إقـــــرار

أنا الموقع أدناه مقدم الرسالة التي تحمل العنوان:

Genetic Polymorphisms in Lung Cancer Susceptibility between Palestinian Population (Gaza Strip)

أقر بأن ما اشتملت عليه هذه الرسالة إنما هو نتاج جهدي الخاص، باستثناء ما تمت الإشارة إليه حيثما ورد، وإن هذه الرسالة ككل أو أي جزء منها لم يقدم من قبل لنيل درجة أو لقب علمي أو بحثي لدى أي مؤسسة تعليمية أو بحثية أخرى.

DECLARATION

The work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degree or qualification

Student's name: اسم الطالب/ة: رامي عبد الرحيم سالم المصري

Signature: التوقيع:

Date: التاريخ: 14 / 12 / 2015

Islamic University of Gaza
Deanery of Higher Education
Faculty of Science
Department of Biotechnology

الجامعـــة الاسلاميـــة غـــزة
عمـــادة الدراســات العليـــا
كليـــة العلـــوم
قسم التكـــنولوجيـــا الحـيويـــة

Genetic Polymorphisms in Lung Cancer Susceptibility between Palestinian Population (Gaza Strip)

Prepared by:

Rami Al-Masri

Supervisors

Dr. Basim Ayesh, PhD Molecular Biochemistry

Al-Aqsa University, Faculty of Science, Biology Dept.

Dr. Abdalla Abed, PhD Human genetics

Islamic University, Faculty of Science, Biology Dept.

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Biotechnology

October 2015

بسم الله الرحمن الرحيم

الجامعة الإسلامية – غزة
The Islamic University - Gaza

مكتب نائب الرئيس للبحث العلمي والدراسات العليا هاتف داخلي: 1150

الرقم..... ج س غ/35/ Ref
التاريخ..... 2015/12/02 Date

نتيجة الحكم على أطروحة ماجستير

بناءً على موافقة شئون البحث العلمي والدراسات العليا بالجامعة الإسلامية بغزة على تشكيل لجنة الحكم على أطروحة الباحث/ **رامي عبدالرحيم سالم المصري** لنيل درجة الماجستير في **كلية** *العلوم* قسم **التكنولوجيا الحيوية** وموضوعها:

خطورة التنوع الجيني لدى مرضى السرطان في قطاع غزة

Genetic Polymorphisms in Lung Cancer Susceptibility between Palestinian Population Gaza Strip

وبعد المناقشة التي تمت اليوم الأربعاء 20 صفر 1437هـ، الموافق 2015/12/02م الساعة التاسعة صباحاً، اجتمعت لجنة الحكم على الأطروحة والمكونة من:

د. عبد الله عساف عابد	مشرفاً ورئيساً
د. باسم محمد عايش	مشرفاً
د. صائب حسين العويني	مناقشاً داخلياً
د. نبيل عبدالله العيلة	مناقشاً خارجياً

وبعد المداولة أوصت اللجنة بمنح الباحث درجة الماجستير في **كلية** *العلوم*/ قسم **التكنولوجيا الحيوية**.

واللجنة إذ تمنحه هذه الدرجة فإنها توصيه بتقوى الله ولزوم طاعته وأن يسخر علمه في خدمة دينه ووطنه

والله ولي التوفيق،،،

نائب الرئيس لشئون البحث العلمي والدراسات العليا

أ.د. عبدالرؤوف علي المناعمة

P.O. Box 108, Rimal, Gaza, Palestine fax: +970 (8) 286 0800 فاكس Tel: +970 (8) 286 0700 هاتف ص.ب. 108 الرمال، غزة، فلسطين
public@iugaza.edu.ps www.iugaza.edu.ps

﴿رَبَّنَا لَا تُزِغْ قُلُوبَنَا بَعْدَ إِذْ هَدَيْتَنَا وَهَبْ لَنَا

مِنْ لَدُنْكَ رَحْمَةً ۚ إِنَّكَ أَنتَ الْوَهَّابُ

آل عمران 8

Declaration

I hereby declare that submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree of the university or other degree of the university or other institute, except where due acknowledgment has been made in the text.

Signature	Name	Date
Rami Ms.	Rami Abd A-Raheem Al-Masri	October 2015

Genetic Polymorphisms in Lung Cancer Susceptibility between Palestinian Population (Gaza Strip)

Abstract:

Lung cancer is a leading cause of cancer-related deaths throughout the world. Recent GWAS and consecutive validation supported that SNPs in the 15q25, 5p15.33 and 6p22 regions showed significant association with lung cancer risk in multiple populations. Therefore, in this study six SNPs in three genomic regions were investigated because of their already established involvement in different populations. The relationship between these SNPs and smoking has been investigated.

Methods:

The SNPs rs16969968, rs1051730, rs402710, rs2736100, rs9295740 and rs4324798 were genotyped in a case-control study with 30 cancer cases and 60 age- and gender-matched cancer-free controls, all from Palestinian population of Gaza strip. Whole blood samples were collected from all subjects and relevant personal and clinical data was also collected from the participants and/or their medical record. Gene polymorphisms at the specified SNPs were determined by allele specific PCR and RFLP-PCR methodology.

Results:

The SNP rs16969968 G>A allele and rs1051730 C>T allele were found significantly related to lung cancer risk (*P-value* = 0.022, OR= 3.23; 0.001, OR=3.05 respectively). The allelic frequencies for the susceptibility alleles of rs16969968 and rs1051730 were higher in the cases than in the controls (36.7% vs 23.3% and 46.7 vs 22.5% respectively). Homo- and heterozygote nicotine consuming individuals have higher rate of cigarettes consumption in case of rs16969968 G>A (40.0±15.81 and 36.04±12.2 CPD compared to the wild type *P-Value*=0.006), and in case of rs1051730 C>T (39.17±15.62 and 33.00±10.36 CPD *P-Value*=0.015). The rs402710 has significant relation with lung cancer susceptibility (*P-Value* =0.015, OR= 3.00) between cases and controls. However, No statistically significant differences could be found in the distribution of rs402710 SNP between smoker control and smoker cases (*P*-value = 0.621). In

addition, rs402710 T allele can be considered as risk allele for NSCLC. The rs2736100 was also associated with risk of lung cancer (P=0.044, OR= 4.33) and the frequency of the risk allele between cases and controls was 35% vs 18.3% respectively. However, it didn't have correlation with smoking quantity (*P- value* = 0.269). Also, the rs9295740 distribution for the risk allele among case (30.0%) and control (15.0%) is not statistically significant (*P*-value = 0.075), but between the smoker cases and smoker controls is statistically significant (P-value = 0.016), so it may be risk factor for lung cancer between smokers. The allele frequency for rs4324798 was 20% in controls and 0.0% in cases, so it may have a role in protection from lung cancer.

Conclusion:

Our findings provide further evidence supporting the genetic variants in 15q25, 5p15.33 and 6p22 regions associated with the risk of lung cancer. Smoking and lung infections are important triggers for lung cancer. Further studies with larger cohorts of Palestinian lung cancer patients are recommended to extend the outcomes of this study.

ملخص الدراسة:

سرطان الرئة هو السبب الرئيسي للوفيات المرتبطة بالسرطان في جميع أنحاء العالم. وقد عملت دراسات رابطة الجينوم على التحقق من صحة أن مناطق **الكروموسومات الثلاث** 15q25، 5p15.33و 6p22 تظهر ارتباطا كبيرا مع سرطان الرئة في عدة مجتمعات مختلفة. لذلك تم تصميم هذه الدراسة للتحقق من ستة اشكال من النيوكلوتايد في ثلاثة مناطق من الجينوم البشري، حيث انه تم اثبات علاقة بعضها بالتدخين وسرطان الرئة.

الطريقة:

تم تحديد الأنماط الجينية لستة النيوكلوتايد rs16969968، rs1051730، rs402710، rs2736100، rs9295740 و rs4324798 في دراسة الحالات والشواهد مع 30 عينة من مرضى سرطان الرئة و 60 عينة خالية من سرطان الرئة كعينة ضابطة للدراسة بين فلسطينيين من قطاع غزة . حيث تم جمع عينات الدم من المرضى بالاضافة الى جمع البيانات الشخصية والسريرية ذات الصلة من المرضى أو من السجلات الطبية الخاصة بهم كما تم تحديد الأنماط الجينية عن طريق RFLP- أو AS-PCR.

النتائج:

وجدنا ان الأنماط الجينية ل rs16969968 و rs1051730 (15q25) لها علاقة كبيرة مع خطر الاصابة بسرطان الرئة (قيمة احتمال= 0.022 ومعدل خطورة= 3.23؛ قيمة احتمال= 0.001 ومعدل خطورة= 3.05) على التوالي، علاوة على ذلك، وجدت ان نسبة النمط الجيني الخطر كانت مرتفعة بين مرضى سرطان الرئة (36.7% و 46.7%) عنها في العينات الضابطة (23.3% و 22.5%) بالترتب، كما ظهر ان الأنماط الجينية **الخطرة** لديها أعلى معدل استهلاك للسجائر 15.81±40.0 (قيمة احتمال =0.006)، 15.62 ± 39.17 (قيمة احتمال =0.015) على التوالي.

اما بالنسبة ل rs402710 فلديها علاقة مع زيادة خطر الاصابة بسرطان الرئة (قيمة احتمال= 0.015 و معدل خطورة= 3.00)، ولكن لم يتم ايجاد علاقة لها مع التدخين (قيمة احتمال= 0.621)، كما لوحظ ان النمط الجيني السئ يرتبط بنوع معين من سرطان الرئة NSCLC. كما وجد ان rs2736100 لديها علاقة ذات دلالة

إحصائية انها تزيد من خطر الاصابة بسرطان الرئة (قيمة احتمال= 0.044 و معدل خطورة= 4.33) ولا توجد لديها علاقة مع كمية التدخين (قيمة احتمال= 0.269)، كما أن نسبة النمط الجيني الخطر بين الحالات المرضية عالية (35%) مقارنة بالعينات الضابطة (18.3%) .أما النمط الجيني ل rs9295740 فارتباطه مع خطر الاصابة بسرطان الرئة غير مقبول (قيمة احتمال= 0.075) مع العلم ان نسبة النمط الجيني الخطير بين المرضى (30%) أعله منه بين العينات الضابطة (15%)، لكن يعتبر النمط الجيني الخطير ل rs9295740 بين المدخنين عامل لزيادة خطر الاصابة بسرطان الرئة (قيم احتمال= 0.016). بينما النمط الجيني لrs4324798 كان على النقيض، حيث كانت نسبة تواجده في العينات الضابطة (20%) مقابل ولا حالة من الحالات المرضية، لذلك انه من المتوقع أن يكون له دور في الحماية من سرطان الرئة.

الاستنتاج:

تظهر النتائج التي توصلنا إليها مزيدا من الأدلة المؤيدة لارتباط الأنماط الجينية في المناطق 15q25، 5p15.33 و 6p22 بخطر الإصابة بسرطان الرئة، إلا أن rs4324798 يجب أن تكون قيد المزيد من الدراسات لتأكيد دورها في خطر الاصابة بسرطان الرئة بين الفلسطينيين.

Dedication

To my beloved parents

To my wife Aseel

To my sons
Omar, Kareem & Maya

Acknowledgment

Foremost, I would like to express my deepest gratitude to my supervisor **Dr. Basim Ayesh,** for his mentorship and continuous support. I am extremely grateful for his patience, kindness, and willingness to share his knowledge. His enthusiasm for science has greatly inspired and encouraged me. I consider myself extremely fortunate to have his as my mentor.

I am also grateful for my supervisor **Dr. Abdalla Abed**. He has been generous with his knowledge, experience and precious time, and been a source of support and encouragement.

I would like to thank the medical technologist's staff in laboratory departments of **European Gaza Hospital, Al-Shifa Hospital and Molecular Biology department of central laboratory** about the great effort for help me in collection of the specimens and data.

I would also like to thank my friends and classmates both in the University and in Jordan for their friendship, help and supports.

Lastly but most importantly, my great love and respect to **my wife** who has supported me to make this work possible and **my parents** for all the support and encouragement they provided me.

Great thanks for all people who participate and help me in the success of this study**.**

List of Contents

Contents	Page
Declaration……	III
Abstract……	IV
Arabic Abstract……..	VI
Dedication…….	VIII
Acknowledgment……	IX
List of Contents……..	X
List of Tables……..	XIII
List of Figures……	XV
Abbreviations……..	XVI
Genetic Terms Glossary	XVII
Chapter One- Introduction	
1.1 Overview……..	1
1.2 Objectives of the study……..	2
1.3 Significance of the study……	3
Chapter Two- Literature Review	
2.1 Lung Cancer……..	4
2.1.1 Types of lung cancer…….	4
2.1.2 Epidemiology of lung cancer……	5
2.1.3 Risk Factors……	5
2.1.3.1 Smoking……	7
2.1.3.2 Age……	7
2.1.4 Other factors ……..	7
2.2 Genetic polymorphisms……..	8
2.3 Genetic susceptibility……..	8
2.3.1 Genetic Susceptibility to Lung Cancer…….	8
2.4 Molecular Biology of Lung Cancer Susceptibility Regions…….	9
2.4.1 Region 15q25…….	10

2.4.1.1 rs16969968 (c.1192G>A)…….. 11
2.4.1.2 rs1051730 (c.645C>T)…… 11
2.4.2 Region 5p15…… 12
2.4.2.1 rs402710 (g.1320607C>T, c.1532+9G>A)…….. 14
2.4.2.2 rs2736100 (c.1574-3777G>T)…… 14
2.4.3 Region 6p21…… 15
2.4.3.1 rs4324798 (g.79518G>A)……... 16
2.4.3.2 rs9295740 (g.27689502G>A)…… 16

Chapter Three- Materials and Methods

3.1 Materials…… 18
3.1.1 Reagents……... 18
3.2 Instruments…… 18
3.3 Study Ethics…….. 19
3.4 Samples…… 19
3.5 Settings and place of work…… 20
3.6 Patients data…… 20
3.7 DNA Isolation…… 20
3.8 Molecular investigation…… 20
3.8.1 Primers selections…….. 20
3.8.2 Allele Specific PCR…… 21
3.8.2.1 AS-PCR for rs16969968…….. 21
3.8.2.2 AS-PCR for rs1051730…… 22
3.8.2.3 AS-PCR for rs402710…….. 22
3.8.2.4 AS-PCR for rs2736100…… 23
3.8.2.5 AS-PCR for rs4324798…… 23
3.8.3 Restriction Enzyme Length Polymorphism (RFLP)…… 23
3.8.3.1 rs9295740-RFLP…….. 23
3.8.4 β-Globin Gene…….. 24
3.9 Quality Control…… 24
3.10 Agarose gel electrophoresis…… 24

3.11 Biostatistics/ Data analysis…….. 25
3.12 Limitations of the study…… 25

Chapter Four- Results and Discussion

4.1 Study population description…… 26
4.1.1 Gender distribution…….. 26
4.1.2 Living area…… 27
4.1.3 Age distribution…… 27
4.1.4 Case distribution according to place of sample collection…… 28
4.1.5 Distribution of cases by type of lung cancer…… 28
4.1.6 History of lung infectious diseases…… 29
4.1.7 Smoking status…… 30
4.1.8 Cigarettes per day…… 30
4.1.9 Duration of smoking…… 31
4.1.10 Qualitative detection of SNPs…….. 31
4.2 SNPs Genotyping and Relations…….. 33
4.2.1 rs16969968, G>A…… 33
4.2.2 rs1051730, C>T…… 37
4.2.3 rs402710, C>T…….. 41
4.2.4 rs2736100, G>T……... 45
4.2.5 rs9295740, G>A……... 49
4.2.6 rs4324798, G>A……... 52

Chapter Five - Conclusions and Recommendations

5.1 Conclusion…… 56
5.2 Recommendation…… 57

Chapter Six – References 59

Appendixes 72

LIST OF TABLES

Table Title	Page No.
Table 3.1 Reagent used in the study…….	18
Table 3.2 The Instruments used in the study……..	18
Table 3.3: Sequence of primers for each SNP…….	21
Table4.1: Age distribution of cases and controls…….	27
Table 4.2: distribution by history of lung infectious diseases…….	29
Table4.3: Number of cigarettes per day……...	30
Table4.4: Duration of smoking……...	31
Table 4.5: The distribution of alleles among case and control……..	34
Table 4.6: Variant genotype with lung cancer risk……..	34
Table 4.7: Allele distribution among smoker cases and smoker controls…….	35
Table 4.8: SNPs in nAChR gene vs number of cigarettes per day (CPD)……..	37
Table 4.9: Distribution of the rs16969968 between patients depends on lung cancer type……..	37
Table 4.10: The distribution of alleles among case and control…….	38
Table 4.11: Variant genotype with lung cancer risk…….	39
Table 4.12: Allele distribution among smoker cases and smoker controls…….	39
Table 4.13: SNPs in nAChR gene vs number of cigarettes per day (CPD)	41
Table 4.14: Distribution of the rs1051730 between patients depends on lung cancer type……..	41
Table 4.15: The distribution of alleles among case and control…….	42
Table4.16: Variant genotype with lung cancer risk……..	43
Table 4.17: Allele genotypes and its correlations…….	43
Table 4.18: rs402710 vs number of cigarettes per day (CPD)……..	44
Table 4.19: rs402710 vs duration of smoking……..	44

Table 4.20: Distribution of the rs402710 between patients depends on lung cancer type…….. 45
Table 4.21: The distribution of alleles among case and control……. 46
Table 4.22: Variant genotype with lung cancer risk…… 47
Table 4. 23: Allele distribution among smoker cases and smoker controls…… 47
Table 4.24: rs2736100 vs number of cigarettes per day (CPD)……. 48
Table 4.25: Distribution of the rs2736100 between patients depends on lung cancer type…….. 69
Table 4.26: The distribution of alleles among case and control……. 50
Table 4.27: Variant genotype with lung cancer risk…… 50
Table 4. 28: Allele distribution among smoker cases and smoker controls…… 51
Table 4.29: Allele distribution among smoker and non-smoker controls…… 52
Table 4.30: rs9295740 vs number of cigarettes per day (CPD)……. 52
Table 4.31: rs9295740 vs years of smoking between smoker cases and controls…… 52
Table 4.32: Distribution of the rs9295740 between patients depends on lung cancer type…….. 53
Table 4.33: The distribution of alleles among case and control……. 54
Table 4.34: Variant genotype with lung cancer risk…… 54
Table 4. 35: Allele distribution among smoker cases and smoker controls…… 55
Table 4.36: Distribution of rs4324798 among controls……... 55
Table 4.37: rs4324798 vs number of cigarettes per day (CPD)……. 56
Table 4.38: rs4324798 vs years of smoking between smoker cases and controls…… 56

