8714	33.4129	24.1073	24.6075	48.7154
			Std. Deviation	2.65255	4.73866	3.49497	3.56412	7.71230	6.41809
		Most Extreme Differences	Absolute	.047	.062	.082	.075	.066	.068
			Positive	.047	.040	.048	.040	.066	.068
			Negative	-.030	-.062	-.082	-.075	-.056	-.040
		Kolmogorov-Smirnov Z		.436	.570	.758	.689	.610	.631
		Asymp. Sig. (2-tailed)		.991	.901	.614	.729	.850	.820
	never drink	N		7	7	7	7	7	7
		Normal Parameters[a,,b]	Mean	40.6600	15.6143	34.2914	24.2800	25.8143	50.0957
			Std. Deviation	1.00904	4.64411	5.97597	4.88742	9.19665	10.51194
		Most Extreme Differences	Absolute	.196	.220	.189	.186	.178	.186
			Positive	.167	.220	.189	.186	.136	.160
			Negative	-.196	-.115	-.189	-.166	-.178	-.186
		Kolmogorov-Smirnov Z		.518	.583	.500	.491	.472	.493
		Asymp. Sig. (2-tailed)		.951	.886	.964	.969	.979	.968
never smoke	current drink	N		15	15	15	15	15	15
		Normal Parameters[a,,b]	Mean	41.5586	14.8136	31.6971	25.8229	27.6664	53.4886
			Std. Deviation	2.98782	5.82908	2.78271	3.71750	8.81023	6.90737
		Most Extreme Differences	Absolute	.190	.191	.142	.157	.261	.177
			Positive	.103	.132	.084	.157	.261	.177
			Negative	-.190	-.191	-.142	-.141	-.132	-.122
		Kolmogorov-Smirnov Z		.712	.713	.530	.587	.977	.661
		Asymp. Sig. (2-tailed)		.691	.689	.941	.881	.295	.774
	never drink	N		45	45	45	45	45	45
		Normal Parameters[a,,b]	Mean	41.8171	15.8782	32.2429	28.4062	23.4733	51.8793
			Std. Deviation	1.99496	2.87150	2.91071	2.68541	4.34080	4.18652

	Most Extreme Differences	Absolute	.171	.106	.158	.077	.092	.189
		Positive	.085	.078	.158	.064	.092	.189
		Negative	-.171	-.106	-.114	-.077	-.088	-.149
	Kolmogorov-Smirnov Z		1.150	.709	1.061	.515	.615	1.266
	Asymp. Sig. (2-tailed)		.142	.697	.210	.954	.844	.081

a. Test distribution is Normal.

b. CAlcoholulated from data.

1. The interaction effect of smoking and alcohol by two way ANOVA

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
Smoking	1	current smoke	96
	3	never smoke	60
Alcohol	1	current drink	99
	3	never drink	52

Tests of Between-Subjects Effects

Dependent Variable:mC

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	21.149[a]	3	7.050	1.167	.325
Intercept	111345.504	1	111345.504	18425.973	.000
Smoking	.953	1	.953	.158	.692
Alcohol	6.920	1	6.920	1.145	.286
Smoking * Alcohol	13.457	1	13.457	2.227	.138
Error	888.300	147	6.043		
Total	266912.371	151			
Corrected Total	909.449	150			

a. R Squared = .023 (Adjusted R Squared = .003)

Tests of Between-Subjects Effects

Dependent Variable:mCmC

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	199.689[a]	3	66.563	3.470	.018
Intercept	16589.626	1	16589.626	864.737	.000
Smoking	31.441	1	31.441	1.639	.202
Alcohol	5.728	1	5.728	.299	.586
Smoking * Alcohol	44.444	1	44.444	2.317	.130
Error	2820.136	147	19.185		
Total	46092.197	151			
Corrected Total	3019.826	150			

a. R Squared = .066 (Adjusted R Squared = .047)

Tests of Between-Subjects Effects

Dependent Variable:uCuC

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	75.286[a]	3	25.095	2.153	.096
Intercept	69803.295	1	69803.295	5987.464	.000
Smoking	57.075	1	57.075	4.896	.028
Alcohol	8.170	1	8.170	.701	.404
Smoking * Alcohol	.446	1	.446	.038	.845
Error	1713.761	147	11.658		
Total	165689.269	151			
Corrected Total	1789.047	150			

a. R Squared = .042 (Adjusted R Squared = .023)

Tests of Between-Subjects Effects

Dependent Variable:mCuC

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	556.357[a]	3	185.452	15.967	.000
Intercept	42413.534	1	42413.534	3651.772	.000
Smoking	137.456	1	137.456	11.835	.001
Alcohol	30.595	1	30.595	2.634	.107
Smoking * Alcohol	23.407	1	23.407	2.015	.158
Error	1707.332	147	11.615		
Total	100879.285	151			
Corrected Total	2263.690	150			

a. R Squared = .246 (Adjusted R Squared = .230)

Tests of Between-Subjects Effects

Dependent Variable:uCmC

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	199.078[a]	3	66.359	1.329	.267
Intercept	41546.075	1	41546.075	831.839	.000
Smoking	2.076	1	2.076	.042	.839
Alcohol	35.921	1	35.921	.719	.398
Smoking * Alcohol	117.445	1	117.445	2.351	.127
Error	7341.891	147	49.945		
Total	98987.545	151			
Corrected Total	7540.968	150			

a. R Squared = .026 (Adjusted R Squared = .007)

Tests of Between-Subjects Effects

Dependent Variable:ᵘCᵐC+ᵐCᵘC

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	465.543[a]	3	155.181	4.137	.008
Intercept	167916.514	1	167916.514	4476.099	.000
Smoking	173.162	1	173.162	4.616	.033
Alcohol	.211	1	.211	.006	.940
Smoking * Alcohol	35.998	1	35.998	.960	.329
Error	5514.563	147	37.514		
Total	385973.208	151			
Corrected Total	5980.106	150			

a. R Squared = .078 (Adjusted R Squared = .059)

APPENDIX E

Statistic Output

The Percentage of Loci of Each LINE-1 Methylation Pattern in NS and CS

1. Normality test in the value of each methylation pattern in CS and NS

One-Sample Kolmogorov-Smirnov Test

Smoking			^{m}C	$^{m}C^{m}C$	$^{u}C^{u}C$	$^{m}C^{u}C$	$^{u}C^{m}C$	$^{m}C^{u}C+^{u}C^{m}C$
current smokers	N		96	96	96	96	96	96
	Normal Parameters[a,,b]	Mean	42.1626	17.8248	33.4994	24.0331	24.6420	48.6757
		Std. Deviation	2.61746	4.79364	3.73846	3.64298	7.81012	6.81946
	Most Extreme Differences	Absolute	.063	.055	.065	.065	.068	.060
		Positive	.063	.036	.062	.042	.068	.060
		Negative	-.033	-.055	-.065	-.065	-.052	-.052
	Kolmogorov-Smirnov Z		.612	.532	.632	.636	.662	.584
	Asymp. Sig. (2-tailed)		.848	.940	.819	.813	.772	.885
never smokers	N		60	60	60	60	60	60
	Normal Parameters[a,,b]	Mean	41.7617	15.6312	32.1070	27.8498	24.4123	52.2618
		Std. Deviation	2.22468	3.72061	2.84267	3.13420	5.87578	4.90433
	Most Extreme Differences	Absolute	.144	.104	.127	.092	.166	.172
		Positive	.078	.084	.127	.042	.166	.172
		Negative	-.144	-.104	-.084	-.092	-.078	-.133
	Kolmogorov-Smirnov Z		1.118	.809	.984	.711	1.289	1.331
	Asymp. Sig. (2-tailed)		.164	.530	.288	.692	.072	.058

a. Test distribution is Normal.

b. Calculated from data.

2. Comparison in the value of each methylation pattern between CS and NS by independent sample *t*-test

Group Statistics

	Smoking	N	Mean	Std. Deviation	Std. Error Mean
^{m}C	current smokers	96	42.1626	2.61746	.26855
	never smokers	60	41.7617	2.22468	.28721
$^{m}C^{m}C$	current smokers	96	17.8248	4.79364	.49182
	never smokers	60	15.6312	3.72061	.48033
$^{u}C^{u}C$	current smokers	96	33.4994	3.73846	.38356
	never smokers	60	32.1070	2.84267	.36699
$^{m}C^{u}C$	current smokers	96	24.0331	3.64298	.37376
	never smokers	60	27.8498	3.13420	.40462
$^{u}C^{m}C$	current smokers	96	24.6420	7.81012	.80130
	never smokers	60	24.4123	5.87578	.75856
$^{m}C^{u}C+^{u}C^{m}C$	current smokers	96	48.6757	6.81946	.69966
	never smokers	60	52.2618	4.90433	.63315

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
^{m}C	Equal variances assumed	2.029	.156	.983	153	.327	.40096	.40787	-.40482	1.20675
	Equal variances not assumed			1.020	140.065	.310	.40096	.39320	-.37640	1.17833
$^{m}C^{m}C$	Equal variances assumed	8.132	.005	3.016	153	.003	2.19368	.72737	.75669	3.63066
	Equal variances not assumed			3.191	146.496	.002	2.19368	.68746	.83506	3.55229
$^{u}C^{u}C$	Equal variances assumed	3.852	.052	2.468	153	.015	1.39237	.56412	.27790	2.50684
	Equal variances not assumed			2.623	147.688	.010	1.39237	.53084	.34334	2.44140
$^{m}C^{u}C$	Equal variances assumed	1.698	.194	-6.698	153	.000	-3.81678	.56985	-4.94257	-2.69099
	Equal variances not assumed			-6.929	139.084	.000	-3.81678	.55083	-4.90587	-2.72769
$^{u}C^{m}C$	Equal variances assumed	9.210	.003	.195	153	.845	.22967	1.17521	-2.09206	2.55140
	Equal variances not assumed			.208	148.263	.835	.22967	1.10340	-1.95076	2.41009
$^{m}C^{u}C+$ $^{u}C^{m}C$	Equal variances assumed	8.913	.003	-3.535	153	.001	-3.58615	1.01448	-5.59035	-1.58195
	Equal variances not assumed			-3.800	150.352	.000	-3.58615	.94361	-5.45060	-1.72170