LIST OF FIGURES

Figure Title	Page No.
Figure 4.1 Distribution of cases by gender……..	26
Figure 4.2 Distribution of cases by living area…….	27
Figure 4.3 Distribution of lung cancer samples by place of collection.	28
Figure 4.4 Distribution of cases by type of lung cancer…….	28
Figure 4.5 Distribution of cases by status of smoking…….	30
Figure 4.6 The electrophoresis profiles for representative amplifications…….	32

ABBREVIATIONS

Abbreviation	Term description
IARC	International agency for research on cancer
PHIC	Palestinian health information center
SCLC	Small cell lung cancer
NSCLC	Non-small cell lung cancer
NNK	Nicotine-derived nitrosamine ketone
DNA	Deoxyribonucleic acid
SNP	Single Nucleotide Polymorphism
UTR	Untranslated region
TFBS	Transcription factor binding site
CNV	Copy number variation
APC	Adenomatous polyposis coli
BRCA	Breast cancer
LD	Linkage disequilibrium
GWAS	Genome wide association studies
TP53	Tumor protein p53 gene
RB1	Retinoblastoma 1 gene
GSTM1	Glutathione S transferase M1
CHEK2	Checkpoint kinase 2
nAChR	Nicotinic acetylcholine receptor
CHRNB4	Cholinergic Receptor, Nicotinic, Beta 4 gene
CHRNA3	Cholinergic Receptor, Nicotinic, alpha 3 gene
CHRNA5	Cholinergic Receptor, Nicotinic, alpha 5 gene
CLPTM1L	Cleft lip and palate transmembrane protein 1-like protein gene
TERT	Telomerase reverse transcriptase gene
MNS16A	Minisatellites
MSH5	MutS protein homolog 5 gene

BAT3	HLA-B associated transcript 3 gene
GPC5	Glypican 5 gene
CPD	Cigarettes per day
CRR9	Cisplatin resistance related gene 9
p	Short arm for chromosome
q	Long arm for chromosome
TRNAL16	Transfer RNA leucine 16
µM	Micro Molar
PCR	Polymerase Chain Reaction
ng	Nano gram
dNTPs	Dinuclotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
U/L	Unite/Litter
USA	United Stat Of America
MOH	Ministry of health
ALFRED	ALlele FREquency Database
NCBI	National Center for Biotechnology Information
ACS	American Cancer Society
WHO	World Health Organization
GWAS	Genome Wide Association Studies
EGH	European Gaza Hospital
WASP	Web-based Allele-Specific PCR assay
AS-PCR	Allele Specific Polymerase Chain Reaction
RFLP	Restriction Enzyme Length Polymorphism
ORs	Odds ratios
$¡^{2}$	Chi-squared test

GENETIC TERMS GLOSSARY

Allele	Alternate forms of a gene or a specific variant/base at a particular locus in the genome that differ in DNA sequence.
Candidate gene	A gene believed to be involved in a complex trait or disease based on known biological and/or physiological properties of its products or its location near a region of association or linkage.
Complex traits	A trait that is influenced by multiple genes, environmental factors and the interaction between them.
Copy number variant (CNV)	A form of structural variation of the DNA where stretches of genomic sequence (1kb 3Mb in size) are deleted or duplicated in varying numbers.
Deoxyribonucleic acid (DNA)	A double helix molecule consisting of 4 bases; Adenine (A), Thymine (T), Guanine (G) and Cytosine (C), together, forming the molecular basis of the genome.
Gene	Traditionally, the basic physical unit of heredity; a sequence of DNA that gives the coding instructions for the synthesis of RNA. The human genome contains approximately 25,000 genes distributed on 23 pairs of chromosomes. New research from the ENCODE project show that about 75% of the genome is transcribed at some point in some cells, and that genes are highly interlaced with overlapping transcripts that are synthesized from both DNA strands.
Genetic code	The set of rules by which information encoded in genetic material (DNA or mRNA sequences) is translated into amino acid sequences. A specific sequence of three nucleotides, a codon, determines the amino acid.
Genetic variation	Variation in alleles of genes, both within and among populations. Provides the “raw material” for natural selection
Genome	The total of an individual organism’s entire genetic material.
Genome wide association studies	The study of genetic variation across the entire genome aimed at identifying genetic variation associated with a complex disease or trait.
Genomics	Genomics is a discipline in genetics concerning the study of the genomes of organisms. Traditionally genomics concerns everything that has to do with DNA. A broader definition is used by the United States Environmental Protection Agency, to also include mRNA and proteins.
Genotype	The combination of alleles on corresponding loci in the two copies of the chromosomes. When two sequence alternatives exist at a given locus, e.g. A and G 3 different genotypes are possible, AA and GG when the allele is

	identical on each chromosome and AG when the allele differs.
Haplotype	A combination of alleles at adjacent loci on the chromosome that are transmitted together.
HapMap	A genome- wide database of patterns of common human genetic sequence variation among multiple ancestral population samples.
Heritability	The proportion of observable differences between individuals that is due to genetic differences.
Hardy Weinberg Equilibrium	The population distribution of 2 alleles (with frequencies p and q) such that the distribution is stable from generation to generation. Genotypes occur at frequencies of p^2, 2pq and q^2 for the major allele homozygote, heterozygote and minor allele homozygote.
Linkage disequilibrium (LD)	The non-random association of allele at two or more loci. Occurs when two or more loci on a chromosome have reduced recombination between them because of their physical proximity to each other. LD describes the extent to which a variant at one locus predicts the variant at another locus.
Locus	Any given specific site in a genome. Often used to describe a particular site where sequence or functional alternatives exist.
Minor allele	The allele with the lowest frequency of a biallelic polymorphisms
Minor allele frequency	The frequency of the least common of 2 alleles in a population.
Mutation	A change in the genomic sequence of DNA as a result of DNA damage, replication error, incomplete repair or other intrinsic events.
Phenotype	A phenotype is the composite of an organism´s observable characteristics or traits and result from the expression of the organism's genes as well as the influence of environmental factors and the interactions between the two.
Single Nucleotide Polymorphism (SNP)	A type of genetic variation where, at a specific locus in the genome two sequence alternatives exists and where the least common alternative is found in minimum 1% of the population in question.
Penetrance	In genetics, is the proportion of individuals carrying a particular variant of a gene (allele or genotype) that also expresses an associated trait (phenotype). In medical genetics, the penetrance of a disease-causing mutation is the proportion of individuals with the mutation who exhibit

	clinical symptoms**.**
DNA adduct	A central tenet of chemical carcinogenesis is that the covalent binding of carcinogens to DNA is causally related to tumorigenesis.

Chapter One

Introduction

1.1 Overview

Cancer is considered one of the most important health problems in both developing and developed countries for its high incidence, cost and associated mortality. Cancer is the third leading cause of death among Palestinians (About 9%) after other accidents (35.5%) and heart diseases (17.4%), and it is one of the major causes of morbidity among Palestinian population (PHIC, 2014).

Lung cancer is a complex and heterogeneous disease dependent on many genes, environmental and lifestyle factors. It is also one of the leading causes of death worldwide. Approximately 1.59 million people die from lung cancer per year; about 750,000 in a year die from liver cancer, the second cause of cancer deaths worldwide (WHO, 2012). In Palestine (West bank), lung cancer caused more than 187 deaths in 2014; of those, 147 were men, and 40 were women (PHIC, 2014). In Palestine, lung cancer is the most common cancer diagnosed in men, and the second most common cancer diagnosed in women.

Lung cancer is divided into two main categories: Small Cell Lung Cancer (SCLC) accounts for about 10-15% of all lung cancers and Non-Small Cell Lung Cancer (NSCLC) accounts for 85-90% of all lung cancers. NSCLC is generally slower growing than SCLC and is divided into subgroups based on the cells that make up the tumor, including adenocarcinoma, squamous cell carcinoma, and poorly differentiated or large cell carcinoma (ACS, 2015).

Lung cancer is a multifaceted disease with a number of possible risk factors. Tobacco smoke is the most well-known risk factor responsible for 80% to 90% of lung cancers (Parkin DM. 2011, CDC, 2015). This relationship has been definitively established as the main cause of lung cancer through a wide variety of case control and cohort studies throughout the world. Relative risk estimates for lung cancer risk among cigarette smokers range from 20-40 times the never smoking (Shopland et al., 1991, Crispo A. et al., 2004).

Furthermore, both intensity, in terms of number of cigarettes smoked, and duration of smoking appear to influence the risk (Knoke J. et al., 2004).

Tobacco smoke is a complex mixture of about 5300 organic compounds that have been identified (Rodgman & Perfetti, 2009). Among these, more than 70 are considered carcinogenic (Hecht S. and Samet M. 2007, Stampfil R. and Anderson P. 2009, Stellman D. and Djordjevic V. 2009). The carcinogenicity of tobacco smoke may result from direct genotoxic damage to DNA; activation of growth and proliferation and suppression of apoptosis pathways and chronic inflammation with subsequent cellular proliferative response to inflammatory injury (Hecht S. and Samet M. 2007, Stampfil R. and Anderson P. 2009, Stellman D. and Djordjevic V. 2009, Chen RJ et. al. 2011).

Cigarette smoke carcinogens lead to DNA adducts formation and induce oxidative damage to DNA, resulting in genome instability that represent the first steps of carcinogenesis (Hecht, 2003). Excessive DNA damage as a result of multiple mutations may fail to be repaired and lead to loss of normal controls on cell growth unless the cell undergoes apoptosis (Wu H. et al., 2004).

The fact that 10-20% of lung cancer patients are never smoking (CDC. 2005, Thun MJ. 2006, Jemal A. 2008), and that a considerable number of smokers never get lung cancer suggests that the underlying individual's genetic makeup may be responsible for this variability. The search for genetic factors that are implicated in lung cancer pathology have highlighted the possible role of a number of loci (15q25, 5p15, and 6p21). These include rs16969968, rs1051730, rs402710, rs2736100, rs4324798 and rs9295740.

1.2 Objective:

This research aimed at studying the interplay between smoking and genetic susceptibility to lung cancer in Gaza strip-Palestinian, with a number of specific aims:

1. To determine the frequency of SNPs previously known to have a role in increasing the genetic risk for lung cancer rs16969968 (c.1192G>A),

rs1051730 (c.645C>T), rs402710 (c.1532+9G>A), rs2736100 (c.1574-3777G>T), rs4324798 and rs9295740).

2. To evaluate the role of smoking and other non-genetic risk factors such as gender, age and family history in susceptibility to lung cancer.
3. To evaluate the combination effect of both genetic and/or non-genetic factors in increasing the susceptibility for lung cancer.

1.3 Significance:

Most of lung cancer patients go undiagnosed as the symptoms associated with lung cancer can be rather subtle, and many patients remain at a silent or latent stage. As lung cancer is associated with increased morbidity and mortality, it puts a severe health and economic burden on patients, their families, and society. Improved diagnosis and prediction of lung cancer and early intervention would alleviate, or reverse, these negative effects. Identification of a part of the genetic risk factors for lung cancer may help in identifying individuals at high risk for lung cancer before the disease manifests.

The genetic susceptibility for lung cancer may be exacerbated when other risk factors coexist (e.g. smoking, family history, diet, alcohol and occupational exposure have played a critical role in the tumorigenesis in lung (CDC, 2013). People with increased risk may be advised to modify their life style in a way that reduces the incidence of cancer.

To the best of our knowledge, there is no previous data describing the frequency of risk alleles for lung cancer among the Palestinian population, neither this issue was tackled what so ever in our population.

Chapter Two

Literature Review

2.1 Lung Cancer

Lung cancer is characterized by uncontrolled proliferation of the lung cells as a result of genetic damage. Like all cancers, lung cancer cells have the ability to invade neighboring tissues and spread or metastasize to distant anatomical parts of the body. If left untreated, lung cancer eventually causes death.

2.1.1 Types of lung cancer

Lung cancer is classified into two main histological groups: Small Cell Lung Cancer (SCLC) which characterized by its growing more rapidly and spreads to other parts of the body earlier than non-small cell lung cancer. It is also more responsive to chemotherapy (CancerCare 2015). and Non-Small Cell Lung Cancer (NSCLC) (Cooper S. & Spiro SG, 2006, Herbst RS et. al., 2008). The NSCLCs generally include adenocarcinomas, squamous cell carcinomas, and large cell carcinomas (ACS, 9.2015).

The adenocarcinoma is the most common form of lung cancer that forms in mucus-secreting glands throughout the body; also it can spread without destroying other tissues. While squamous cell carcinoma (also called epidermoid carcinoma) forms in the lining of the bronchial tubes and its characterized by thin, flat cells that line the passages of the respiratory tract, then large cell carcinomas which refer to non-small cell lung cancers that are neither adenocarcinomas nor epidermoid cancers its characterized by faster growing form of NSCLC (Lung Cancer Alliance. 2015).

The majority of tumors are NSCLC, with SCLC comprising only about 10-30% (Janssen L. & Coebergh W., 2003). Among NSCLC, squamous cell carcinomas were the most common in the mid to late 20th century (Youlden R. et al., 2008).

2.1.2 Staging of Lung Cancer

Staging lung cancer is based on whether the cancer is local or has spread from the lungs to the lymph nodes or other organs. Because the lungs are large, tumors can grow in them for a long time before they are found. Even when symptoms—such as coughing and fatigue—do occur, people think they are due to other causes. For this reason, early-stage lung cancer (stages I and II) is difficult to detect. Most people with lung cancer are diagnosed at stages III and IV (CancerCare 2015).

Non-small cell lung cancer is divided to four stages. In the first stage, the cancer may be present in the underlying lung tissues, but the lymph nodes remain unaffected. While in the second stage, the cancer may have spread to nearby lymph nodes or into the chest wall. And the third stage is divided for two subtypes, the first is characterized by spreading of cancer only to lymph nodes on the same side of the chest where the cancer started, and the second subtype is distinguished by the spreading of cancer to the lymph nodes on the opposite side of the chest, or above the collar bone. However in the fourth stage, the cancer has metastasized throughout the body and may now affect the liver, bones or brain (CancerCenter 2015, Lung Cancer Alliance. 2015).

Small cell lung cancer is divided to two stages, the first stage is limited stage which characterized by the cancer is in one lung, sometimes including nearby lymph nodes, while the second stage is extinctive stage, which characterized by the spreading of cancer to the other lung, the fluid around the lung (the pleura) or to other organs in the body (CancerCare 2015, CancerCenter 2015).

2.1.3 Epidemiology of lung cancer

Cancer in general can be the second cause for death in Gaza strip, after the cardiovascular diseases, about 1595 cases in Gaza strip and 2295 cases in west bank are registered in Palestinian ministry of health in 2014, and the incidence of the cancer in Gaza strip was 86.8/100,000 people, and 82.2/100,000 in west bank (PHIC, 2014).

Worldwide, lung cancer is the most common form of cancer (WCRF International, 2012); accounting for 1.83 million new cases annually representing 13% of all new cancers. It is also the most common cause of death from cancer, with 1.61 million deaths (WHO, 2012).

No recent data was published regarding the incidence of lung cancer in Gaza strip. However by extrapolating the data on the west bank published by the Palestinian Health Information Center (PHIC), one can expect the incidence of lung cancer to be about 9.8% of all cancers (PHIC, 2014). Furthermore, the mortality of lung cancer was the highest between the other cancers with 19.6%, followed by the colon cancer with 13%. The report of lung cancer cases in Gaza was issued by the Palestinian Health Information Center in 2009; in which lung cancer occupied the first leading cause of death, with 13.4% and 21.3% in the year 2008, 2009 respectively (PHIC, 2009). The mortality of lung cancer between males and females in year 2009 was 27.7% with mortality rate 9.2/100,000 males and 13.9% with mortality rate 4.6/100,000 females (PHIC, 2009).

The incidence and mortality of lung cancer have been significantly and constantly increasing over the past two decades in Palestine. The age-standardized incidence and mortality rate for lung cancer of Palestinian population was 10.8/100,000 and 10.0/100,000 for men and 1.8/100,000 and 1.9/100.000 for women in 2008 (IARC, 2008). According to the Palestinian Health Information Center in west bank, number of deaths from lung cancer between cancer cases in 2014 was 19.8%, the second malignant tumor was colon cancer with 13%, following by breast cancer with 10.7% (PHIC, 2014).

Lung cancer remains highly lethal with very low survival with only 17.4% of US lung cancer patients surviving 5 years after diagnosis with best condition of treatment within period from 2005-2011 (ACS, 2015), while in Palestine is approximately 7.7% (IARC, 2008).

2.1.4 Risk Factors

Lung cancer develops when the cells of the lung undergo a sustained genetic damage. Several different chemicals and environmental factors that are capable of causing the genetic damage can lead to lung cancer.

2.1.4.1 Smoking

The most well-known risk factor, and perhaps the most prevalent worldwide, is tobacco smoke. About 80% to 90% of all lung cancer cases occur among people who are either current or former tobacco smokers (Shopland et al., 1991, IRAC 2004, Thun MJ 2005, Freedman LS et. al. 2006). The relationship between smoking and lung cancer is caused by the carcinogens present in tobacco smoke. The risk of developing lung cancer from smoking is influenced by length and intensity of smoking history (Peto R. 1986, Flanders WD. et. al. 2003, Doll R. et. al. 2005, Vandenbroucke J. 2009). On average, a lifetime smoker has a 20-fold increase in the risk of developing lung cancer compared with a lifetime non-smoker (Alberg AJ. and Samet JM. 2003, Doll R et. al. 2005 and Pirie K et. al. 2013).

Cigarette smoke is a mixture of approximately 5300 different chemicals (Rodgman & Perfetti, 2009), more than 70 of which are established carcinogens in both *in vitro* and *in vivo* models such as nitrosamines, polycyclic aromatic hydrocarbons, acroleins and aryl amines (Hecht S. and Samet M. 2007, Stampfil R. and Anderson P. 2009, Stellman D. and Djordjevic V. 2009).

2.1.4.2 Age

The individual's age is an established risk factor for a number of diseases including cancer. Older individuals are probably exposed to more endogenous and exogenous genotoxic influences that ultimately lead to genetic damage accumulation over time. Accordingly, the probability of accumulating enough genetic damage can lead to cancer. In addition, the immune system works less effectively as someone gets older. The average age of newly diagnosed lung cancer patients is around 60 years of age (Cancer Stats. 2007), also in Palestinian population the rate is about 60 years of age (PHIC. 2011).

2.1.5 Other factors

While tobacco smoke is the leading cause of lung cancer, it is noteworthy mentioning that people who have never smoked can develop lung cancer. Furthermore, it is also true that most smokers long-live without lung cancer. This can be in part explained by variable genetic susceptibility to tobacco carcinogens (Amos C. et al., 1999). In addition, past lung infections, and exposure to environmental factors such as radon, radiation and asbestos may play a role in the risk of lung cancer (ACS. 2014).