APPENDIX F

Statistic Output

The Percentage of Loci of Each LINE-1 Methylation Pattern in matched cases

1. Normality test in the value of each methylation pattern in matched cases of CS and NS

One-Sample Kolmogorov-Smirnov Test

Smoking			^{m}C	$^{m}C^{m}C$	$^{u}C^{u}C$	$^{m}C^{u}C$	$^{u}C^{m}C$	$^{m}C^{u}C + ^{u}C^{m}C$
current smokers	N		29	29	29	29	29	29
	Normal Parameters[a,,b]	Mean	42.3248	17.4934	32.8428	23.8869	25.7759	49.6641
		Std. Deviation	3.12412	5.35418	4.31375	3.76678	7.99187	7.45065
	Most Extreme Differences	Absolute	.151	.118	.141	.158	.127	.122
		Positive	.151	.115	.141	.158	.127	.086
		Negative	-.099	-.118	-.113	-.105	-.105	-.122
	Kolmogorov-Smirnov Z		.814	.637	.757	.848	.687	.656
	Asymp. Sig. (2-tailed)		.522	.812	.615	.468	.734	.783
never smokers	N		29	29	29	29	29	29
	Normal Parameters[a,,b]	Mean	42.2762	17.3759	32.8234	26.9666	22.8345	49.8014
		Std. Deviation	2.22889	4.53804	3.39886	2.93338	5.91975	6.66327
	Most Extreme Differences	Absolute	.151	.113	.143	.135	.134	.219
		Positive	.071	.113	.143	.094	.126	.113
		Negative	-.151	-.097	-.105	-.135	-.134	-.219
	Kolmogorov-Smirnov Z		.813	.607	.772	.724	.721	1.180
	Asymp. Sig. (2-tailed)		.523	.855	.590	.670	.676	.124

a. Test distribution is Normal.

b. Calculated from data.

2. Comparison in the value of each methylation pattern between matched cases of current smokers and non-smokers by paired *t*-test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	^{m}C current smokers	42.3248	29	3.12412	.58013
	^{m}C non- smokers	42.2762	29	2.22889	.41389
Pair 2	$^{m}C^{m}C$ current smokers	17.4934	29	5.35418	.99425
	$^{m}C^{m}C$ non- smokers	17.3759	29	4.53804	.84269
Pair 3	$^{u}C^{u}C$ current smokers	32.84276	29	4.313748	.801043
	$^{u}C^{u}C$ non- smokers	32.8234	29	3.39886	.63115
Pair 4	$^{m}C^{u}C$ current smokers	23.8869	29	3.76678	.69947
	$^{m}C^{u}C$ non- smokers	26.9666	29	2.93338	.54472
Pair 5	$^{u}C^{m}C$ current smokers	25.7759	29	7.99187	1.48405
	$^{u}C^{m}C$ non- smokers	22.8345	29	5.91975	1.09927
Pair 6	$^{m}C^{u}C+^{u}C^{m}C$ current smokers	49.6641	29	7.45065	1.38355
	$^{m}C^{u}C+^{u}C^{m}C$ non- smokers	49.8014	29	6.66327	1.23734

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	^{m}C current smokers & ^{m}C non- smokers	29	-.253	.186
Pair 2	$^{m}C^{m}C$ current smokers & $^{m}C^{m}C$ non- smokers	29	.219	.253
Pair 3	$^{u}C^{u}C$ current smokers & $^{u}C^{u}C$ non- smokers	29	-.169	.382
Pair 4	$^{m}C^{u}C$ current smokers & $^{m}C^{u}C$ non- smokers	29	.238	.214
Pair 5	$^{u}C^{m}C$ current smokers & $^{u}C^{m}C$ non- smokers	29	.239	.213
Pair 6	$^{m}C^{u}C+^{u}C^{m}C$ current smokers & $^{m}C^{u}C+^{u}C^{m}C$ non- smokers	29	.257	.179