2.2 Genetic polymorphisms

Genetic polymorphism is an established DNA sequence variation occurring in 1% or more of the population not necessarily resulting in a phenotypic chance. Single nucleotide polymorphisms (SNP) occur about once every 1000 base pairs in the human genome (International HapMap Consortium, 2007). SNPs in protein encoding genes can influence a phenotype either by changing the function or quantity of the encoded protein (Buckland R. 2006). Moreover SNPs can be found in non-coding sequences such as the gene regulatory sequences (promoters, enhancers, and silencers); in introns; within the 5' and 3'UTRs and in intergenic regions (Buckland R. 2006, Chorley BN. et. al., 2008).

SNPs in the coding region are classified into two types: synonymous and nonsynonymous. Synonymous SNPs do not affect the protein sequence while nonsynonymous SNPs change the amino acid sequence of the protein. The nonsynonymous SNPs in turn are divided into two forms: missense and nonsense (Strachan T. and Read A. 1999).

2.3 Genetic susceptibility

The genetic susceptibility for a particular disease (sometimes also called genetic predisposition) is an increased probability of developing that disease due to the individual's genetic makeup. This susceptibility frequently results from specific genetic polymorphisms found in the coding regions of metabolizing enzymes or sometimes in intergenic sequences. These polymorphisms play a role in the development of a disease but do not directly

cause it. Cancer susceptibility genes have been divided into three different classes: rare high-penetrance genes (low frequency), rare intermediate/moderate-penetrance genes (low frequency), and common low-penetrance genes (high frequency) (Stratton and Rahman, 2008).

Variation in the personal risk of a disease outcome may be due to both differing levels of exposure to environmental factors, with subsequent epigenetic effects on gene expression, and to genetic variation among individuals (Christiani C., 2006).

Family history is one of the strongest predisposing factors for cancer. For example around 5–10% of all cancer cases are of the familial type (ACS. 2014).

Genetic susceptibility can be identified by genome-wide linkage analysis, genome-wide association studies (GWAS) and others techniques (Wang Y. et. al., 2008, McKay D. et. al., 2008, Turnbull & Rahman 2008)

GWAS are cornerstone in studying cancer risk as well as clinical outcome and treatment response. It requires some essential prerequisites be fulfilled, such as a sufficient number of cases and controls, approved statistical significant levels and stratification for confounders (Miyagawa et al., 2008; Ziegler et al., 2008). So far, 315 cancer phenotypes have been subjected to GWAS, with a large number of risk alleles identified (GWAS. 2015).

2.3.1 Genetic Susceptibility to Lung Cancer

The genetic variation can contribute to the inter-individual differences in susceptibility to lung cancer in two aspects. Firstly, common SNPs modify the activity of enzymes that metabolize environmental agents such as chemicals found in cigarettes smoke, and maintain cell cycle control and immune function (Rothman et al., 2001). Second, inter-individual differences in susceptibility may result from the ability of the cell to repair the genetic damage from tobacco carcinogenic exposures and the cellular ability to remove adducts (Wei and Spitz, 1997).

Genetic susceptibility to lung cancer had been tested in a different number of epidemiologic studies. The later have identified three low-penetrance loci, 15q25.1, 5p15.33 and 6p21, associated with lung cancer, and accounting for 7% of the familial risk (Varghese and Easton, 2010). A common locus variant near rs1051730 and rs16969968 on chromosome 15q contains 3 genes encoding subunits of the nicotinic acetylcholine receptor (*nAChR*), *CHRNA3, CHRNA5,* and *CHRNB4* (Amos CI. et. al., 2008, Hung J. et. al., 2008, Thorgeirsson E. et. al., 2008, Shiraishi et al. 2009, Lips et al. 2010, Timofeeva et al. 2012).

The chromosomal region 5p15.33 containing the cleft lip and palate transmembrane protein 1-like protein gene (*CLPTM1L*) and the telomerase reverse transcriptase (*TERT*) gene, which were found to be associated with lung cancer, especially with adenocarcinoma (Okazaki et al. 2014). The identified SNPs in this region rs402710 and rs2736100 were found to modulate lung cancer risk (McKay et al. 2008, Jin et al. 2009, Landi et al. 2009, Hsiung et al. 2010, Ito et al. 2012).

GWAS also demonstrated that 6p21 is a new lung cancer susceptibility locus in populations of European descent (Wang Y. et. al., 2008, Zienolddiny S. et. al., 2009). This relationship has identified the SNPs rs4324798 and rs9295740 as risk alleles for lung cancer in various ethnic groups with varying degrees (TanTai J. et. al. 2013).

2.4 Molecular Biology of Lung Cancer Susceptibility Regions

As previously indicated, the three human genomic regions at chromosomes 15q25, 5p15, and 6p21 have been identified to be associated with susceptibility to lung cancer. Genes that regulate the acetylcholine nicotinic receptors, the *BAT3-MSH5* gene locus and the *TERT-CLPTM1L* locus were mapped to theses genomic locations (Wang et al., 2010, Truong et al. 2010, Wang Y. et. al., 2008, McKay D. et. al., 2008). Additionally, the 13q31.3 locus mapping to the *GPC5* gene was found to associate with lung cancer in never smokers (Li Y. et. al., 2010). Additional risk loci at 3q28, 13q12.12 and 22q12.2 were also identified in a study conducted on Han

Chinese subjects (Hu Z. et. al., 2008). The following sections discuss some of these loci in more details.

2.4.1 Region 15q25.

The region at 15q24–25.1 contains the nicotinic acetylcholine receptors (*nAChRs*) which is a member of a superfamily of ligand-gated ion channels that mediate fast signal transmission at synapses. The *nAChRs* are hetero-pentamers of homologous subunits encoded by separate genes and have different primary structures (muscle and neuronal forms). There are several subtypes of neuronal *nAChRs* that vary based on which homologous subunits are arranged around the central channel. The subunits having a pair of adjacent cysteines at the presumed acetylcholine binding site are classified as alpha-subunits (*like muscle alpha-1*), while subunits lacking these cysteine residues are classified as beta-subunits (Groot Kormelink and Luyten, 1997).

Three genes for nicotinic acetylcholine receptor subunit, namely *CHRNA5*, *CHRNA3*, and *CHRNB4*), were found associated with the risk of lung cancer in two independently conducted GWASs. The studies gave consistent results for associations between two SNPs at this locus (rs16969968 in *CHRNA5* gene, and rs1051730 in *CHRNA3*) and the risk of lung cancer (Amos et al., 2008; Hung et al., 2008). Both SNPs are in strong linkage disequilibrium (Amos et al., 2008; Hung et al., 2008).

2.4.1.1 rs16969968 (c.1192G>A)

The SNP rs16969968 results in the substitution of aspartate for asparagine at the highly conserved codon 398 (D398N) of the CHRNA5 protein, located in the central part of the second intracellular loop of the receptor. The SNP rs16969968 is in strong linkage disequilibrium with a SNP in the *CHRNA3* gene (Hung et al., 2008; Thorgeirsson et al., 2008). A significant association between the rs16969968 variant and smoking was established among 149 smokers and 148 nonsmokers (odds ratio of 1.84, $p = 0.03$ for 1 allele; odds ratio of 3.59, $p = 0.032$ for 2 alleles) (Hong et al. (2010). The rs16969968 exhibited evidence of a recessive mode of inheritance, resulting in individuals having a 2-fold increase in risk of developing nicotine dependence once exposed to cigarette smoking (Saccone F. et al., 2007).

The rs16969968 SNP leads to a D398N amino acids substitution resulting from an exchange of a guanine (G allele) into adenine (A allele) at position 1192 of the coding sequence (c.1192G>A) (Bierut et al., 2008; Saccone et al., 2007). Expression of the A 'risk-allele' reduces *nAChR* function, which may enhance nicotine dependence susceptibility (Bierut et al., 2008). It has been suggested that reduced *nAChR* function due to A-allele expression may regulate dopamine-mediated reward signaling, thereby facilitating dependence (Bierut et al., 2008). Others have suggested that smokers expressing the A allele may smoke to ameliorate cognitive impairments (Winterer et. al., 2010).

An association study involving more than 2000 individuals together with other functional studies have demonstrated that the risk allele decreases response to a nicotine agonist and thus may be involved in nicotine dependence (Bierut LJ. et al., 2008). Another study replicated the association between this SNP and nicotine dependence in 200 individuals, and also concluded that the rs16969968 (A) allele was significantly associated with "enhanced pleasurable responses" to a person's first cigarette (Sherva R. et al., 2008).

2.4.1.2 rs1051730 (c.645C>T)

In a large genome-wide association study, a SNP in the *CHRNA3* gene (rs1051730) was found highly associated with smoking quantity and with risk of lung cancer and peripheral arterial occlusive disease (Thorgeirsson et al., 2008). Such association reflects an example of gene-environment interaction in which nicotine dependence as a result of the SNP confers risk of lung and cardiovascular diseases through an effect on behavior. The rs1051730 SNP is located in a linkage disequilibrium block containing two other candidate genes, encoding the nicotinic acetylcholine receptor alpha 5 (*CHRNA5*) and beta 4 (*CHRNB4*) subunit genes (Thorgeirsson et al., 2008). The SNP rs1051730 represents a synonymous C>T transition in exon-5 at position 645 of the coding sequence.

Keskitalo et al. (2009) measured the number of cigarettes smoked per day (CPD) and the serum level of an immune-reactive nicotine metabolite

(cotinine) in 516 Finnish daily smokers (aged 30-75 years; 303 males and 213 females (Keskitalo et al. 2009). The associations of 21 SNPs in 15q25.1 were examined with cotinine and CPD. The SNP rs1051730 showed the strongest association with cotinine level and CPD. However, this SNP accounted for nearly a 5-fold larger proportion of variance in cotinine levels than in CPD ($R^{(2)}$ 4.3% vs 0.9%). The effect size of the SNP was 0.30 for cotinine level, whereas it was 0.13 for CPD. Keskitalo et al. (2009) concluded that variation at the *CHRNA5/CHRNA3/CHRNB4* cluster influences nicotine level, measured as cotinine, more strongly than smoking quantity, measured by CPD, and appears thus to be involved in regulation of nicotine levels among smokers.

In two recent studies, together comprising over 6,000 lung cancer patients of European ancestry, the rs1051730 (T) allele was significantly associated with increased risk for lung cancer (Hung RJ. et al., 2008). Having one copy of the allele increased risk for lung cancer about 1.3 times, while having two copies exhibited a 1.8 times increased risk. Up to 14% of lung cancer incidence may be attributable to this allele (Hung RJ. et al., 2008, Amos CI. et al., 2008).

A different study published at the same time concluded that the (T) allele carriers are not at higher risk of becoming smokers compared to (C) carriers (Thorgeirsson E. et. al. 2008). However, if they do smoke, (T) carriers are quite likely to smoke more cigarettes than (C) carriers, and as an apparent consequence, they are at higher risk for lung cancer.

Thorgeirsson et al. (2008) linked the same allele to the risk of 2 smoking-related diseases (lung cancer and peripheral artery disease) in populations of European descent. The SNP was strongly associated with the smoking quantity ($P = 5 \times 10^{-16}$). Thorgeirsson et al. (2008) concluded that their findings provided a case study of a gene-environment interaction, highlighting the role of nicotine addiction in the pathology of other serious diseases.

2.4.2 Region 5p15.

The 5p15.33 locus contains two genes, the *TERT* and the *CLPTM1L*, which have been reported to be implicated in lung carcinogenesis (Chanock J. and Hunter J. 2005, Li C. et. al. 2013).

The *TERT* gene product contains three distinct structural domains: the RNA-binding domain (*TRBD*), the reverse transcriptase domain and the carboxy- terminal extension (CTE), which represents the putative thumb domain of *TERT* (Gillis J. et. al., 2008). Tumor cells can prevent telomere loss upregulation of telomerase (Blasco A., 2005). Activation of telomerase induced by the catalytic component *TERT* is a pivotal step during cellular immortalization and malignant transformation of human cells (Zhang A. et. al., 2000). Telomerase has been found to be reactivated in the majority of cancers, including those of the lung (Lantuejoul S. et. al., 2007).

Palate transmembrane 1-like (*CLPTM1L*), also called cisplatin resistance related gene 9 (*CRR9*), was identified among the genes involved in resistance to the anticancer drug cisplatin in ovarian cancer cells (Yamamoto K. et. al., 2001). *CLPTM1L* is located at the 5p15.33 locus near telomerase reverse transcriptase (*TERT*). Recent genetic studies revealed that this locus is a susceptibility region for lung and several other cancers (Ni Z. et. al., 2012). Several studies have shown *CLPTM1L* to be highly expressed in renal carcinoma cell line and laryngeal squamous cell carcinoma (Asakura T. et. al., 2005, Colombo J. et. al., 2009).

2.4.2.1 rs402710 (g.1320607C>T, c.1532+9G>A).

The SNP rs402710 is located at 5p15.33 and in a region of high LD that includes the proximal and putative promoter regions of *TERT*, as well as the entire coding region of the *CLPTM1L* gene locus. It is located in the intron region of *CLPTM1L* which was is expressed in various tissue, including lung tissue and overexpressed in cisplatin-resistant cell lines, encodes an enzyme—cleft lip and palate trans-membrane 1-like that may be associated with apoptosis (Yamamoto K. et. al. 2001).

The SNP was found associated with the susceptibility to lung cancer prominently in never-smokers (P = 0.01), ex-smokers (P = 0.0007) and

current smokers (P = 0.0001) (McKay et. al. 2008). There was no apparent geographical heterogeneity in the allele frequencies of rs402710. Adjustment for smoking exposure (packs/ year) had no effect on the observed association with a smoking adjusted OR per allele of 1.19 (1.12-1.26). The same study investigated rs402710 in the context of smoking intensity among controls and did not observe any association between number of cigarettes consumed per day and rs402710 (p = 0.74).

Several studies have replicated the significant association of the SNP with lung cancer risk, in Caucasian and Asian population (Jin G et. al., 2009, Hsiung CA. et. al. 2010, Truong T. et. al. 2010, Yoon KA. et.al. 2010, Jaworowska E. et.al. 2011, Pande M. et. al. 2011, Ito H. et.al. 2011, Chen XF. et. al., 2012; Zhao DP. et. al., 2014), while some other replication studies showed an inconsistent outcome (Jin G. et.al. 2009, Yang P. et. al. 2010, Chen XF. et. al., 2012).

Zhao D. et. al. (2014) found both rs402710 and rs401681 conferred significantly greater risks for adenocarcinoma and squamous cell carcinoma when stratified by histological type of tumors. Furthermore, associations of these polymorphisms with lung cancer risk were observed among current smokers and former smokers, as well as never smokers. These findings demonstrated that rs402710 and rs401681 are risk-conferring factors for the development of lung cancer.

The results from a meta-analysis study provided evidence that rs402710 T allele significantly contributed to decreased lung cancer risk, and the case-control study implied that the variant may yield stronger effect on NSCLC and never smokers (Lu X. et. al. 2013).

2.4.2.2 rs2736100 (c.1574-3777G>T).

The rs2736100 polymorphism is localized to intron 2 of the *TERT* gene with a possible contribution to an increased risk of lung cancer (McKay et al. 2008). In follow-up replication studies, several research groups have reported associations between this SNP and cancer risk, but with inconclusive results (Jin G. et. Al., 2009, Gago-Dominguez M. et. al., 2011).

Zou P. et. al., 2012 concluded that the *TERT* rs2736100 polymorphism is associated with cancer risk. Also, Bae EY. et. al., 2012 demonstrated that rs2736100 SNP in the 5p15 region has effect on the risk of lung cancer and it was significant only for adenocarcinoma.

Miki D. et. al., 2010 conducted a genome-wide association study in a Japanese cohort, with replication in two independent studies in Japanese and Korean individuals, in a total of 2,098 lung adenocarcinoma cases and 11,048 controls. The combined analyses identified two susceptibility loci for lung adenocarcinoma: *TERT* (rs2736100, combined P = 2.91×10^{-11}), odds ratio (OR) = 1.27) and *TP63* (rs10937405, combined $P = 7.26 \times 10^{-12}$), OR = 1.31).

2.4.3 Region 6p21.

A recent study from the UK reported that the 6p21.33 locus is associated with lung cancer risk (Wang Y. et al., 2008), and a suggestive association for another SNP in the region (rs4324798 on 6p22.1).

2.4.3.1 rs4324798 (g.79518G>A).

The rs4324798 SNP is located near the major histocompatibility complex locus with the risk allele (A) possibly associated with lung adenocarcinoma in European descent cases (Landi T. et al., 2009). The association of the SNP with lung cancer was found weak and inconsistent in different studies (Wang Y. et. al., 2008, landi t. et. al., 2009). TanTai J. et. Al. (2014) demonstrated that the three common variations (rs4324798, rs3117582, and rs9295740) on 6p21 are risk factors associated with increased lung cancer susceptibility, but these associations vary in different ethnic populations.

Yang P. et. al., 2010 conclude that, the SNP rs4324798, as an independent prognostic factor for overall survival in SCLC patients, an outcome that needs to be validated by other studies.

2.4.3.2 rs9295740 (g.27689502G>A).

The SNP rs9295740 is located in the transfer RNA locus for leucine 16 (anticodon UAA) (*TRNAL16*), on chromosome 6 and the ancestral allele is A (ALFRED. 2014, NCBI. 2014).

Wang et al., 2008 reported that linkage disequilibrium (LD) in the 6p21.33–6p22.1 region was extensive, and rs9295740 at 6p22.1, is in moderate LD with lung cancer risk in the European population. In another study in Koreans population, there was no association between rs9295740 and the risk of lung cancer (Bae EY. et al., 2013).

A significant association between the SNP and the overall NSCLC risk was reported mainly in males with OR =1.921, 95% CI: 0.942–3.919 (de Mello et al, 2013). In the same study, the GA+AA genotype group represented a nearly statistically significant increase in NSCLC risk when compared with the rs9295740 AA genotype group (OR =1.742, 95% CI: 0.979–3.098). Also, they observed the rs9295740 GA genotype in the 6p21 locus demonstrated a trend toward higher Progression-free Survival (PFS) than the rs9295740 GG and rs9295740 AA genotypes: 6 months (range: 3.21–8.78) versus 4 months (range: 0.39–7.60), respectively, p= 0.074. However, the rs9295740 GA+AA genotype group had a higher PFS than the rs9295740 GG genotype group: 6 months (range: 4.31–7.68) versus 4 months (range: 2.06–5.93), respectively, p= 0.034. So, they concluded that the rs9295740 GA+AA genotype group in the 6p21 locus had a higher overall survival (OS) than the rs9295740 GG genotype group: 13 months (range: 9.38–16.61) versus 9 months (range: 5.91 12.08), p =0.045. But, another study in a Korean population concluded there was no association between the rs9295740 and lung cancer risk (Bae EY. et. al., 2012).

Chapter Three

Materials and Methods

3.1 Materials

3.1.1 Reagents

Table 3.1 Reagent used in the study.

No.	Items	Manufacture	Country
1.	QIAamp DNA Blood Mini kit	Qiagen	Germany
2.	GoTaq Green Master Mix, 2X	Promega	USA
3.	DNA Ladder (100 bp)	GeneDirex	USA
4.	Primers	GeneDirex	USA
5.	Restriction Enzyme (TP45I)	New England Biolabs	USA
6.	Agarose	Promega	USA
7.	Ethanol 96%	Local	Local
8.	Ethidum Bromide	Sigma	Germany
9.	Tris-Acetate EDTA, 10X	Promega	USA

3.2 Instruments

Table 3.2 The Instruments used in the study**.**

No.	Instruments	Supplier
1.	Thermal-Cycler: MasterCycler Gradient	
2.	Vortex	
3.	Micropipette 0.5--10 µL	
4.	Micropipette 5--50 µL	Eppendrof, Germany
5.	Micropipette 20--200 µL	
6.	Micropipette 100--1000 µL	
7.	Micro-centrifuge	
8.	Refrigerator (- 80 C^{o})	Ing. Climas, Germany

9.	Gel Electrophoresis Chambers and Power Supply	Biorad, USA
10.	Gel Documentation System	UVP, USA
11.	Nano-Drop (ND-1000)	ThermoScientific, USA
12.	Microwave Oven	LG, Korea

3.3 Study Ethics

The collection of the patient and control materials was approved by the ethical committees of the Palestinian Ministry of Health and Helsinki Declaration of 1975 (106). All enrolled participants were informed about the study according to the study protocol and was give written informed consent.

The approval letter for the present study was obtained from the Helsinki committee and the Palestinian Ministry of Health (MOH) **(Annex 1, 2, 3)**.

3.4 Samples

This study is a case control study with convenience sample. The sample set was consisted of two groups: lung cancer patients and control individuals. Then samples was classified according to gender, age, height, weight, smoking history (current smoking (%), age started (yr), smoking duration, package/years, cigarettes/day), history of other exposures and family history.

Samples were collected from all patients being managed for lung cancer in the oncology clinics of Shifa and EGH hospitals during the period from February 2014 to June 2014.

Any lung cancer case was excluded if the cancer in lung result from metastasis from other organs, and any patient or control must not make blood transfusion for six months at least.

An equal number of matched healthy individuals in age, gender were serving as a control group.

3.5 Settings and place of work

The practical parts of this work were performed in the Molecular Biology Department of the central laboratory, MOH-Gaza.

3.6 Patients data

Patient's medical data were collected from their records in the relevant hospitals and by questionnaire. The data included personal, medical, management and family information (e.g. age, weight, gender, cigarettes per day, type of cancer, smoking status, suffering from infectious diseases, years of smoking, period of quit smoking and other).

3.7 DNA Isolation

For all patients, about 5 ml peripheral blood was collected in tubes containing EDTA, and genomic DNA was extracted from the peripheral lymphocytes using a QIAamp DNA Blood Mini kit (Qiagen, Germany).

3.8 Molecular investigation

3.8.1 Primers Selections

Some of primer pair's sequences were designed by Web-based Allele-Specific PCR assay (WASP) online software, while the remaining primers were designed manually. And the Tm was calculated using Promega Biomath calculater online software (Table 3.1).

Table 3.3: Sequence of primers for each SNP

Gene	SNP ID	Name	Sequance 5'-3'	Product Size(bp)
CHRNA5	rs16969968 (G/A)	rs16Rw	CTTGTAATGTAGCGAATAGAAG**C**	215
		rs16Rm	CTTGTAATGTAGCGAATAGAAG**T**	
		rs16Fc	CGCTATCAACATTCATCATC	
CHRNA3	rs1051730 (C/T)	rs10Rw	TTGTACTTGATGTCGTGTTG**G**	150
		rs10Rm	TTGTACTTGATGTCGTGTTG**A**	
		rs10Fc	ACTGTACCATGAAGTTCGGT	
CLPTM1L	rs402710 (C/T)	rs4Rw	CCAGCGGTGGTGAGTGA**G**	175
		rs4Rm	CCAGCGGTGGTGAGTGA**A**	
		rs4Fc	CAACTGGAACCCAAGTTTAG	
TRET	rs2736100 (G/T)	rs27Fw	CCGTGTTGAGTGTTTCC**G**	223
		rs27Fm	CCGTGTTGAGTGTTTCC**T**	
		rs27Rc	CACATCTTCATCTGTGCATC	
Inron	rs4324798 (G/A)	rs43Rw	GAAAAAAGAACTACCTCGCG**C**	212
		rs43Rm	GAAAAAAGAACTACCTCGCG**T**	
		rs43Fc	TGCTTGGTCACCGTATAAAA	
vWF	rs9295740 (A/G)	F	AGCACTTAGAGTTCTGGGCG	175
		R	AAACGAGTTCATGCTGGGGT	
Beta-Globin Gene		F	CAACTTCATCCACGTTCACC	270
		R	GAAGAGCCAAGGACAGGTAC	

3.8.2 Allele Specific PCR

The genotyping method which used to detect the selected SNPs variants was developed using AS-PCR. PCR was carried out using a thermal cycler (Eppendorf Mastercycler Gradient; Eppendorf, Hamburg, Germany). The final volume of all PCR protocols was 20 µL.

3.8.2.1 AS-PCR for rs16969968

The AS-PCR required the use of two separate tubes for the amplification of wild-type and variant-type allele. The first tube containing the

PCR mixture for the wild-type amplification comprised of 1X GoTaq Green Master Mix, 2 µL of extracted DNA, 0.5µM of rs16Rw and 0.5µM of rs16Fc primer. The PCR mixture for the variant-type amplification comprised similar elements as in the first tube except for the rs16Rw primer, which was replaced with 0.5µM of rs16Rm primer. The PCR cycling was performed with an initial denaturation at 95°C for 5 minutes, followed by 38 cycles of amplification; 95°C for 30 sec, 58°C for 30 sec and 72°C for 45 sec. The final extension step was performed at 72°C for 10 min.

3.8.2.2 AS-PCR for rs1051730.

The AS-PCR required the use of two separate tubes for the amplification of wild-type and variant-type allele. The first tube containing the PCR mixture for the wild-type amplification comprised of 1X GoTaq Green Master Mix, 2 µL of extracted DNA, 0.5µM of rs10Rw and 0.5µM of rs10Fc primer. The PCR mixture for the variant-type amplification comprised similar elements as in the first tube except for the rs10Rw primer, which was replaced with 0.5µM of rs10Rm primer. The PCR cycling was performed with an initial denaturation at 95°C for 5 minutes, followed by 38 cycles of amplification; 95°C for 30 sec, 59°C for 30 sec and 72°C for 45 sec. The final extension step was performed at 72°C for 10 min.

3.8.2.3 AS-PCR for rs402710.

The AS-PCR required the use of two separate tubes for the amplification of wild-type and variant-type allele. The first tube containing the PCR mixture for the wild-type amplification comprised of 1X GoTaq Green Master Mix, 2 µL of extracted DNA, 0.5µM of rs4Rw and 0.5µM of rs4Fc primer. The PCR mixture for the variant-type amplification comprised similar elements as in the first tube except for the rs4Rw primer, which was replaced with 0.5µM of rs4Rm primer. The PCR cycling for rs402710 is the same as rs1051730.

3.8.2.4 AS-PCR for rs2736100.

The AS-PCR required the use of two separate tubes for the amplification of wild-type and variant-type allele. The first tube containing the PCR mixture for the wild-type amplification comprised of 1X GoTaq Green Master Mix, 2 µL of extracted DNA, 0.5µM of rs27Fw and 0.5µM of rs27Rc primer. The PCR mixture for the variant-type amplification comprised similar elements as in the first tube except for the rs27Fw primer, which was replaced with 0.5µM of rs27Fm primer. The PCR cycling was performed with an initial denaturation at 95°C for 5 minutes, followed by 38 cycles of amplification; 95°C for 30 sec, 57°C for 30 sec and 72°C for 45 sec. The final extension step was performed at 72°C for 10 min.

3.8.2.5 AS-PCR for rs4324798.

The AS-PCR required the use of two separate tubes for the amplification of wild-type and variant-type allele. The first tube containing the PCR mixture for the wild-type amplification comprised of 1X GoTaq Green Master Mix, 2 µL of extracted DNA, 0.5µM of rs43Rw and 0.5µM of rs43Fc primer. The PCR mixture for the variant-type amplification comprised similar elements as in the first tube except for the rs43Rw primer, which was replaced with 0.5µM of rs43Rm primer. The PCR cycling for rs402710 is the same as rs16969968.

3.8.3 Restriction Enzyme Length Polymorphism (RFLP)

3.8.3.1 rs9295740-RFLP

The SNP rs9295740 was genotyped by using polymerase chain reaction (PCR)–restriction fragment length polymorphism method. The reaction was composed of a pair of primers flanking the polymorphic site 0.5µM of forward and 0.5µM of reverse primer to amplify a 175-bp PCR product for rs9295740.

The PCR cycling was performed with an initial denaturation at 95°C for 5 minutes, followed by 38 cycles of amplification; 95°C for 30 sec, 61°C for 30

sec and 72°C for 45 sec. The final extension step was performed at 72°C for 10 min.

The amplicon was digested by restriction enzyme of *Tsp45I* (New England BioLabs, Beverly, MA) by incubation at 65 °C for 60 min with reaction volume 30 µl, then separated on a 2% agarose gel. The G allele was produce two fragments of 102- and 73-bp while the A allele results in a single 175-bp fragment.

3.8.4 β-Globin Gene

β-Globin gene was amplified by the primers that target a 270-bp region of the gene were done behind each reaction at same concentration that used in SNP genotyping.

The cycling conditions for β-Globin gene were carried out with a first denaturation step at 95ºC for 5 min, followed by 38 cycles at 95°C for 30 sec, 61°C for 30 sec, 72°C for 45 sec and then a final extension step at 72ºC for 10 min

3.9 Quality Control

In each PCR run a positive control of a sample with a known heterozygous genotype and a blank water (H_2O) sample was run simultaneously. To validate the genotyping results, 10% randomly selected samples was repeated and the concordance rate was determined.

3.10 Agarose gel electrophoresis

After the amplification, electrophoresis was performed at 100 V for 50 min in 1X tris-borate-EDTA buffer on 2% agarose gels stained with ethidium bromide (0.5 µg/ µL). The amplified PCR products were visualized under UV light.

3.11 Biostatistics/ Data analysis

The data was analyzed by the SPSS software (version 19). The one way ANOVA test was used for mean comparisons when indicated with a 95% confidence interval. The figures were prepared and presented by Microsoft office excel 2010 program.

Differences in the distribution of categorical variables were analyzed using the χ^2 test. To test for population stratification the deviation of the genotype frequencies in the controls from those expected under Hardy-Weinberg Equilibrium (HWE) was assessed by χ^2 test. Odds ratios (ORs) and associated 95% confidence intervals (CIs) were calculated by unconditional logistic regression, adjusting for age, sex and smoking. To avoid issues associated with non-normally distributed adduct data and to take into account co-variates the data were log-transformed and the impact of genotypes on adduct levels was assessed by generalized linear modeling. Two-sided P values <0.05 were considered as statistically significant.

3.12 Limitations of the study

One major limitation of the present study is that the relatively small sample size, especially in the stratified analyses, limited our ability to detect moderate interactions. Large-scale epidemiological studies are needed in the future to confirm our findings. Secondly, Most of patients hadn't known about its disease, so it was difficult to take the sample or information from them. Thirdly, the war on Gaza and psychological factors effect on the work direction of the study. Lastly, the controls were not uniformly defined. Although the most of the controls were selected mainly from healthy smoker populations.

Chapter Four

Results and Discussion

4.1 Study population description

The study population comprised 30 lung cancer patients that can be further classified into 21 SCLC and 9 NSCLC patients. The control group comprised 60 participants and consisted of 30 smoker and 30 non-smoker individuals matched in age and gender to the cases. The cases and controls were defined by basic descriptions of age, sex, cigarette pack/day smoking, period of smoking, family history of lung cancer and prior medical history of pneumonia.

4.1.1 Gender distribution

A total of 30 samples patients were studied and the results show a clear gender-specific difference in lung cancer rates, with much higher rates in men (93%) than in women (Figure 4.1).

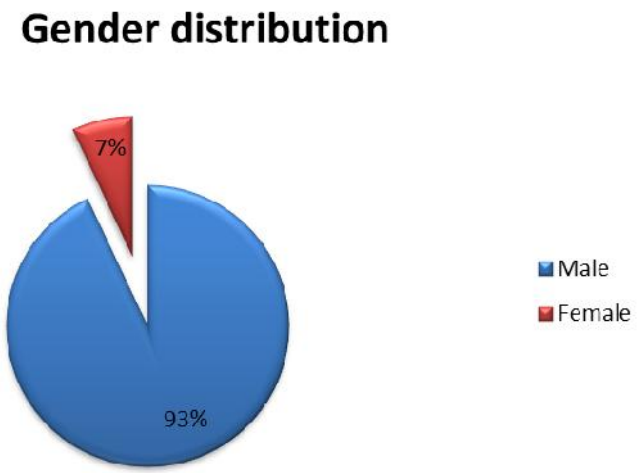


Figure 4.1 Distribution of cases by gender.

4.1.2 Living area

Lung cancer patients were distributed allover Gaza strip as follows: 5 cases (17%) from North Gaza, 10 cases (33%) from Gaza City, 3 cases (10%) from Middle Area, 8 cases (27%) from Khan Younis City and 4 cases (13%) from Rafah City (Figure 4.2).

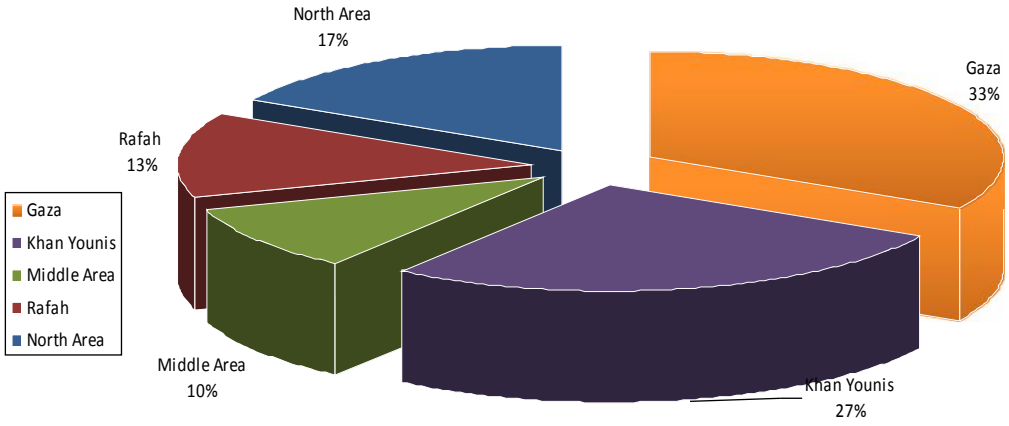


Figure 4.2 Distribution of cases by living area.

4.1.3 Age distribution

The age distribution between the cases and controls is described in table (4.1). The mean age of case and control groups was not different as both groups were matched in age.

Table4.1: Age distribution of cases and controls

	Lung cancer cases	Control
Age		
Mean ±SD	61.4 ± 9.1	60.4 ± 7.9
Median (min-max)	61.5 (45-80)	61(43-80)

4.1.4 Case distribution according to place of sample collection

The patient samples were collected from lung cancer patients admitted to Oncology departments of two main hospitals in Gaza Strip, Al-Shifa Hospital (17 cases, 56%) and the European Gaza Hospital (13 cases, 44%, figure 4.4).

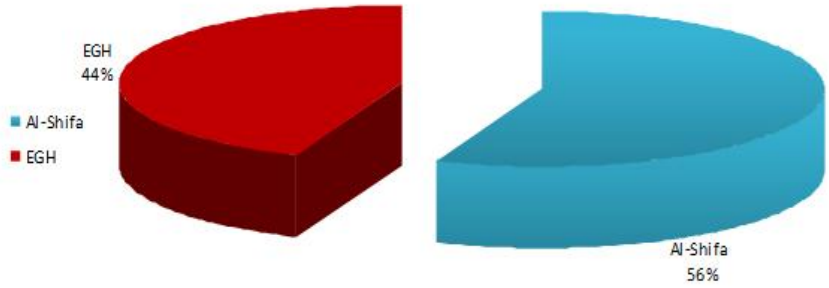


Figure 4.3 Distribution of lung cancer samples by place of collection

4.1.5 Distribution of cases by type of lung cancer.

Two main types of lung cancer were encountered in this study. Twenty one cases (70%) were the small cell lung cancer (SCLC) type and nine cases (30%) were of the non-small cell lung cancer (NSCLC) type. Two cases of the NSCLC were Adenocarcinoma (Figure 4.4). The diagnosis of each type of lung cancer was confirmed previously by diagnostic tools in the managing hospitals.

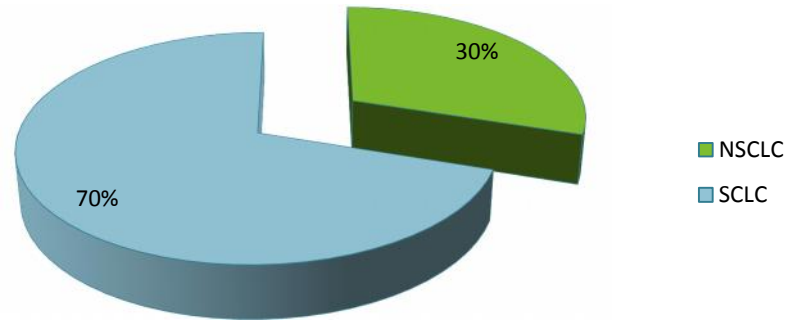


Figure 4.4 Distribution of cases by type of lung cancer.

4.1.6 History of lung infectious diseases

The distribution of cases and controls by history of lung infectious diseases was statistically significant with 40% of cases compared to none of the control group (*P*-value = 0.00, OR = 1.765, table 4.2). Moreover, most of those with a history of lung infectious diseases are among smoker cases (*P*-value = 0.00). We found 11 cases suffering from lung infectious diseases was a smoker case while one case suffering from lung infectious diseases was in non-smoker cases.

Table 4.2: Distribution of cases and controls by history of lung infectious diseases

		Controls	Cases	*P*-Value
lung infectious diseases	Yes	0 (0%)	12 (40%)	0.00
	No	60 (100%)	18 (60%)	

In our study, the small number of cases with a history of pneumonia emphasizes the need for further studies with larger study population to confirm such a relationship. Nevertheless, this association can't be neglected, because it is consistent with other studies that have also reported increased risks associated with pneumonia (Denholm R. et al., 2014, Ibrahim M. et. al., 2013, Brenner R. et. al., 2011, Zhan P. et. al., 2011). Chronic inflammation has been shown to increase cancer risk by influencing every stage of cancer from initiation, promotion, invasion, and metastasis via induction of oncogenic mutations and genomic instability, local immunosuppression, and angiogenesis (reviewed in Grivennikov et al 2010). Chronic inflammation triggered by tobacco smoke has been shown to promote lung carcinogenesis (Takahashi et al 2010). Inflammation induced by cigarette smoke also promotes COPD, a disease associated with increased lung cancer risk (Grivennikov et al 2010, Punturieri et al 2009). In contrast, nicotine itself appears to suppress immune function and has been shown to be protective against inflammatory diseases such as pneumonia and ulcerative colitis (Rubin and Hanauer 2000, Blanchet et al 2004, Shivji et al 2005). Suppression of the immune response by nicotine may impact immune

surveillance, preventing the clearance of nascent tumor cells (Gahring and Rogers 2006, Grivennikov et al 2010).