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	^{m}C current smokers & ^{m}C non- smokers	.04862	4.27137	.79317	-1.57612	1.67336	.061	28	.952
Pair 2	$^{m}C^{m}C$ current smokers & $^{m}C^{m}C$ non- smokers	.11759	6.21266	1.15366	-2.24558	2.48076	.102	28	.920
Pair 3	$^{u}C^{u}C$ current smokers & $^{u}C^{u}C$ non- smokers	.019310	5.925247	1.100291	-2.234533	2.273154	.018	28	.986
Pair 4	$^{m}C^{u}C$ current smokers & $^{m}C^{u}C$ non- smokers	-3.07966	4.18820	.77773	-4.67276	-1.48655	-3.960	28	.000
Pair 5	$^{u}C^{m}C$ current smokers & $^{u}C^{m}C$ non- smokers	2.94138	8.73700	1.62242	-.38200	6.26476	1.813	28	.081
Pair 6	$^{m}C^{u}C+^{u}C^{m}C$ current smokers & $^{m}C^{u}C+^{u}C^{m}C$ non- smokers	-.13724	8.62572	1.60176	-3.41829	3.14381	-.086	28	.932

APPENDIX G

Statistics Output

The Percentage of Loci of Each LINE-1 Methylation Pattern in FS and NS

1. Normality test in the value of each methylation pattern in FS and NS

One-Sample Kolmogorov-Smirnov Test

Smoking			^{m}C	$^{m}C^{m}C$	$^{u}C^{u}C$	$^{m}C^{u}C$	$^{u}C^{m}C$	$^{m}C^{u}C + ^{u}C^{m}C$
former smokers	N		17	17	17	17	17	17
	Normal Parameters[a,,b]	Mean	41.1563	15.4731	33.1631	24.5131	26.8488	51.3625
		Std. Deviation	2.38288	4.61016	1.96863	3.96856	5.67322	5.24888
	Most Extreme Differences	Absolute	.102	.138	.206	.265	.113	.130
		Positive	.074	.138	.137	.265	.113	.092
		Negative	-.102	-.103	-.206	-.153	-.096	-.130
	Kolmogorov-Smirnov Z		.407	.550	.822	1.061	.453	.521
	Asymp. Sig. (2-tailed)		.996	.923	.509	.211	.986	.949
never smokers	N		60	60	60	60	60	60
	Normal Parameters[a,,b]	Mean	41.7617	15.6312	32.1070	27.8498	24.4123	52.2618
		Std. Deviation	2.22468	3.72061	2.84267	3.13420	5.87578	4.90433
	Most Extreme Differences	Absolute	.144	.104	.127	.092	.166	.172
		Positive	.078	.084	.127	.042	.166	.172
		Negative	-.144	-.104	-.084	-.092	-.078	-.133
	Kolmogorov-Smirnov Z		1.118	.809	.984	.711	1.289	1.331
	Asymp. Sig. (2-tailed)		.164	.530	.288	.692	.072	.058

a. Test distribution is Normal.

b. Calculated from data.

2. Comparison in the value of each methylation pattern between FS and NS by independent sample *t*-test

Group Statistics

	Smoking	N	Mean	Std. Deviation	Std. Error Mean
^{m}C	former smokers	16	41.1563	2.38288	.59572
	never smokers	60	41.7617	2.22468	.28721
$^{m}C^{m}C$	former smokers	16	15.4731	4.61016	1.15254
	never smokers	60	15.6312	3.72061	.48033
$^{u}C^{u}C$	former smokers	16	33.1631	1.96863	.49216
	never smokers	60	32.1070	2.84267	.36699
$^{m}C^{u}C$	former smokers	16	24.5131	3.96856	.99214
	never smokers	60	27.8498	3.13420	.40462
$^{u}C^{m}C$	former smokers	16	26.8488	5.67322	1.41830
	never smokers	60	24.4123	5.87578	.75856
$^{m}C^{u}C+^{u}C^{m}C$	former smokers	16	51.3625	5.24888	1.31222
	never smokers	60	52.2618	4.90433	.63315

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
^{m}C	Equal variances assumed	.217	.642	-.953	74	.344	-.60542	.63522	-1.87113	.66030
	Equal variances not assumed			-.915	22.475	.370	-.60542	.66134	-1.97527	.76444
$^{m}C^{m}C$	Equal variances assumed	1.068	.305	-.143	74	.886	-.15804	1.10219	-2.35420	2.03811
	Equal variances not assumed			-.127	20.506	.901	-.15804	1.24862	-2.75851	2.44243
$^{u}C^{u}C$	Equal variances assumed	1.764	.188	1.396	74	.167	1.05612	.75647	-.45117	2.56342
	Equal variances not assumed			1.720	33.672	.095	1.05612	.61392	-.19196	2.30421
$^{m}C^{u}C$	Equal variances assumed	.003	.956	-3.572	74	.001	-3.33671	.93422	-5.19819	-1.47523
	Equal variances not assumed			-3.114	20.262	.005	-3.33671	1.07148	-5.56991	-1.10350
$^{u}C^{m}C$	Equal variances assumed	.085	.771	1.484	74	.142	2.43642	1.64185	-.83504	5.70788
	Equal variances not assumed			1.515	24.303	.143	2.43642	1.60842	-.88100	5.75383
$^{m}C^{u}C+$ $^{u}C^{m}C$	Equal variances assumed	.537	.466	-.642	74	.523	-.89933	1.40010	-3.68910	1.89043
	Equal variances not assumed			-.617	22.487	.543	-.89933	1.45698	-3.91714	2.11847

APPENDIX H

Statistics Output
The Percentage of Loci of Each LINE-1 Methylation Pattern in Pack-Year Smoking Groups

1. Normality test in the value of each methylation pattern in pack-year smoking groups

One-Sample Kolmogorov-Smirnov Test

Pack year group (mean)			^{m}C	$^{m}C^{m}C$	$^{u}C^{u}C$	$^{m}C^{u}C$	$^{u}C^{m}C$	$^{m}C^{u}C+^{u}C^{m}C$
Group I <=13.23	N		54	54	54	54	54	54
	Normal Parameters[a,,b]	Mean	42.3193	17.6863	33.0480	24.2970	24.9678	49.2656
		Std. Deviation	2.82850	5.10220	4.00190	3.43219	8.41840	7.21790
	Most Extreme Differences	Absolute	.075	.108	.082	.124	.083	.084
		Positive	.075	.049	.068	.064	.083	.084
		Negative	-.063	-.108	-.082	-.124	-.068	-.052
	Kolmogorov-Smirnov Z		.552	.791	.604	.910	.607	.614
	Asymp. Sig. (2-tailed)		.921	.559	.859	.379	.855	.846
Group II >13.23	N		42	42	42	42	42	42
	Normal Parameters[a,,b]	Mean	42.0877	18.3931	34.2166	22.7051	24.6851	47.3900
		Std. Deviation	2.43684	4.65204	3.43592	3.04518	7.35800	6.56707
	Most Extreme Differences	Absolute	.127	.081	.161	.132	.088	.088
		Positive	.127	.079	.161	.067	.088	.054
		Negative	-.080	-.081	-.079	-.132	-.085	-.088
	Kolmogorov-Smirnov Z		.749	.481	.952	.779	.518	.520
	Asymp. Sig. (2-tailed)		.629	.975	.325	.579	.951	.950

a. Test distribution is Normal.

b. Calculated from data.

2. Comparison in the value of each methylation pattern between pack-year groups by independent sample *t*-test

Group Statistics

	Pack year group (mean)	N	Mean	Std. Deviation	Std. Error Mean
^{m}C	Group I <=13.23	54	42.3193	2.82850	.38491
	Group II >13.23	42	42.0877	2.43684	.41190
$^{m}C^{m}C$	Group I <=13.23	54	17.6863	5.10220	.69432
	Group II >13.23	42	18.3931	4.65204	.78634
$^{u}C^{u}C$	Group I <=13.23	54	33.0480	4.00190	.54459
	Group II >13.23	42	34.2166	3.43592	.58078
$^{m}C^{u}C$	Group I <=13.23	54	24.2970	3.43219	.46706
	Group II >13.23	42	22.7051	3.04518	.51473
$^{u}C^{m}C$	Group I <=13.23	54	24.9678	8.41840	1.14560
	Group II >13.23	42	24.6851	7.35800	1.24373
$^{m}C^{u}C+^{u}C^{m}C$	Group I <=13.23	54	49.2656	7.21790	.98223
	Group II >13.23	42	47.3900	6.56707	1.11004

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
^{m}C	Equal variances assumed	.507	.478	.398	87	.692	.23154	.58205	-.92535	1.38844
	Equal variances not assumed			.411	80.116	.682	.23154	.56375	-.89034	1.35343
$^{m}C^{m}C$	Equal variances assumed	.030	.863	-.661	87	.511	-.70685	1.07007	-2.83373	1.42004
	Equal variances not assumed			-.674	77.473	.502	-.70685	1.04900	-2.79548	1.38178
$^{u}C^{u}C$	Equal variances assumed	.349	.556	-1.421	87	.159	-1.16861	.82261	-2.80363	.46642
	Equal variances not assumed			-1.468	80.267	.146	-1.16861	.79617	-2.75295	.41573
$^{m}C^{u}C$	Equal variances assumed	1.860	.176	2.232	87	.028	1.59189	.71315	.17443	3.00936
	Equal variances not assumed			2.290	78.778	.025	1.59189	.69505	.20837	2.97542
$^{u}C^{m}C$	Equal variances assumed	.641	.426	.162	87	.871	.28263	1.74051	-3.17682	3.74209
	Equal variances not assumed			.167	79.470	.868	.28263	1.69093	-3.08278	3.64805
$^{m}C^{u}C+^{u}C^{m}C$	Equal variances assumed	.114	.737	1.240	87	.218	1.87556	1.51268	-1.13105	4.88217
	Equal variances not assumed			1.265	77.577	.210	1.87556	1.48222	-1.07556	4.82667