4.1.7 Smoking status

Only 4 lung cancer cases were non-smoker cases while the rest 26 were smoker cases (Figure 4.5).

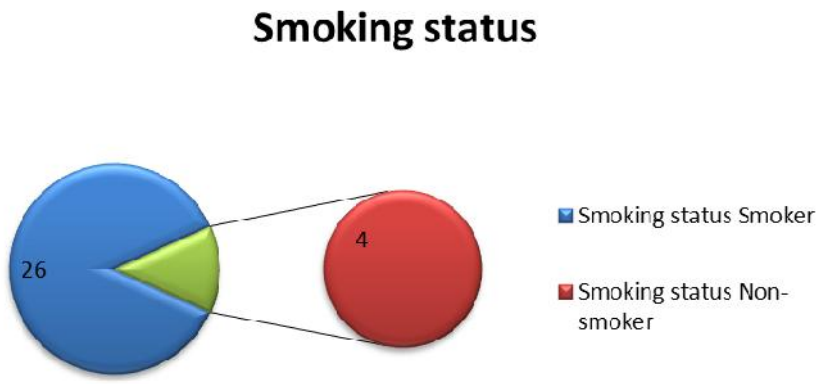


Figure 4.5 Distribution of cases by status of smoking.

4.1.8 Cigarettes per day

The amount of cigarettes smoked per day by lung cancer patients and smoker controls was not statistically significantly different (table 4.3, *P*-value = 0.823).

Table4.3: Number of cigarettes per day

	Smoker Lung cancer cases	**Smoker control**	***P-value***
Mean ±SD	30.2±13.2	31.0±12.5	
Median (min-max)	30 (10-60)	30 (15-57)	0.823

4.1.9 Duration of Smoking

Few patients have quitted smoking just after being diagnosed and thus were combined with current smokers in one group (Smokers).

The cumulative period of smoking (in years) was found to have significant relation to lung cancer susceptibility (mean = 28.0 ±10.5 years for smoker controls, compared to 39.04 ±12.7 years in smoker cases (*P*-value =0.001, figure 4.7).

Table4.4: Distribution of cases and controls by duration of smoking (Years)

	Smoker Lung cancer cases	**Smoker control**	***P-value***
Mean ±SD	39.04 ±12.7	28.0 ±10.5	0.001

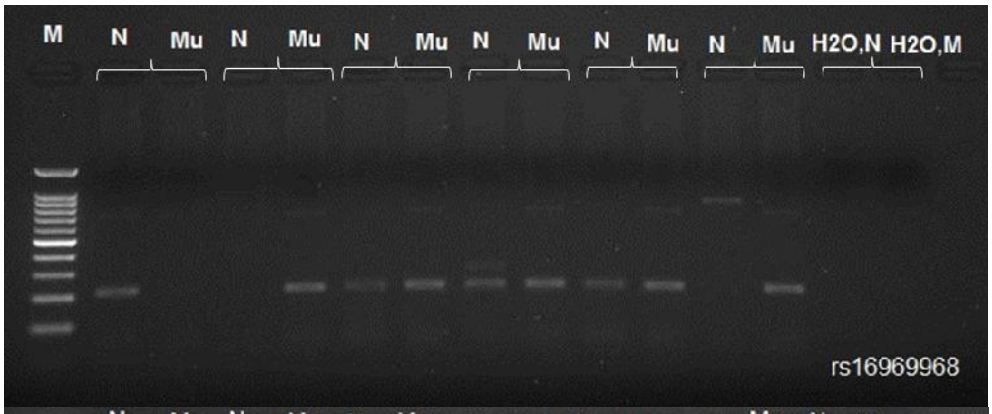
These result support that the duration of smoking (period of exposure to carcinogens) also can be considered as risk factor for lung cancer and this finding is supported by other researches (Cancer Research UK Statistical Information Team. 1978, Doll R. et. al., 2005, Lubin JH. et. al., 2006, Chen LS. et. al., 2015).

The age at which one starts smoking thus can be associated with the risk for lung cancer. Smoking at an earlier age exposes the smoker more and more to carcinogenic substance present in cigarettes that may initiate the process of cellular transformation. This speculation is supported by results of others (Flanders WD. et. al., 2003; Al-Moustafa A. 2012).

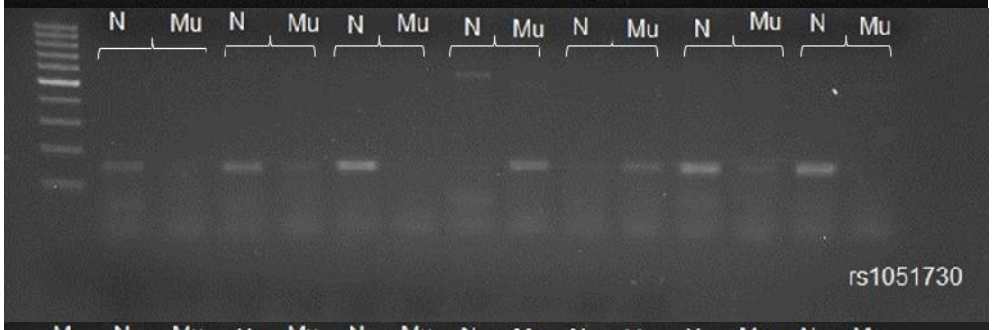
4.1.10 Qualitative detection of SNPs.

Ninety subjects were recruited for the study. Samples were genotyped for rs16969968, rs1051730, rs402710, rs2736100, rs4324798 and rs9295740. The allele specific amplification and analysis of each SNP was performed successfully as shown in Figure 4.6. The presence of PCR bands with different sizes in the agarose gel indicated the genotype of the samples.

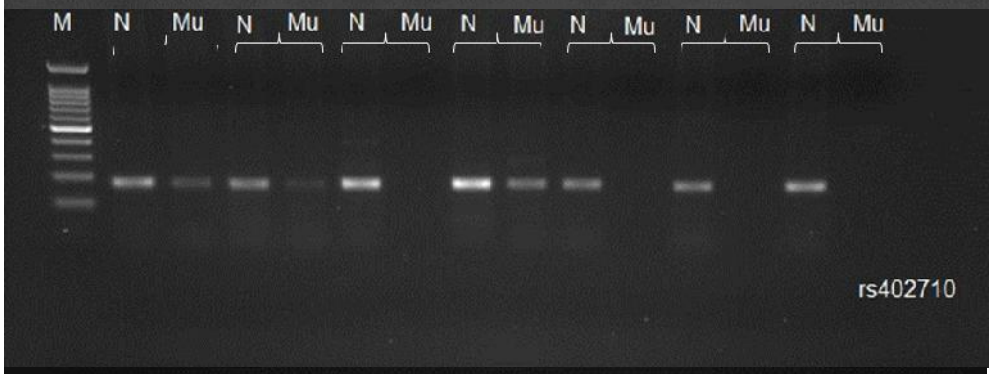
A.


M
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
H2O,N
H2O,M
rs16969968

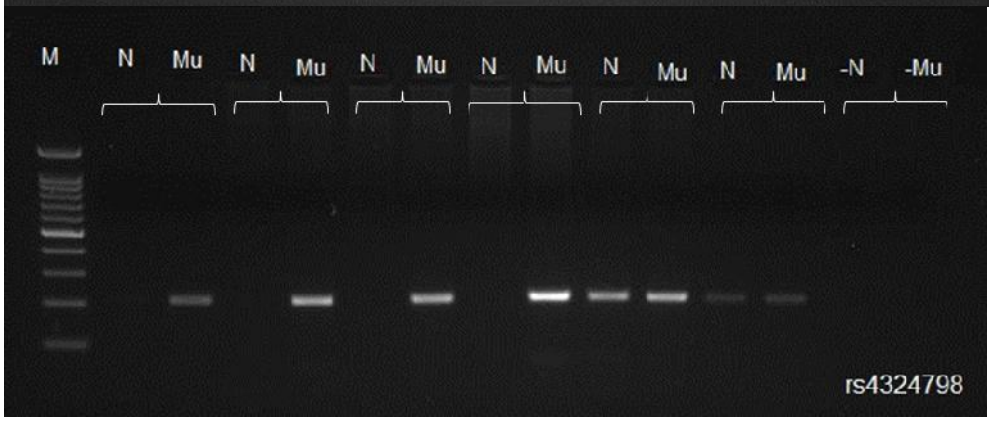
B.


N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
rs1051730

C.


M
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
rs402710

D.


M
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
N
Mu
-N
-Mu
rs4324798

E.

F.

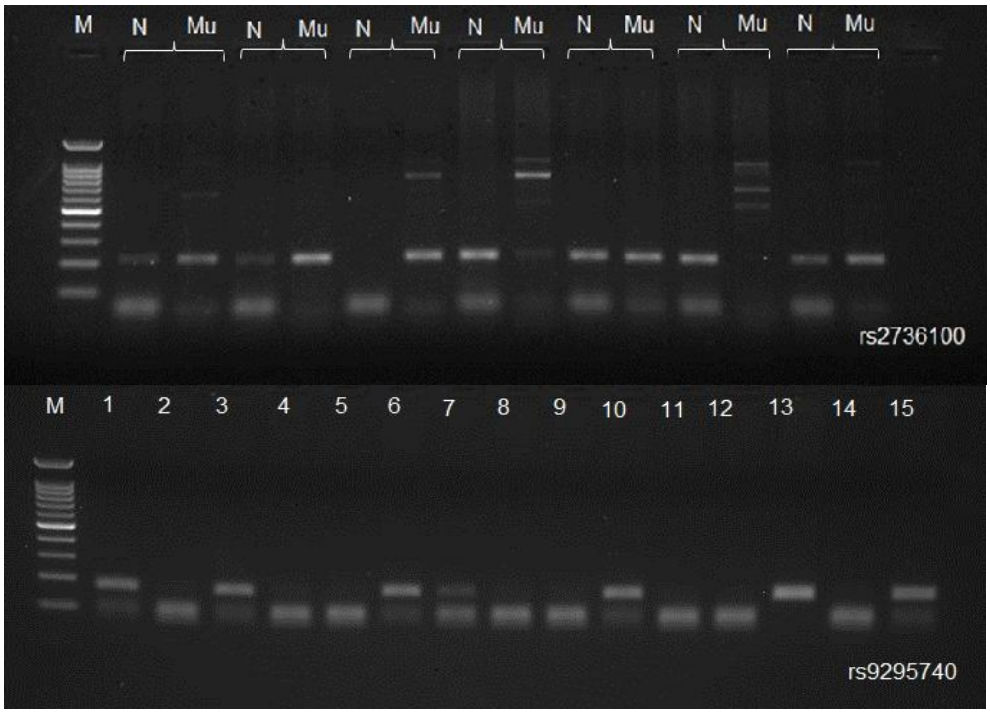


Figure 4.6 The electrophoresis profiles for representative amplifications. rs16969968 (**A**), rs1051730 (**B**), rs402710 (**C**), rs4324798 (**D**), rs2736100 (**E**) and rs9295740 (**F**). **N** = is the lane for amplification with primers specific for the normal allele, **Mu** = is the lane for amplification with primers specific for the mutant allele, **M**= DNA ladder (marker). See materials and methods for more details.

4.2 SNPs Genotyping and Relations

4.2.1 rs16969968, G>A

The genotype of the SNP (rs16969968) in the cases and the unrelated smoker and non-smoker controls is illustrated in table 4.5 and figure 4.6A.

When the Hardy-Weinberg equilibrium is applied in our study population, the allelic frequency for the “A” allele of the SNP rs16969968 was 23.89% (*P*-value= 0.6165). In comparison, the documented frequency was 39% in Utah residents with ancestry from northern and western Europe (CEPH) populations, 1% in Japanese in Tokyo (JPT) populations, 3% Han Chinese in Beijing (CHB), 8% in African American population, 40.2% in Toscani in Italia (TSI) populations, 22.4% in Mexican ancestry in Los Angeles (MEX) populations and 21.3% in Gujarati Indians in Houston (GIH) populations (PharmGKB, 2010).

The allelic frequency of the susceptibility allele is higher in the case group (36.7%) than in the control group (17.5%, table 4.5). The distribution of alleles among case and control is statistically significant (*P*-value = 0.022)

Table 4.5: The distribution of alleles among case and control

		Control (n=60)	**Case** (n=30)	**Total** (n=90)	***P-value***
rs16969968	G	41(68.3%)	12 (40%)	53 (58.9%)	0.022
	GA	17 (28.3%)	14 (46.7%)	31 (34.4%)	
	A	2 (3.4%)	4 (13.3%)	6 (6.7)	
Allele Frequency		17.5%	36.7%	23.9%	0.6165

By examining the potential role of carriage of the A- allele as a risk factor for lung cancer, heterozygousity and particularly homozygousity for the susceptibility allele 'A' was found to increase the risk for lung cancer when compared to the wild type genotype (OR= 2.81 and 6.83 respectively; table 4.6).

Lips E. et. al. (2011) has explained that the strong association between the SNP and Lung cancer risk (OR= 1.30, 95% CI1.23-1.38, P = 1 $x10^{-18}$), was virtually unchanged after adjusting for smoking (Smoking adjusted OR= 1.27, 95%CI 1.19-1.35, P= $5x10^{-13}$). In addition, Chen S. et. al., 2015 emphasized the clinical significance of the *CHRNA5* variant rs16969968 as a predictor of delayed smoking cessation and an earlier age of lung cancer.

Table 4.6: Variant genotype with lung cancer risk

Genotype		**Risk**		***P*-value**
		OR	**95%CI**	
rs16969968 G/A	GA+AA vs GG	3.23	1.30-8.05	0.010
	GA vs GG	2.81	1.08-7.32	0.029
	AA vs GG	6.83	1.11-41.97	0.041

As previously pointed out, the mean number of cigarettes/day (CPD) was not significantly different between the smoker controls and cases (table 4.3, P-value = 0.823). However, the allelic frequency of the susceptibility allele was higher in the case than in the control (36.7% vs 23.3% respectively; table

4.7). This suggests a role for this allele in predisposition to lung cancer apart from smoking. This suggested role was previously pointed out (Timofeeva N. et. al. 2012).

Table 4.7: Allele distribution among smoker cases and smoker controls

	Control (CPD= 31.0±12.5)	**Case (CPD= 30.2±13.2)**	***P-Value***
GG	18	10	
GA	10	13	0.272
AA	2	3	
frequency	23.3%	36.7%	

The differences in distribution of the rs16969968 between the smoker control and smoker cases was not statistically significant (*P*-value = 0.272; table 4.7).

Because of the small sample size, analysis of statistical significance could not be applied to compare between non-smoker case (n=4) and non-smoker control (n=30).

By calculating the allelic frequency of the A allele in nonsmokers and smokers regardless of lung cancer, the frequency was higher in smokers (29.1%) than nonsmokers (15.7%). Furthermore, table 4.8 shows that the higher the proportion of the A allele, the higher the mean CPD (P- value = 0.006). Accordingly, we may look at homo- or heterozygousity for the A- allele as a predisposing factor for smoking dependence, which in turn confers the individual susceptible to lung cancer. This assumption strengthens previously reported conclusions that association of this variant with lung cancer risk is likely mediated largely, if not wholly, via tobacco exposure (Munafò R. et. al., 2012).

Our results also confirm and extend previous findings for associations between rs16969968 and lung cancer, and smoking quantity (Amos et. al., 2008; Spitz R. et al., 2008, Gabrielsen E. et. al., 2013). The rs16969968 was shown to exhibit evidence of a recessive mode of inheritance, with homozygotes having a 2-fold increase in risk of developing nicotine

dependence once exposed to cigarette smoking (Saccone F. et al., 2007). Hong et al. (2010) found a significant association (p = 0.020) between the rs16969968 variant and smoking among 149 smokers and 148 nonsmokers (odds ratio of 1.84, P = 0.03 for 1 allele; odds ratio of 3.59, P = 0.032 for 2 alleles). The two variants of the nicotinic acetylcholine receptor subunit genes *CHRNA5* and *CHRNA3* on 15q25, rs16969968 and rs578776, were previously found associated with cotinine (P = 0.001 and 0.03, respectively) in current smokers and with lung cancer risk (P < 0.001 and P = 0.001, respectively) (Timofeeva N. et. al. 2012). Furthermore, the serum nicotine level was reported to be significantly correlated with the results of the Fagerström test for Nicotine Dependence (P<0.001) establishing for the first time a relationship between (rs16969968) and the strength of nicotine addiction (Wojas-Krawczyk K. et. al. 2012). Carriers of the A allele expressed significantly higher levels of nicotine when compared with the carriers of the GG genotype (P=0.05).

Similar to our results (tables 4.3 and 4.4), the increasing nicotine levels were previously reported to indicated an increased risk of lung cancer with each additional copy of the rs1051730 and rs16969968 risk allele (per-allele odds ratio = 1.31, 95% CI = 1.21 to 1.42) (Munafò R. et. al., 2012).

On the other hand, few other reports concluded that cigarette consumption/day is not an adequate measure of smoking dose (Le Merchand L. et al., 2008). Moreover, other research yet demonstrated that the *CHRNA5-A3-B4* gene variants (rs16969968, rs588765 and rs578776) do not exhibit a robust association with smoking cessation and are unlikely to be useful for clinically optimizing smoking cessation pharmacotherapy for Caucasian smokers (Tyndale RF. et. al., 2015).

Table 4.8: SNPs in *nAChR* gene vs number of cigarettes per day (CPD)

SNP	CPD	*P-Value*
	ALL case and smoker control	
rs16969968		
GG	25.54±11.6	
GA	36.04±12.2	0.006
AA	40.0±15.81	

The distribution of rs16969968 in the case group according to the type of lung cancer (SCLC or NSCLC) is shown in table 4.9. The allelic frequency of the susceptibility allele "A" in SCLC (35.7%) was lower than in NSCLC (38.9%). However, this difference was not statistically significant (*P*-value=0.815). Similarly, the distribution of cases by cancer type was not significant in terms of the status of ever suffering from pneumonia, (*P*-value=0.100).

Table 4.9: Distribution of the rs16969968 between patients depends on lung cancer type.

SNP	Cancer Type		*P-Value*
	SCLC	NSCLC	
GG	9 (42.8%)	3 (33.4%)	
GA	9 (42.8%)	5 (55.5%)	0.815
AA	3 (14.4%)	1 (11.1%)	
Allele Frequencies	35.7%	38.9%	

4.2.2 rs1051730, C>T

The genotype of the SNP (rs1051730) in the cases and the unrelated smoker and non-smoker controls is illustrated in table 4.10 and figure 4.6B.

When the Hardy-Weinberg equilibrium is applied in our study population, the allelic frequency for the "T" allele of the SNP rs16969968 was 31.11% (*P-value*= 0.887). In comparison, the documented frequency was

39% in Palestinian population (ALFRED. 2014) 39% in Utah residents with ancestry from northern and western Europe (CEPH) populations, 10% in YRI (Yoruba) populations, 1% in Japanese in Tokyo (JPT) populations, 4% Han Chinese in Beijing (CHB), 19% in African American population, 41.2% in Toscani in Italia (TSI) populations, 21.6% in Mexican ancestry in Los Angeles (MEX) populations and 21.3% in Gujarati Indians in Houston (GIH) populations (PharmGKB. 2008).

The allelic frequency of the susceptibility allele is higher in the case group (46.7%) than in the control group (22.5%, table 4.10). The distribution of alleles among case and control is statistically significant (*P*-value = 0.001).

Table 4.10: The distribution of alleles among case and control

		Control (n=60)	**Case** (n=30)	**Total** (n=90)	**P-value**
rs1051730	**CC**	34 (56.7%)	9 (30.0%)	43 (47.8%)	0.001
	CT	25 (41.7%)	14 (46.7%)	39 (43.3%)	
	TT	1 (1.6%)	7 (23.3%)	8 (8.9%)	
Allele Frequency		22.5%	46.7%	30.5%	0.8870

By examining the potential role of carriage of the T- allele as a risk factor for lung cancer, the form of heterozygousity and particularly homozygousity for the susceptibility allele 'T' was found to increase the risk for lung cancer when compared to the wild type genotype (OR= 3.05; table 4.11).

Previous studies was suggested that *CHRNA3* rs1051730 polymorphism is a risk factor associated with increased lung cancer susceptibility among Asian population (Liu P. et. al. 2010, Ping Z. and Yong S. 2015), but these associations vary in different ethnic populations (Gu M. et. al. 2012; Hu B. et. al. 2014), also it associated with the risk for lung adenocarcinoma (p-value=1.9×10^{-10}, Tseng S. et. al. 2014)

In the other side, other research concluded that the rs1051730 was not associated with lung cancer risk (Girard N. et. al. 2010).

Table 4.11 : Variant genotype with lung cancer risk

Genotype		Risk		*P-value*
		OR	**95%CI**	
rs1051730 C/T	CT+TT vs CC	3.05	1.20–7.75	0.015
	CT vs CC	2.20	0.82–5.91	0.113
	TT vs CC	13.22	2.33–74.9	0.001

As previously pointed out, the mean number of cigarettes/day (CPD) was not significantly different between the smoker controls and cases (table 4.3, *P*-value = 0.823). But, the allelic frequency of the susceptibility allele was higher in the smoker case than in the smoker control (46.7% vs 22.5% respectively; table 4.10). This suggests a role for this allele in predisposition to lung cancer apart from smoking. These results supported by previously study (Ren JH. et. al. 2013).

Table 4.12: Allele distribution among smoker cases and smoker controls

	Control (CPD= 31.0±12.5)	**Case (CPD= 30.2±13.2)**	***P-Value***
CC	14	7	
CT	16	13	0.015
TT	0	6	
frequency	26.7%	41.7%	

The differences in distribution of the rs1051730 between the smoker control and smoker cases was statistically significant (*P*-value = 0.015; table 4.12).

Because of the small sample size, analysis of statistical significance could not be applied to compare between non-smoker case (n=4) and non-smoker control (n=30).

By calculating the allelic frequency of the T allele in nonsmokers and smokers regardless of lung cancer, the frequency was higher in smokers (26.6%) than nonsmokers (18.3%). Furthermore, table 4.13 shows that the higher the proportion of the T allele, the higher the mean CPD (*P- Value* = 0.015). Accordingly, we may look at homo- or heterozygousity for the T- allele

as a predisposing factor for smoking which in turn confers the individual susceptible to lung cancer.

Our results also confirm and extend previous findings for smoking intensity, which demonstrate that in the smokers the heterozygotes (CT) and homozygotes (TT) for rs1051730 genotype had higher smoking intensity compared with non-carriers (CC) (Wium-Andersen K. et. al., 2015). Also the previous study explained that the tobacco consumption was 21.1 pack-years in non-carriers, 22.8 in heterozygotes and 24.8 in homozygotes (P-trend<0.001, Rode L. et. al., 2014).

In other study, it's showed a stronger association of rs1051730-rs16969968 (mentioned previously) genotype with objective measures of tobacco exposure compared with self-reported cigarette consumption. The association of these variants with lung cancer risk is likely to be mediated largely, if not wholly, via tobacco exposure. And it's explained that the increasing in nicotine levels indicated an increased risk of lung cancer with each additional copy of the rs1051730-rs16969968 risk allele (per-allele odds ratio = 1.31, 95% CI = 1.21 to 1.42; Munafò R. et. al., 2012; Ware J. et. al., 2015).

Previous study, suggested to that the rs1051730 functions as a genetic modifier of the risk of developing lung adenocarcinoma (ADC) in the Chinese population, particularly in nonsmoking females (He P. et. al., 2014; Tseng S. et. al., 2014)

But in different researches, they was demonstrate that no association between the smoking habit and the *CHRNA3* rs1051730 polymorphism was observed (Li C. et. al. 2013), and no correlation with smoking cessation (Leung T. et. al., 2015).

Table 4.13: SNPs in *nAChR* gene vs number of cigarettes per day (CPD)

SNP	CPD ALL case and smoker control	*P-value*
rs1051730		
CC	24.86±12.97	0.015
CT	33.00±10.36	
TT	39.17±15.62	

The distribution of rs1051730 in the case group according to the type of lung cancer (SCLC or NSCLC) is shown in table 4.14. The allelic frequency of the susceptibility allele "T" in SCLC (50.0%) was higher than in NSCLC (38.8%). However, the distribution of rs1051730 in cases group depending on the type of lung cancer (SCLC or NSCLC) was not statistically significant (*P*-value=0.582). Similarly, the distribution was not significant according to the status of ever suffering from pneumonia, (*P*-value=0.350).

Table 4.14: Distribution of the rs1051730 between patients depends on lung cancer type.

SNP	Cancer Type		*P-Value*
	SCLC	**NSCLC**	
CC	6 (28.5%)	3 (33.3%)	
CT	9 (43.0%)	5 (55.6%)	0.582
TT	6 (28.5%)	1 (11.1%)	
Allele Frequencies	50.0%	38.8%	

4.2.3 rs402710, C>T

The genotype of the SNP (rs402710) in the cases and unrelated smoker and non-smoker controls is illustrated in table 4.15 and figure 4.6C.

When the Hardy-Weinberg equilibrium is applied in our study population, the allelic frequency for the "T" risk allele of the SNP rs402710 was 17.2% and the allelic distribution was in equilibrium (*P*-value= 0.807). In

comparison, according to the Allele Frequency Database, the documented frequency was 23% in Palestinian population (ALFRED, 2014), 48% in African populations, 35% in Mixed American, 31% in East Asian, 33% in European population and 16% in South Asian populations (ENSEMBLE, 2015).

The allelic frequency of the susceptibility allele is higher in the case group (28.3%) than in the control group (14.2%, table 4.15). The distribution of alleles among case and control is statistically significant (*P*-value = 0.015).

Table 4.15: The distribution of alleles among case and control

		Control (n=60)	**Case** (n=30)	**Total** (n=90)
rs402710	CC	45(75.0%)	15 (50.0%)	60 (66.7%)
	CT	13 (21.7%)	13 (43.3%)	26 (28.9%)
	TT	2 (3.3%)	2 (6.7%)	4 (4.4)
Allele Frequency		14.2%	28.3%	18.9%

By examining the potential role of carriage of the T- allele as a risk factor for lung cancer, heterozygousity or homozygousity for the susceptibility allele 'T' was found to increase the risk for lung cancer when compared to the wild type genotype (OR= 3.00; table 4.16). in consistence with this result, previous studies found that rs402710, confers significantly greater risks for adenocarcinoma and squamous cell carcinoma when stratified by the histological type of tumor, (Chen XF. et. al., 2012; Wu H. and Zhu R., 2014), or for lung cancer in general in different populations (Jin G et. al., 2009; Li H. et. al. 2012; Zhao DP. et. al., 2014).

The risk allele of rs402710 was reported to be associated with significantly higher levels of bulky aromatic/hydrophobic DNA adducts (P = 0.02); demonstrating a potential association between the allele and DNA adducts formation and hence a basis for susceptibility to the development of lung cancer (Zienolddiny S. et. al., 2009).

Interestingly, two studies found rs402710 to be associated with decreased lung cancer risk (Lu X. et. al., 2013 and Xun X. et. al., 2014).

Table4.16: Variant genotype with lung cancer risk

Genotype		Risk		*P*-value
		OR	95%CI	
rs402710 C/T	CT+TT vs CC	3.00	1.19–7.55	0.018
	CT vs CC	3.00	1.14–7.88	0.023
	TT vs CC	3.00	0.38–23.19	0.273

No statistically significant differences could be found in the distribution of rs402710 SNP between smoker control and smoker cases (*P*-value = 0.621; table 4.17). In addition, because of the small sample size, analysis of statistical significance to compare between non-smoker case (n=4) and non-smoker control (n=30) was not feasible.

Table 4.17: Allele genotypes and its correlations

rs402710 C/T		Type		*P-value*
		Control	Case	
Smoker	CC	22	16	0.621
	CT	6	8	
	TT	2	2	
Allele Frequency		16.7%	23.1%	

By calculating the allelic frequency of the T allele in nonsmokers and smokers regardless of lung cancer, the frequency was higher in smokers (20.0%) than nonsmokers (8.3%), which may point out to a possible role of the SNP in smoking. However, table 4.18 shows that the mean number of cigarettes smoked per day was not significantly different in smoker cases compared to smoker controls (*P-value* = 0.621). Similarly the duration of smoking was not found to be significantly different (*P-value* = 0.437, table4.19). Therefore, we may conclude that the risk allele independently increases the risk for lung cancer with no relation to smoking. This conclusion is confirmed by findings of a previous study that reported lack of association between the amount of cigarettes and the carriage of the SNP rs402710 (*p* = 0.74, McKay et. al., 2008).

Table 4.18: rs402710 vs number of cigarettes per day (CPD)

SNP	CPD ALL case and smoker control	*P-value*
rs402710		
CC	29.97±13.85	0.621
CT	32.25±11.17	
TT	26.67±11.55	

Table 4.19: rs402710 vs duration of smoking

SNP	**Duration (**Yrs**)** ALL case and smoker control	*P-value*
rs402710		
CC	34.7±13.1	0.437
CT	30.2±12.8	
TT	35.0±5.0	

The distribution of rs402710 in the case group according to the type of lung cancer (SCLC or NSCLC) is shown in table 4.20. The allelic frequency of the susceptibility allele “T” in SCLC (21.4%) was lower than in NSCLC (44.2%). However, the distribution of rs402710 in cases group depending on the type of lung cancer (SCLC or NSCLC) was not statistically significant (*P*-value=0.137). Similarly, the distribution was not significant according to the status of ever suffering from lung infectious diseases (*P*-value=0.266). As mentioned earlier, some studies have established a statistically significant relation between the CT genotype at rs402710 and lung adenocarcinoma, a type of NSCLC (Truong T. et. al., 2010, Chen XF. et. al., 2012).

Table 4.20: Distribution of the rs402710 between patients depends on lung cancer type.

SNP	Cancer Type		*P-Value*
	SCLC	NSCLC	
CC	13 (61.9%)	2 (22.2%)	
CT	7 (33.3%)	6 (66.7%)	0.137
TT	1 (4.8%)	1 (11.1%)	
Allele Frequencies	21.4%	44.2%	

4.2.4 rs2736100, G>T

The genotype of the SNP (rs2736100) in the cases and the unrelated smoker and non-smoker controls is illustrated in table 4.21 and figure 4.6E.

When the Hardy-Weinberg equilibrium is applied in our study population, the allelic frequency for the "T" allele of the SNP rs2736100 was 23.89% (*P-value*= 0.9371). In comparison, the documented frequency was 37% in Palestinian population (ALFRED. 2014), while it was 47.3% in Utah residents with Northern and Western European ancestry (CEU) populations, 62.4% in Japanese in Tokyo (JPT) populations, 56.4% Han Chinese in Beijing (CHB), 52% in African American population, 46.6% in Toscani in Italia (TSI) populations, 61.2% in YRI (Yoruba) populations, 35.5% in Indian American population (GIH) and 41.4% in Maasai Kenya population (MKK, NCBI. 2015).

The allelic frequency of the susceptibility allele is higher in the case group (35.0%) than in the control group (18.3%, table 4.21). The distribution of alleles among case and control is statistically significant (*P*-value = 0.044)

Table 4.21: The distribution of alleles among case and control

		Control (n=60)	**Case** (n=30)	**Total** (n=90)
rs2736100	GG	40 (66.7%)	12 (40.0%)	52 (57.8%)
	GT	18 (30.0%)	15 (50.0%)	33 (36.7%)
	TT	2 (3.3%)	3 (10.0%)	5 (5.5%)
Allele Frequency		18.3%	35.0%	23.8%

By examining the potential role of carriage of the T- allele as a risk factor for lung cancer, heterozygousity and homozygousity for the susceptibility allele ‘T’ was found to increase the risk for lung cancer when compared to the wild type genotype (OR= 4.33; table 4.22).

These results was supported by previous study which showed that rs2736100 was associated with the risk of lung cancer not only in an additive model (OR=1.19, 95% CI: 1.04-1.35; p=0.01), but also in a dominant model (OR=1.14, 95% CI: 1.01-1.28; p=0.03) (Wang HM. et. al., 2013). And in another research, its results indicated that a genetic variation rs2736100 may increase lung cancer risk, which is consistent with earlier prospective studies relating longer telomere length with increased lung cancer risk (Machiela MJ. et. al., 2015).

Gene variations rs2736100 and rs2853676 in *TERT* and rs401681 and rs31489 in *CLPTM1L* had significant direct associations on lung adenocarcinoma without indirect effects through nicotine dependence (Tseng TS et. al., 2014). Also, Yin Z et. al., 2014 found that *TERT* polymorphism (rs2736100) might be a genetic susceptibility factor for lung cancer in non-smoking females in China. Yang J. and its team demonstrated there an association between rs2736100 and increased risk of lung cancer. Subgroup analysis by ethnicity demonstrated a significant association among both Asian and Caucasian populations. They additionally founded an increased risk of non-small cell lung cancer and lung adenocarcinoma strongly associated with rs2736100 (Yang J. et. al., 2014).

The other researches demonstrated that *TERT* rs2736100 polymorphism is a risk factor associated with increased lung cancer susceptibility, particularly for lung adenocarcinoma (Yuan Y. et. al., 2014; Nie W. et. al., 2014; Wu H. and Zhu R., 2014).

While the Liu SG. and his team suggested that rs2736100 on *TERT-CLPTM1L* indicates a poor prognosis for lung cancer in the Chinese Han population (Liu SG. et. al., 2015).

Table 4.22: Variant genotype with lung cancer risk

Genotype		Risk		*P-value*
		OR	**95%CI**	
rs2736100 G/T	GT+TT vs GG	4.33	1.53–12.29	0.005
	GT vs GG	4.82	1.53–15.14	0.005
	TT vs GG	2.89	0.38–22.04	0.287
	GT+GG vs TT	2.071	0.27–15.48	0.468

As previously pointed out, the mean number of cigarettes/day (CPD) was not significantly different between the smoker controls and cases (table 4.3, *P-value* = 0.823). However, the allelic frequency of the susceptibility allele was higher in the case than in the control (35% vs 18.3% respectively; table 4.21). So it can be consider as risk factor for lung cancer susceptibility apart of smoking.

Table 4. 23: Allele distribution among smoker cases and smoker controls

	Control (CPD= 31.0±12.5)	**Case (CPD= 30.2±13.2)**	
GG	17	11	
GT	12	12	0.366
TT	1	3	
Allele frequency	23.3%	34.6%	

The differences in distribution of the rs2736100 between the smoker control and smoker cases was not statistically significant (*P-value* = 0.366; table 4.23).

Because of the small sample size, analysis of statistical significance could not be applied to compare between non-smoker case (n=4) and non-smoker control (n=30).

By calculating the allelic frequency of the T allele in nonsmokers and smokers regardless of lung cancer, the frequency was higher in smokers (23.3%) than nonsmokers (13.3%). Furthermore, we also investigated rs2736100 in the context of smoking intensity among controls and cases, and

did not observed any association between number of cigarettes consumed per day and rs2736100 (*P- value* = 0.269, table 4.24).

Table 4.24: rs2736100 vs number of cigarettes per day (CPD)

SNP	CPD	*P-Value*
	ALL case and smoker control	
rs2736100		
GG	33.21±12.73	
GT	27.46±12.28	0.269
TT	31.25±14.36	

The distribution of rs2736100 in the case group according to the type of lung cancer (SCLC or NSCLC) is shown in table 4.25. The allelic frequency of the susceptibility allele “T” in SCLC (35.7%) was higher than in NSCLC (33.3%).

Table 4.25: Distribution of the rs2736100 between patients depends on lung cancer type.

SNP	Cancer Type	
	SCLC	NSCLC
GG	9 (42.8%)	3 (33.3%)
GT	9 (42.8%)	6 (66.7%)
TT	3 (14.4%)	0 (0.0%)
Allele Frequencies	35.7%	33.3%

However, the distribution of rs2736100 in cases group depending on the type of lung cancer (SCLC or NSCLC) was not statistically significant (*P*-value=0.366). Similarly, the distribution was not significant according to the status of ever suffering from pneumonia, (*P*-value=0.180).

4.2.5 rs9295740, G>A

Table 4.26 presents the distribution of cases and smoker and non-smoker controls by genotype of the SNP (rs9295740). The distribution of alleles among case and control is not statistically significant (*P*-value = 0.075).

The overall frequency of the risk allele "A" is 20.0%, and the alleles are in Hardy-Weinberg equilibrium (*P*-value= 0.3564), but the allele frequency according to ALFRED database was 24% in Palestinian population (ALFRED, 2014). The allele frequency was 40% in African American populations, 27% in Chinese in Beijing (CHB) population, 17% in Japanese population, 18% in Indian American population (GIH) and 20% in Italian populations (TSI, ensemble 2015).

The allelic frequency of the susceptibility allele is higher in the case group (30.0%) than in the control group (15.0%, table 4.26, figure 4.6F).