APPENDIX I

Microarray Expression Experiments of the Airway Epithelia of Smokers

Experiment 1 GSE 4302-2 (Genome-Wide Profiling of Airway Epithelial Cells in Asthmatics, Smokers and Healthy Controls) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.90	1.67	2.15	6.48E-24	0.80	0.70	0.91	8.99E-04

Case	Control
GSM98205	GSM98206
GSM98231	GSM98207
GSM98232	GSM98209
GSM98233	GSM98211
GSM98237	GSM98212
GSM98238	GSM98213
GSM98246	GSM98217
GSM98247	GSM98218
GSM98248	GSM98219
GSM98249	GSM98221
GSM98250	GSM98222
GSM98252	GSM98223
GSM98253	GSM98225
GSM98255	GSM98229
GSM98256	GSM98230
GSM98257	GSM98234
	GSM98235
	GSM98236
	GSM98239
	GSM98240
	GSM98241
	GSM98242
	GSM98243
	GSM98244
	GSM98245
	GSM98251
	GSM98254
	GSM98258

Experiment 2 GSE 19667-1 (Threshold of Biologic Response of the Small Airway Epithelium to Low Levels of Tobacco Smoke) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.68	1.48	1.90	5.53E-16	1.10	0.94	1.28	2.27E-01

Case	Control
GSM252871	GSM302396
GSM101115	GSM190155
GSM101116	GSM190156
GSM252876	GSM254149
GSM114089	GSM298220
GSM114090	GSM298221
GSM252879	GSM298222
,GSM252880	GSM298223
GSM252881	GSM298224
GSM252882	GSM254150
GSM252884	GSM298225
GSM298231	GSM302397
GSM252885	GSM254151
GSM254157	GSM298226
GSM254158	GSM298227
GSM254159	GSM254152
GSM298232	GSM298228
GSM298233	GSM298229
GSM298234	GSM300859
GSM298235	GSM469989
GSM298236	GSM350871
GSM298237	GSM350873
GSM298239	GSM434049
GSM298240	GSM350955
GSM254160	,GSM350956
GSM298241	GSM434050
GSM298242	GSM410161
GSM298243	GSM434051
GSM298244	GSM458579

Case	Control
GSM298245	GSM434052
GSM254161	GSM410162
GSM298246	GSM469990
	GSM410163
	GSM469991
	GSM469992
	GSM458580
	GSM469993
	GSM458581
	GSM458582
	GSM469994
	GSM469995
	GSM469996
	GSM469997
	GSM469998
	GSM469999

Experiment 3 GSE 19667-2 (Threshold of Biologic Response of the Small Airway Epithelium to Low Levels of Tobacco Smoke) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.74	1.51	2.01	2.97E-14	0.74	0.58	0.95	1.82E-02

Case	Control
GSM491043	GSM302396
GSM300878	GSM190155
GSM300871	GSM190156
GSM300872	GSM254149
GSM300874	GSM298220
GSM300880	GSM298221
GSM491044	GSM298222
	GSM298223
	GSM298224
	GSM254150
	GSM298225

Case	Control
	GSM302397
	GSM254151
	GSM298226
	GSM298227
	GSM254152
	GSM298228
	GSM298229
	GSM300859
	GSM469989
	GSM350871
	GSM350873
	GSM434049
	GSM350955
	GSM350956
	GSM434050
	GSM410161
	GSM434051
	GSM458579
	GSM434052
	GSM410162
	GSM469990
	GSM410163
	GSM469991
	GSM469992
	GSM458580
	GSM469993
	GSM458581
	GSM458582
	GSM469994
	G SM469995
	GSM469996
	GSM469997
	GSM469998
	GSM469999

Experiment 4 GSE 11906-8 (Quality Control in Microarray Assessment of Gene Expression in Human Airway Epithelium) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.63	1.41	1.87	1.05E-11	1.18	0.94	1.48	1.56E-01

Case	Control
GSM300913	GSM300881
GSM300914	GSM300882
GSM300915	GSM300883
	GSM300884
	GSM300885
	GSM300886
	GSM300887
	GSM300888
	GSM300889
	GSM300890
	GSM300891
	GSM300892
	,GSM300893
	GSM300894
	GSM300895
	GSM300896
	GSM300897

Experiment 5 GSE 11906-5 (Quality Control in Microarray Assessment of Gene Expression in Human Airway Epithelium) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
2.37	1.63	3.46	3.93E-06	1.32	1.16	1.49	1.79E-05

Case	Control
GSM101111	GSM101096
GSM101108	GSM101106

Case	Control
GSM101112	GSM101100
GSM101109	GSM101104
GSM101110	GSM190149
GSM252871	GSM101103
GSM101114	GSM101105
GSM101115	GSM101102
GSM101113	GSM101101
GSM101116	GSM190150
GSM252876	GSM190151
GSM114089	GSM190152
GSM114090	GSM190153
GSM252878	GSM298219
GSM252879	GSM190155
GSM298230	GSM190156
GSM252880	GSM254149
,GSM252881	GSM298220
GSM252882	GSM298221
GSM252884	GSM298222
GSM298231	GSM298223
GSM252885	GSM298224
GSM254157	GSM254150
GSM254158	GSM298225
GSM254159	GSM254151
GSM298232	GSM298226
GSM298233	GSM298227
GSM298234	GSM254152
GSM298235	GSM298228
GSM298236	GSM298229
,GSM298237	GSM300859
GSM298238	GSM300860
GSM298239	
GSM298240	
GSM254160	
GSM298241	
GSM298242	
GSM298243	
GSM298244	

Case	Control
G SM298245	
GSM254161	
,GSM298246	
GSM298247	
GSM300861	

Experiment 6 GSE 4498 (Expression data of small airway epithelium from phenotypically normal smokers and non-smokers) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
0.74	0.62	0.89	9.96E-04	1.54	1.25	1.89	4.85E-05

Case	Control
GSM101107	GSM101095
GSM101108	GSM101096
GSM101109	GSM101097
GSM101110	GSM101098
GSM101111	GSM101099
GSM101112	GSM101100
GSM101113	GSM101101
GSM101114	GSM101102
GSM101115	GSM101103
GSM101116	GSM101104
	GSM101105
	GSM101106

Experiment 7 GSE 13933-2 (Trachea Epithelium as a ??? Canary??? for Cigarette Smoking-induced Biologic Phenotype of Small Airway Epithelium) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.24	1.09	1.40	1.00E-03	1.52	1.32	1.75	2.58E-09

Case	Control
GSM300900	GSM300881
GSM350970	GSM300882
GSM300905	GSM300883
GSM300907	GSM300884
GSM300910	GSM300885
GSM300911	GSM300886
GSM350971	GSM300887
GSM350972	GSM300888
GSM350973	GSM300889
GSM350974	GSM300890
GSM300912	GSM350959
GSM350975	GSM350960
GSM350976	,GSM300893
GSM350977	GSM300894
GSM350978	GSM300895
GSM350979	GSM300896
GSM350980	GSM350961
GSM350981	GSM350962
GSM350982	GSM350963
	GSM350964
	GSM350965
	GSM300897
	GSM350966
	GSM350967
	GSM350968
	GSM350969

Experiment 8 GSE 11906-7 (Quality Control in Microarray Assessment of Gene Expression in Human Airway Epithelium) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.26	1.07	1.49	5.50E-03	1.21	1.02	1.44	2.94E-02

Case	Control
GSM300898	GSM300881
GSM300899	GSM300882
GSM300900	GSM300883
GSM300901	GSM300884
GSM300902	GSM300885
GSM300903	GSM300886
GSM300904	GSM300887
GSM300905	GSM300888
GSM300906	GSM300889
GSM300907	GSM300890
GSM300908	GSM300891
GSM300909	GSM300892
GSM300910	GSM300893
GSM300911	GSM300894
GSM300912	GSM300895
	GSM300896
	GSM300897

Experiment 9 GSE 3320 (Gene expression profile of small airway epithelium of normal non-smokers and normal smokers) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.51	1.12	2.04	6.11E-03	1.21	0.89	1.64	2.15E-01

Case	Control
GSM74802	GSM74797
GSM74819	GSM74798
GSM74820	GSM74799
GSM74821	GSM74800
GSM74822	GSM74801
GSM74823	

Experiment 10 GSE 8545-2 (Variability in Small Airway Epithelial Gene Expression Among Normal Smokers) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
0.68	0.47	0.98	3.84E-02	1.33	1.18	1.51	3.89E-06

Case	Control
GSM114089	GSM101095
GSM114090	GSM101097
GSM252879	GSM101098
GSM252882	GSM101096
GSM252884	,GSM101106
GSM252885	GSM101100
GSM254157	GSM101104
GSM254158	GSM101103
GSM254159	GSM101105
GSM254160	GSM101102
GSM254161	GSM101101
	GSM190151
	GSM252867
	GSM190153
	GSM254149
	GSM254150
	GSM254151
	GSM254152

Experiment 11 GSE 13933-1 (Trachea Epithelium as a ??? Canary??? for Cigarette Smoking-induced Biologic Phenotype of Small Airway Epithelium) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.19	1.00	1.42	4.47E-02	1.63	1.41	1.88	2.60E-11

Case	Control
GSM252871	GSM252855,
GSM252872	GSM252856,
GSM252874	GSM252857,
GSM252876	GSM252860,
GSM252879	GSM252861,
GSM252881	GSM252863,
GSM252882	GSM252865,
GSM252884	GSM190151,
GSM252885	GSM252867,
GSM254157	GSM190156
GSM298235	GSM298220
GSM298236	GSM298224
GSM298240	GSM254151
GSM298243	GSM298226
GSM298245	GSM298227
GSM350957	GSM254152
GSM410164	GSM298228
GSM350958	GSM300859
GSM410165	GSM350955
	GSM350956
	GSM410161
	GSM410162
	GSM410163