Table 4.26: The distribution of alleles among case and control

		Control (n=60)	**Case** (n=30)	**Total** (n=90)
rs9295740	G	44 (73.3%)	15 (50.0%)	59 (65.6%)
	GA	14 (23.3%)	12 (40.0%)	26 (28.9%)
	A	2 (3.4%)	3 (10.0%)	5 (5.5%)
Allele Frequency		15.0%	30.0%	20.0%

In order to determine the role of this SNP in conferring susceptibility to lung cancer the odds ratio was calculated for each pair of genotypes and listed in table 4.27. The A-allele, in homozygous or heterozygous genotype, was found to increase the risk for lung cancer when compared to the wild type genotype (OR= 2.75; table 4.27).

Table 4.27: Variant genotype with lung cancer risk

Genotype		**Risk**		***P-value***
		OR	**95%CI**	
rs9295740 G/A	GA+AA vs GG	2.75	1.10–6.87	0.026
	GA vs GG	2.51	0.95–6.62	0.052
	AA vs GG	4.40	0.67–28.91	0.099
	GA+GG vs AA	3.22	0.51–20.42	0.193

Other reports have identified the SNP rs9295740 at the region of chromosome 6p (6p22.1) to be associated with risk to lung cancer (Wang et. al., 2008). In favor of this result, three common variations (rs4324798, rs3117582, and rs9295740) on 6p21 were previously reported to increases the risk for lung cancer susceptibility with variable extents in different ethnic populations (TanTai J. et. al. 2014). The SNP was identified as a risk locus for lung cancer in Caucasian populations (Zhang M. et. al., 2010). Other reports however, failed to detect possible association between the SNP and lung cancer risk in Chinese populations and in Korean population (Zhang M. et. al., 2010; Bae EY. et. al., 2012).

The allelic distribution of the SNP between smoker cases and smoker controls was analyzed. The allelic frequency of the susceptibility allele was higher in the smoker cases than in the smoker controls (32.7% vs 10.0% respectively; table 4.28), and the distribution was statistically significant (P-value = 0.016). Recall that the mean number of cigarettes/day (CPD) was not significantly different between the smoker controls and cases (table 4.3, *P*-value = 0.823). Therefore, we may claim that the SNP is predisposing to lung cancer separately from smoking. Further support for this conclusion comes from analysis of the distribution of the alleles in nonsmokers and smokers irrespective of lung cancer; the frequency was higher in nonsmokers (20.0%) than smokers (10.0%) (Table 29, P-values=.266). Furthermore, smoking intensity, evaluated by number of cigarettes smoked per day, was not significantly different between cases and controls having different allelic genotypes (P- value = 0.269, table 4.30). Similarly, years of smoking were not significantly different in the different genotypes (P-value = 0.255; table 4.31).

Table 4. 28: Allele distribution among smoker cases and smoker controls

	Control (CPD= 31.0±12.5)	**Case (CPD= 30.2±13.2)**	***P-value***
GG	24	12	
GA	6	11	0.016
AA	0	3	
Allele frequency	10.0%	32.7%	

Table 4.29: Allele distribution among smoker and non-smoker controls

	Smoker Control	Non-Smoker Control	*P-value*
GG	24	20	
GA	6	8	0.266
AA	0	2	
Allele frequency	10.0%	20.0%	

Table 4.30: rs9295740 vs number of cigarettes per day (CPD)

SNP	CPD smoker case and control	*P-Value*
rs9295740		
GG	30.19±11.99	
GA	30.41±14.31	0.269
AA	36.67±15.28	

Table 4.31: rs9295740 vs years of smoking between smoker cases and controls

SNP	Years of Smoking smoker case and control	*P-Value*
rs9295740		
GG	32.31±12.25	
GA	32.76±13.23	0.255
AA	45.00±15.0	

The distribution of rs9295740 in the case group according to the type of lung cancer (SCLC or NSCLC) is shown in table 4.32. The allelic frequency of the susceptibility allele “A” in SCLC (33.3%) was higher than in NSCLC

(22.2%). However, the distribution of rs9295740 in cases grouped by the type of lung cancer (SCLC or NSCLC) was not statistically significant (*P*-value=0.490). Similarly, the distribution was not significant according to the status of ever suffering from lung infectious disease, (*P*-value=0.547).

In a previous study, the SNP rs9295740 G/A was found associated with NSCLC risk and it suggested that rs9295740 may be considered a prognostic biomarker for advanced NSCLC (de Mello RA. et. al., 2013).

Table 4.32: Distribution of the rs9295740 between patients depends on lung cancer type.

SNP	Cancer Type	
	SCLC	**NSCLC**
GG	10 (47.6%)	5 (55.6%)
GA	8 (38.1%)	4 (44.6%)
AA	3 (14.3%)	0 (0.0%)
Allele Frequencies	33.3%	22.2%

4.2.6 rs4324798, G>A

The genotype of the SNP (rs4324798) in the cases and the unrelated smoker and non-smoker controls is depicted in table 4.33, figure 4.6D. The distribution of alleles among case and control is statistically significant (*P*-value = 0.000). The allelic frequency of the allele “A” is lower in the case group (0.0%) than in the control group (20.0%, table 4.33).

The overall frequency of the “A” allele is 13.3% and its distribution among the entire population is in Hardy-Weinberg equilibrium (*P*-value= 1.444). In comparison, the documented frequency was 9% in African (AFR) populations, 2% in American (AMR) populations, 0% East Asian (EAS) population, 7% in European (EUR) populations, 0% in South Asian (SAS) populations (Ensemble 2015) and 2% in Palestinian populations (ALFRED 2014).

Table 4.33: The distribution of alleles among case and control

		Control (n=60)	**Case** (n=30)	**Total** (n=90)
rs4324798	GG	36 (60.0%)	30(100.0%)	66 (65.6%)
	GA	24 (40.0%)	0 (0.0%)	24 (28.9%)
	AA	0 (0.0%)	0 (0.0%)	0 (0.0%)
Allele Frequency		20.0%	0.0%	13.3%

In order to examine the potential role of the "A" allele in lung cancer we calculated the odds ratio for homo- or heterozygousity compared to the wild type allele (table 4.34). The allele 'A' was not found to be compatible with increasing the risk for lung cancer (OR= 0.55). Conversely, its absence from the case group suggests a protective role against lung cancer. Different literature reports concluded that the "A" allele is associated with predicting the overall survival of lung cancer patients and particularly of SCLC patients (Yang P. et. al. 2010, Xun W. et. al., 2011).

Table 4.34: Variant genotype with lung cancer risk

Genotype		**Risk**		***P-value***
		OR	**95%CI**	
rs4324798 G/A	GA+AA vs GG	0.55	0.44-0.68	0.000
	GA vs GG	0.55	0.44-0.68	0.000

Similar results were expectedly obtained by comparing the allelic distribution of the SNP between smoker cases and smoker controls (Table 4.35).

Table 4. 35: Allele distribution among smoker cases and smoker controls

	Control (CPD= 31.0±12.5)	**Case (CPD= 30.2±13.2)**	***P-value***
GG	18	26	
GA	12	0	0.000
AA	0	0	
Allele frequency	20.0%	0.0%	

By calculating the allelic frequency of the A allele in nonsmokers and smokers regardless of lung cancer, the frequency was the same in smokers and nonsmokers (20.0%) (Table: 4.36).

Table 4.36: Distribution of rs4324798 among controls

SNP	Controls		*P-Value*
	Smoker	Non-Smoker	
rs9295740			
GG	18	18	0.604
GA	12	12	
Allele Frequency	20 %	20 %	

Furthermore, we also investigated rs4324798 in the context of smoking intensity among controls and cases, and we did not observed any statistically significant difference between genotypes of the SNP rs4324798 and the number of cigarettes consumed per day (*P- value* = 0.699, table 4.37), or the years of smoking (*P-value* = 0.064; table 38). These results exclude any relationship between the SNP alleles and smoking.

Table 4.37: rs4324798 vs number of cigarettes per day (CPD)

SNP	CPD	*P-Value*
	ALL case and smoker control	
rs4324798		
GG	30.95±13.62	0.699
GA	29.33±8.99	

Table 4.38: rs4324798 vs years of smoking between smoker cases and controls

SNP	Years of Smoking smoker case and control	*P-Value*
rs9295740		
GG	34.77±13.30	0.064
GA	27.08±8.47	

Because the allele was not detected in the cases group, its relation to the type of cancer could not be discussed.

Chapter Five

Conclusion and Recommendation

5.1 Conclusion

Like other cancers, the lung cancer is a multifactorial complex disease in which environmental, genetic and epigenetic factors may interplay and contribute to the disease etiology. The genetic component is frequently polygenic in which sequence changes in a number of loci may be important in increasing the relative risk for the disease. The role of cigarette smoking is the most well-established trigger of lung cancer disease. However, only a subset of smokers will experience the disease during their lifetime, indicating an inter-individual variability in susceptibility to such a potent carcinogen as well as the possible existence of a role for the underlying genetic variation. A number of loci have been shown by GWAS to be implicated in conferring a potential risk for lung cancer. Therefore, this follow up study was designed to investigate whether the results of previous GWAS could be replicated in the Palestinian population of Gaza strip. Six SNPs in three genomic regions were investigated in this study because of their already established involvement in different populations. The relationship between these SNPs and smoking has been investigated. To the best of our knowledge no previous studies have tackled this subject before among the Palestinian population in Gaza strip.

Results of the study have led to formulate the following conclusions:

1. In our study population, the six investigated SNPs (rs16969968 and rs1051730 in 15q25; rs402710 and rs2736100 in 5p15.33; rs9295740 and rs4324798 in 6p22) were found to be related to lung cancer in different aspects.

    a. The SNPs rs16969968 and rs1051730 are associated with increasing the risk for lung cancer via increasing the individual's tendency to smoke more and at an earlier age.

    b. The SNP rs402710 was found associated with increasing the risk for lung cancer with no relation to smoking.

c. The SNPs rs2736100 and rs9295740 were found to increase the risk for lung cancer with no relation to smoking.

d. The SNP rs4324798 was not detected in the case group indicating a possible protective role against lung cancer.

2. The allelic frequency of the 6 SNPs rs16969968, rs1051730, rs402710, rs2736100, rs9295740 and rs4324798 fill within the worldwide range in different populations.

3. In consistence with other studies, 87.7% of the cases were smokers, highlighting the important role of smoking as non-genetic risk factors for lung cancer development.

   a. In agreement with other studies, the duration of smoking is far more important than the amount of cigarettes smoked in conferring risk for lung cancer.

4. The history of lung infections represents a potential risk factor for lung cancer, a result that is in agreement with other studies.

5.2 Recommendation

1 Public awareness campaigns should be carried out to reduce the economic and health burden of smoking particularly in teenagers.

2 The concerned health care personnel should be aware of the possible implications of gene-environment interaction in increasing risk to complex diseases.

3 With more solid and comprehensive results as well as deeper understanding of the influence of the genetic variations on lung cancer clinical outcomes, these identified markers could be incorporated into a prognosis prediction model to increase prediction accuracy in both population and individual level.

4 Further studies are recommended in the following areas:

- Replication of the current study in larger cohorts of lung cancer patients particularly none-smoker ones versus non-smoker controls to eliminate the effect of smoking in risk estimation.
- To investigate the role of nicotine addiction in the pathology of other serious diseases.
- Extend the search for other regions that are also important for the understanding of lung cancer susceptibility risk.

Chapter Six

Reference

Al Moustafa A., (2012). Cigarette Smoking and Lung Cancer: Pediatric Roots. Lung Cancer International, (2012): 790841-7.

Alberg AJ. and Samet JM., (2003). Epidemiology of lung cancer. Chest, 132:29S–55S.

Amos C., Wu X, Spitz M., (1999): Is there a genetic basis for lung cancer susceptibility? Recent Results Cancer Res,151:3–12.

Amos C., Wu X., Broderick P., Gorlov IP., Gu J., (2008): Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet, 40(5):616-22.

Asakura T, Imai A, Ohkubo-Uraoka N, Kuroda M, Iidaka Y., (2005): Relationship between expression of drug-resistance factors and drug sensitivity in normal human renal proximal tubular epithelial cells in comparison with renal cell carcinoma. Oncol Rep, 14: 601–607.

Bae EY., Lee SY., Kang BK., Lee EJ., Choi YY., (2012): Replication of results of genome-wide association studies on lung cancer susceptibility loci in a Korean population. Respirology, 17(4):699-706.

Bierut LJ., Stitzel JA, Wang JC, Hinrichs AL, Grucza RA., (2008): Variants in nicotinic receptors and risk for nicotine dependence. Am J Psychiatry, 165(9):1163-71.

Blanchet MR., Israël-Assayag E., Cormier Y., (2004): Inhibitory effect of nicotine on experimental hypersensitivity pneumonitis in vivo and in vitro. Am J Respir Crit Care Med., 169:903–909.

Blasco MA., (2005): Telomeres and human disease: ageing, cancer and beyond. Nat Rev Genet, 6:611-622.

Brenner DR., McLaughlin JR., Hung RJ., (2011): Previous Lung Diseases and Lung Cancer Risk: A Systematic Review and Meta-Analysis. PLoS ONE, 6(3):e17479.

Buckland PR., (2006): The importance and identification of regulatory polymorphisms and their mechanisms of action. Biochim Biophys Acta, 1762(1):17-28.

CancerCare. (2005): US Center for Disease Control. Annual smoking-attributable mortality, years of potential life lost, and productivity losses -- United States, 1997–2001. MMWR Morb Mortal Wkly Rep., 54:625–8.

Chanock SJ., Hunter DJ. (2008): Genomics: when the smoke clears. Nature, 452(7187): 537–538.

Chen LS., Rayjean H., Timothy B., (2015). CHRNA5 Risk Variant Predicts Delayed Smoking Cessation and Earlier Lung Cancer Diagnosis—A Meta-Analysis. JNCI J Natl Cancer Inst., 107 (5): 100.

Chen RJ., Chang LW., Lin P., Wang YJ., (2011). Epigenetic Effects and Molecular Mechanisms of Tumorigenesis Induced by Cigarette Smoke: An Overview. Oncology, 2011:654931, 14.

Chen XF., Cai S., Chen QG., Ni ZH., Tang JH., (2012): Multiple variants of TERT and CLPTM1L constitute risk factors for lung adenocarcinoma. Genet Mol Res, 11: 370–378.

Chorley BN., Wang X., Campbell MR., Pittman GS., Noureddine MA., Bell DA., (2008): Discovery and verification of functional single nucleotide polymorphisms in regulatory genomic regions: current and developing technologies. Mutat Res, 659(1-2):147-157.

Christiani DC. (2006): Genetic susceptibility to lung cancer. J Clin Oncol, 24: 1651-2.

Colombo J., Fachel AA., De Freitas Calmon M., Cury PM., Fukuyama EE., (2009): Gene expression profiling reveals molecular marker candidates of laryngeal squamous cell carcinoma. Oncol Rep; 21: 649–663.

Cooper S. and Spiro SG., (2006): Small cell lung cancer: treatment review. Respirology, 11:241-8.

Crispo A., Brennan P., Jockel KH., Schaffrath-Rosario A., Wichmann HE., (2004): The cumulative risk of lung cancer among current, ex- and never-smokers in European men. Br J Cancer, 91(7):1280-6.

de Mello RA., Ferreira M., Soares-Pires F., Costa S., Cunha J., (2013): The impact of polymorphic variations in the 5p15, 6p12, 6p21 and 15q25 Loci on the risk and prognosis of portuguese patients with non-small cell lung cancer. PLoS One, 8(9):e72373.

Denholm R., Schüz J., Straif K., Stücker I., Jöckel KH., (2014): Is previous respiratory disease a risk factor for lung cancer? Am J Respir Crit Care Med., 1;190 (5):549-59.

Doll R., Peto R., Boreham J., Sutherland I., (2005): Mortality from cancer in relation to smoking: 50 years observations on British doctors. Br J Cancer, 92(3):426-29.

Flanders WD., Lally CA., Zhu BP., Henley SJ., and Thun MJ., (2003). Lung cancer mortality in relation to age, duration of smoking, and daily cigarette consumption: results from cancer prevention study II. Cancer Research, (63) 19: 6556–6562

Flanders WD., Lally CA., Zhu BP., Henley SJ., Thun MJ., (2003). Lung cancer mortality in relation to age, duration of smoking, and daily cigarette consumption: results from Cancer Prevention Study II. Cancer Res., 63(19):6556-62.

Gabrielsen E., Romundstad P., Langhammer A., Krokan H., Skorpen F., (2013). Association between a 15q25 gene variant, nicotine-related habits, lung cancer and COPD among 56 307 individuals from the HUNT study in Norway. Human Genetics, (21)1293–1299.

Gago-Dominguez M., Jiang X., Conti DV., Castelao JE., Stern MC., (2011): Genetic variations on chromosomes 5p15 and 15q25 and bladder cancer risk: findings from the Los Angeles-Shanghai bladder case control study. Carcinogenesis, 32:197-202.

Gago-Dominguez M., Jiang X., Conti DV., Castelao JE., Stern MC., (2011). Genetic variations on chromosomes 5p15 and 15q25 and bladder cancer risk: findings from the Los Angeles-Shanghai bladder case-control study. Carcinogenesis, 32(2): 197–202.

Gahring LC., Rogers SW., (2006): Neuronal nicotinic acetylcholine receptor expression and function on nonneuronal cells. AAPS Journal, 7:E885–E894.

Gillis AJ., Schuller AP., Skordalakes E., (2008): Structure of the Tribolium castaneum telomerase catalytic subunit TERT. Nature, 455:633-637.

Girard N., Lou E., Azzoli CG., Reddy R., Robson M., Harlan M., (2010).. Analysis of genetic variants in never-smokers with lung cancer facilitated by an internet-based blood collection protocol: A preliminary report. Clinical Cancer Research, 16:755-763.

Grivennikov SI., Greten FR., Karin M., (2010): Immunity, inflammation, and cancer. Cell, 140:883–899.

Groot Kormelink PJ., Luyten L. Cloning and sequence of full-length cDNAs encoding the human neuronal nicotinic acetylcholine receptor (nAChR) subunits beta-3 and beta-4 and expression of seven nAChR subunits in

the human neuroblastoma cell line SH-SY5Y and/or IMR-32. FEBS Lett., 400: 309-314.

Gu M., Dong X., Zhang X., Wang X., Yue Q., Jun Y., Wenquan N., (2012) : Strong Association between Two Polymorphisms on 15q25.1 and Lung Cancer Risk: A Meta-Analysis. PLoS ONE, 7(6).

Hecht S., and Samet M., (2007): Cigarette smoking. In: Rom, W., editor. Environmental and Occupational Medicine. Vol. 4th. Wolters Kluwer; Philadelphia.

Hecht SS., (2003): Tobacco carcinogens, their biomarkers and tobacco induced cancer. Nat Rev Cancer, 3: 733-744.