Experiment 12 GSE 11906-6 (Quality Control in Microarray Assessment of Gene Expression in Human Airway Epithelium) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.38	0.94	2.03	9.70E-02	0.98	0.88	1.10	7.87E-01

Case	Control
GSM300875	GSM101095
GSM300876	GSM101097
GSM300877	GSM101098
GSM300878	GSM101096
GSM300871	GSM101106
GSM300879	GSM101100
GSM300872	GSM101104
,GSM300873	GSM190149
GSM300874	GSM101103
GSM300880	GSM101105
	GSM101102
	GSM101101
	GSM190150
	GSM190151
	GSM190152
	GSM190153
	GSM298219
	GSM190155
	GSM190156
	GSM254149
	GSM298220
	GSM298221
	GSM298222
	GSM298223
	GSM298224
	GSM254150
	GSM298225
	GSM254151
	GSM298226

Case	Control
	GSM298227
	GSM254152
	GSM298228
	GSM298229
	GSM300859
	GSM300860

Experiment 13 GSE 27002 (Chronic Cigarette Smoke Exposure Results in Coordinated Methylation and Gene Expression Changes in Human Alveolar Macrophages) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.13	0.85	1.50	4.10E-01	1.28	0.95	1.72	9.90E-02

Case	Control
GSM665114	GSM665104
GSM665115	GSM665105
GSM665116	GSM665106
GSM665117	GSM665107
GSM665118	GSM665108
GSM665119	GSM665109
GSM665120	GSM665110
GSM665121	GSM665111
GSM665122	GSM665112
GSM665123	GSM665113
GSM665124	
GSM665125	
GSM665126	

Experiment 14 GSE 7895-1 (Reversible and Permanent effects of Tobacco Smoke Exposure on Airway Epithelial Gene Expression) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
1.07	0.82	1.38	6.28E-01	0.92	0.73	1.18	5.23E-01

Case	Control
GSM194224	GSM194203
GSM194225	GSM194204
GSM194226	GSM194205
GSM194227	GSM194206
GSM194228	GSM194207
GSM194229	GSM194208
,GSM194230	,GSM194209
GSM194231	GSM194210
GSM194232	GSM194211
GSM194233	GSM194212
GSM194234	GSM194213
GSM194235	GSM194214
GSM194236	GSM194215
GSM194237	GSM194216
GSM194238	GSM194217
GSM194239	GSM194218
GSM194240	GSM194219
GSM194241	GSM194220
GSM194242	GSM194221
GSM194243	GSM194222
GSM194244	GSM194223
GSM194245	
GSM194246	
GSM194247	
GSM194248	
GSM194249	
GSM194250	
GSM194251	
GSM194252	

Case	Control
GSM194253	
GSM194254	
GSM194255	
GSM194255	
GSM194256	
GSM194257	
GSM194258	
GSM194259	
GSM194260	
GSM194261	
GSM194262	
GSM194263	
GSM194264	
GSM194265	
GSM194266	
GSM194267	
GSM194268	
GSM194269	
GSM194270	
GSM194271	
GSM194272	
GSM194273	
GSM194274	
GSM194275	

Experiment 15 GSE 994-1 (Effects of cigarette smoke on the human airway epithelial cell transcriptome) between smokers and non-smokers.

Down-regulation				Up-regulation			
OR	Lower CI	Upper CI	*p*-value	OR	Lower CI	Upper CI	*p*-value
0.96	0.78	1.20	7.37E-01	0.97	0.76	1.23	8.12E-01

Case	Control
GSM15684	GSM15729
GSM15686	GSM15736
GSM15709	GSM15737

Case	Control
GSM15711	GSM15738
GSM15712	GSM15739
GSM15714	GSM15740
GSM15717	GSM15718
GSM15713	GSM15720
GSM15715	GSM15721
GSM15716	GSM15725
GSM15687	GSM15719
GSM15685	GSM15728
GSM15688	GSM15731
GSM15689	GSM15732
GSM15690	GSM15733
GSM15691	GSM15734
GSM15692	GSM15735
GSM15696	GSM15722
GSM15695	GSM15723
GSM15697	GSM15724
GSM15698	GSM15726
GSM15699	GSM15727
GSM15700	GSM15730
GSM15701	
GSM15702	
GSM15703	
,GSM15704	
GSM15705	
GSM15706	
GSM15707	
GSM15708	
GSM15710	
GSM15693	
GSM15694	

VITAE

SIRIPORN WANGSRI, D.D.S.

Date of birth	March 10, 1980
Place of birth	Mukdaharn, Thailand
Nationality	Thai
Education	Doctor of dental surgery degree from Khonkaen University, Thailand (2000-2005)
Work experience	
2005-2008	Lueamnat Hospital, Amnatcharoen, Thailand
2008-2009	Pathumratchawongsa Hospital, Amnatcharoen, Thailand