Herbst RS., Heymach JV., Lippman SM., (2008): Lung cancer. N Engl J Med, 359(13):1367- 80.

here:http://www.isdscotland.org/Health-Topics/Cancer/Publications/index.asp. Last accessed on 07.05.2015.

Hong E., Hodgkinson A., Yang Y., Sampath H., Ross J., (2010): A genetically modulated, intrinsic cingulate circuit supports human nicotine addiction. Proc. Nat. Acad. Sci., 107: 13509-13514.

Hsiung C., Lan Q., Hong Y., Chen C., Hosgood H., (2010): The 5p15.33 Locus Is Associated with Risk of Lung Adenocarcinoma in Never-Smoking Females in Asia. PLoS Genet, 6(8): e1001051.

http://alfred.med.yale.edu/alfred/recordinfo.asp?condition=sites.site_uid=%27SI624401R, Last access on 21/10/2015.

http://alfred.med.yale.edu/alfred/SiteTable1A_working.asp?siteuid=SI354084Y, Last access 25.10.2015.

http://alfred.med.yale.edu/alfred/SiteTable1A_working.asp?siteuid=SI400886A, Last access 10.07.2015

http://alfred.med.yale.edu/alfred/SiteTable1A_working.asp?siteuid=SI624401R, Last access 25.08.2015

http://asia.ensembl.org/Homo_sapiens/Variation/Population?db=core;v=rs402710;vdb=variation#373431_tablePanel, Last access 31.08.2015

http://asia.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=6:27721223-27722223;v=rs9295740;vdb=variation;vf=106094176, Last access 25.08.2015

http://globocan.iarc.fr/factsheet.asp. Last accessed 10.09.2011.

http://globocan.iarc.fr/factsheet.asp. Last accessed on 15.01.2014.

http://seer.cancer.gov/statfacts/html/lungb.html. last accessed on 05.10.2015.

http://www.cancer.org/cancer/cancercauses/geneticsandcancer/heredity-and-cancer. Last access on 22.10.2015.

http://www.cdc.gov/cancer/lung/basic_info/risk_factors.htm#22. Last accessed on 02/10/2015.

http://www.cdc.gov/cancer/lung/basic_info/risk_factors.htm. Last acsecced 7.12.2015.

http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=2736100. Last accessed on 25.10.2015.

http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=rs9295740. Last accessed on 21.10.2015.

http://www.ons.gov.uk/ons/rel/vsob1/cancer-statistics-registrations--england--series-mb1-/index.html. Last accessed on 8.08.2015.

http://www.org/int/cancer-facts-figures/worldwide-data, Last accessed on 15/10/2015.

http://www.qub.ac.uk/research-centres/nicr/CancerData/OnlineStatistics/. Last accessed on 10.06.2015

http://www.wales.nhs.uk/sites3/page.cfm?orgid=242&pid=59080. Last accessed on 18.06.2015.

http://www.who.int/mediacentre/factsheets/fs297/en/, Last accessed on 09.10.2015

https://www.pharmgkb.org/variant/rs1051730?previousQuery=rs1051730#tabview=tab2&subtab=21. Last accessed 07.12.2015.

https://www.pharmgkb.org/variant/rs16969968?previousQuery=rs16969968#tabview=tab2&subtab=21. Last accessed 07.12.2015

Hu Z., Chen J., Tian T., Zhou X., Gu H., Xu L., (2008): Genetic variants of miRNA sequences and non-small cell lung cancer survival. J Clin Invest., 118(7):2600-2608.

Hu Z., Wu C., Shi Y., Guo H., Zhao X., (2011): A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. Nat Genet, 43: 792–796.

Hung RJ., McKay JD., Gaborieau V., Boffetta P., Hashibe M., Zaridze D., (2008): A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature, 452(7187):633-637.

Ibrahim EM., Kazkaz GA., Abouelkhair KM., Al-Mansour MM., Al-Fayea TM., (2013): Increased risk of second lung cancer in Hodgkin's lymphoma survivors: a meta-analysis. Lung, 191(1):117-34).

Ito H., McKay JD., Hosono S., Hida T., Yatabe Y., (2012): Association between a Genome-Wide Association Study-Identified Locus and the Risk of Lung Cancer in Japanese Population. Thoracic Oncology, 7: 790–798.

Janssen-Heijnen ML and Coebergh JW. (2003): The changing epidemiology of lung cancer in Europe. Lung Cancer, 41(3): 245-58.

Jaworowska E., Trubicka J., Lener MR., Masojc B., Zlowocka-Perlowska E., (2011): Smoking related cancers and loci at chromosomes 15q25, 5p15, 6p22.1 and 6p21.33 in the Polish population. PLoS One, 6: e25057.

Jemal A., Siegel R., Ward E., (2008): Cancer statistics 2008. CA: a cancer journal for clinicians, 58:71–96.

Jiang M., Wu H., Qin C. (2013): Genetic variant rs401681 at 5p15.33 modifies susceptibility to lung cancer but not esophageal squamous cell carcinoma. PLoS One, 8(12):e84277.

Jin G., Xu L., Shu Y., Tian T., Liang J., (2009): Common genetic variants on 5p15.33 contribute to risk of lung adenocarcinoma in a Chinese population. Carcinogenesis, 30:987-990.

Keskitalo K., Broms U., Heliovaara M., Ripatti S., Surakka I., Perola M., Pitkaniemi J., Peltonen L., Aromaa A., Kaprio J., (2009): Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNB4) on chromosome 15. Hum. Molec. Genet., 18: 4007-4012.

Knoke D., Shanks G., Vaughn W., Thun MJ., Burns DM., (2004): Lung cancer mortality is related to age in addition to duration and intensity of cigarette smoking: an analysis of CPS-I data. Cancer Epidemiol Biomarkers Prev., 13(6):949-57.

Landi T., Hsiung C., Matsuo K., Hong Y., Seow A., (2009): A genome-wide association study of lung cancer identifies a region of chromosome 5p15

associated with risk for adenocarcinoma. Am J Hum Genet, 85(5):679-91.

Lantuejoul S., Salon C., Soria JC., Brambilla E., (2007): Telomerase expression in lung preneoplasia and neoplasia. Int J Cancer, 120:1835-1841.

Le Merchand, ML., Derby KS., Murphy SE., Hecht SS., Hatsukami D., Carmella SG., (2008): Smokers with the CHRNA lung cancer associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. Cancer Res., 68,9137-9140.

Leung T., Bergen A., Munafò MR., De Ruyck K., Selby P., De Luca V., (2015). Effect of the rs1051730-rs16969968 variant and smoking cessation treatment: a meta-analysis. Pharmacogenomics, 16(7):713-20.

Li C., Yin Z., Wu W., Li X., Ren Y., Zhou B. (2013): Genetic variations in TERT-CLPTM1L genes and risk of lung cancer in Chinese women nonsmokers. PLoS One, 8(5):e64988.

Li Y., Sheu CC., Ye Y., de Andrade M., Wang L., Chang SC., (2010): Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. Lancet Oncol, 11(4):321-330.

Liu SG., Ma L., Cen QH., Huang JS., Zhang JX., Zhang JJ. (2015): Association of genetic polymorphisms in TERT-CLPTM1L with lung cancer in a Chinese population. Genet Mol Res., 14(2):4469-76.

Lu X., Ke J., Luo X., Zhu Y., Zou L., (2013): The SNP rs402710 in 5p15.33 Is Associated with Lung Cancer Risk: A Replication Study in Chinese Population and a Meta-Analysis. PLoS ONE, 8(10): e76252.

Machiela MJ., Hsiung CA., Shu XO., Seow WJ., Wang Z., (2015): Genetic variants associated with longer telomere length are associated with increased lung cancer risk among never-smoking women in Asia: a report from the female lung cancer consortium in Asia. Int J Cancer, 137(2):311-9.

McKay D., Hung RJ., Gaborieau V., Boffetta P., Chabrier A., (2008): Lung cancer susceptibility locus at 5p15.33. Nature Genetics, 40: 1404–1406.

Miki D., Kubo M., Takahashi A., Yoon KA., Kim J., (2010): Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. Nat Genet, 42: 893–896.

Miyagawa T., Nishida N., Ohashi J., Kimura R., Fujimoto A., Kawashima M., Koike A., Sasaki T., Tanii H., Otowa T., (2008). Appropriate data

cleaning methods for genomewide association study. Journal of human genetics, 53:886-893.

Munafò R. (2012). Genetic Variants, Tobacco Exposure and Lung Cancer Risk. JNCI J Natl Cancer Inst, 104 (10).

Ni Z., Tao K., Chen G., Chen Q., Tang J., (2012): CLPTM1L Is Overexpressed in Lung Cancer and Associated with Apoptosis. PLoS ONE, 7(12): e52598.

Nie W., Zang Y., Chen J., Xiu Q. (2014): TERT rs2736100 polymorphism contributes to lung cancer risk: a meta-analysis including 49,869 cases and 73,464 controls. Tumour Biol., 35(6):5569-74.

Okazaki I., Ishikawa S., Sohara Y., (2014). Genes associated with succeptibility to lung adenocarcinoma among never smokers suggest the mechanism of disease. Anticancer Res., 34(10):5229-40.

Palestinian Health Information Center. 2009, 2011 and 2014. Health annual report. Palestinian Ministry of Health.

Pande M., Spitz MR., Wu X., Gorlov IP., Chen WV., (2011):. Novel genetic variants in the chromosome 5p15.33 region associate with lung cancer risk. Carcinogenesis, 32: 1493–1499.

Parkin DM., (2010): Tobacco-attributable cancer burden in the UK in 2010. Br J Cancer, 105(S2):S6-S13.

Peto, R., (1986). Influence of dose and duration of smoking on lung cancer rates. In: Zaridze, D. & Peto, R., eds, Tobacco: A Growing International Health Hazard (IARC Scientific Publications No. 74), Lyon, IARCPress, pp. 23–33.

Pirie K., Peto R., Reeves GK., Green J., Beral V., (2013). The 21st century hazards of smoking and benefits of stopping: a prospective study of one million women in the UK. Lancet. 2013 Jan 12;381(9861):133-41.

Punturieri A., Szabo E., Croxton TL., Shapiro SD., Dubinett SM., (2009): Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. J Natl Cancer Inst., 101:554–559.

Rodgman A., Perfetti TA., (2009). Alphabetical Component Index. In: The Chemical Components of Tobacco and Tobacco Smoke. Boca Raton, FL: CRC Press, pp. 1483–1784.

Rothman N., Wacholder S., Caporaso N., Garcia-Closas M., Buetow K., Fraumeni J., (2001): The use of common genetic polymorphisms

to enhance the epidemiologic study of environmental carcinogens. Biochim Biophys Acta., 1471:C1–10.

Rubin DT., Hanauer SB., (2000): Smoking and inflammatory bowel disease. Eur J Gastroenterol Hepatol, 12:855–862.

Saccone F., Hinrichs AL., Saccone NL., Chase GA., Konvicka K., (2007): Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Hum. Molec. Genet., 16: 36-49.

Sherva R., Wilhelmsen K., Pomerleau CS., Chasse SA., Rice JP., (2008): Association of a single nucleotide polymorphism in neuronal acetylcholine receptor subunit alpha 5 (CHRNA5) with smoking status and with 'pleasurable buzz' during early experimentation with smoking; Addiction, 1544-52.

Shiraishi K., Kohno T., Kunitoh H., Watanabe S., Goto K., Nishiwaki Y., (2009). Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. Carcinogenesis, 30(1):65-70.

Shivji M., Burger S., Moncada CA., Clarkson ABJ., Merali S., (2005): Effect of nicotine on lung S-adenosylmethionine and development of Pneumocystis pneumonia. J Biol Chem., 280:15219–15228.

Shopland DR., Eyre HJ., Pechacek TF., (1991). Smoking-attributable cancer mortality in 1991: is lung cancer now the leading cause of death among smokers in the United States?. J. Natl. Cancer Inst., 83, 1142–1148.

Spitz R., Wei Q., Dong Q., Amos I., Wu X., (2003): Genetic susceptibility to lung cancer: the role of DNA damage and repair. Cancer Epidemiol Biomarkers Prev., 12(8):689-698.

Spitz, MR., Amos CI., Dong Q., Lin J., Wu X., (2008): The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. J. Natl. Cancer Inst., 100.1552-1556.

Stampfil MR. and Anderson GP. (2009): How cigarette smoke skews immune responses to promote infection, lung disease and cancer. Nature Reviews Immunol., 9:377-84.

Stellman SD. and Djordjevic MV., (2009): Monitoring the tobacco use epidemic II: The agent: Current and emerging tobacco products. Prev Med Jan, 48(1 Suppl):S11-5.

Strachan T., and Read A., (1999). Human molecular genetics (2nd ed). Singapore, New York. Wiley. pp. 576.

Stratton MR., and Rahman N., (2008). The emerging landscape of breast cancer susceptibility. Nat Genet., 40(1):17-22.

Takahashi H., Ogata H., Nishigaki R., Broide D., Karin M., (2010). Tobacco smoke promotes lung tumorigenesis by triggering IKKβ and JNK1 dependent inflammation. Cancer Cell, 19; 17(1): 89.

TanTai J., Shen Y., Zhao H., (2014): Quantitative assessment of the influence of common variations on 6p21 and lung cancer risk. Tumour Biol, 35(1):689-94.

Thorgeirsson TE., Geller F., Sulem P., Rafnar T., Wiste A., Magnusson KP., Manolescu A., Thorleifsson G., Stefansson H., Ingason A., Stacey SN., Ber., (2008): A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature, 452(7187):638-42.

Thun MJ., Henley SJ., Burns D., Jemal A., Shanks TG., Calle EE., (2006): Lung cancer death rates in lifelong nonsmokers. Journal of the National Cancer Institute, 98:691–9.

Timofeeva N., Hung J., Rafnar., (2012). Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls. Hum Mol Genet., 21(22): 4980–4995.

Truong T., Hung RJ., Amos CI., Wu X., Bickeböller H., (2010): Replication of Lung Cancer Susceptibility Loci at Chromosomes 15q25, 5p15, and 6p21: A Pooled Analysis From the International Lung Cancer Consortium. JNCI, (102)13.

Tseng TS., Park JY., Zabaleta J., Moody-Thomas S., Sothern MS., Chen T, Evans DE., Lin HY. (2014): Role of nicotine dependence on the relationship between variants in the nicotinic receptor genes and risk of lung adenocarcinoma. PLoS One, 9(9):e107268.

Turnbull C., and Rahman N., (2008). Genetic predisposition to breast cancer: past, present, and future. Annu Rev Genomics Hum Genet., 9:321-45.

Tyndale RF., Zhu AZX., George TP., Cinciripini P., Hawk LW., Schnoll RA., (2015). Lack of Associations of CHRNA5-A3-B4 Genetic Variants with Smoking Cessation Treatment Outcomes in Caucasian Smokers despite Associations with Baseline Smoking. PLoS ONE, 10(5): e0128109. doi:10.1371.

Vandenbroucke JP., (2009). Commentary: 'Smoking and lung cancer'--the embryogenesis of modern epidemiology. Int J Epidemiol., 38(5):1193-6.

Varghese JS., and Easton DF., (2010). Genome-wide association studies in common cancers--what have we learnt? Current opinion in genetics & development, 20:201-209.

Wang H., Zhao Y., Ma J., Zhang G., Mu Y., Qi G., Fang Z., (2013): The genetic variant rs401681C/T is associated with the risk of non-small cell lung cancer in a Chinese mainland population. Genet Mol Res., 12(1):67-73.

Wang HM., Zhang XY., Jin B. (2013): TERT genetic polymorphism rs2736100 was associated with lung cancer: a meta-analysis based on 14,492 subjects. Genet Test Mol Biomarkers, 17(12):937-41.

Wang Y., Broderick P., Matakidou A., Eisen T., Houlston S., (2011): Chromosome 15q25 (CHRNA3-CHRNA5) Variation Impacts Indirectly on Lung Cancer Risk. PLoS ONE, 6(4): e19085.

Wang Y., Broderick P., Webb E., Wu X., Vijayakrishnan J., Matakidou A., Qureshi M., (2008): Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat. Genet., 40:1407–1409.

Wang Y., Peter B., Athena M., Timothy E. and Houlston R., (2010): Role of 5p15.33 (TERT-CLPTM1L), 6p21.33 and 15q25.1 (CHRNA5-CHRNA3) variation and lung cancer risk in never-smokers. Carcinogenesis, (31)2: 234–238.

Wei Q., and Spitz M., (1997). Cancer metastasis rev. 16:295-307.

Winterer G., Mittelstrass K., Giegling I., Lamina C., Fehr C., Brenner H., Breitling LP., Nitz B., Raum E., Müller H., Gallinat J., Gal A., Heim K., (2010). Risk gene variants for nicotine dependence in the CHRNA5-CHRNA3-CHRNB4 cluster are associated with cognitive performance. Am J Med Genet B Neuropsychiatr Genet., 5:1448–58.

Wojas-Krawczyk K., Krawczyk P., Biernacka B., Grzybek M., Kołodziej P., (2012). The polymorphism of the CHRNA5 gene and the strength of nicotine addiction in lung cancer and COPD patients. Cancer Prevention, 21:111–117.

Wu H. and Zhu R., (2014): Quantitative assessment of common genetic variants on chromosome 5p15 and lung cancer risk. Tumour Biol., 35(6):6055-63.

Wu X., Dave J., Jiang H., Pathak S., Spitz R., (1997): Lung carcinoma patients with a family history of cancer and lymphocyte primary chromosome 9 aberrations. Cancer, 15;79(8):1527-32.

Wu X., Zhao H., Suk R., Christiani D., (2004): Genetic susceptibility to tobacco-related cancer. Oncogene, 23:6500–6523.

Xun W., Brennan P., Tjonneland A., Vogel U., Overvad K., Kaaks R., (2011): Single-nucleotide polymorphisms (5p15.33, 15q25.1, 6p22.1, 6q27 and 7p15.3) and lung cancer survival in the European Prospective Investigation into Cancer and Nutrition (EPIC). Mutagenesis, 26(5):657-66.

Xun X., Wang H., Yang H., Wang H., Wang B., Kang L., Jin T, Chen C., (2014): CLPTM1L genetic polymorphisms and interaction with smoking and alcohol drinking in lung cancer risk: a case-control study in the Han population from northwest China. Medicine (Baltimore), 93(28):e289.

Yamamoto K., Okamoto A., Isonishi S., Ochiai K., Ohtake Y., (2001): A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. Biochem Biophys Res Commun, 280: 1148–1154

Yang J. and Jiao S. (2014): Increased lung cancer risk associated with the TERT rs2736100 polymorphism: an updated meta-analysis. Tumour Biol., 35(6):5763-9.

Yang P. et. al., (2010): A Rigorous and Comprehensive Validation: Common Genetic Variations and Lung Cancer. Cancer Epidemiol Biomarkers Prev., 19(1): 240–244.

Yin Z., Cui Z., Ren Y., Zhang H., Yan Y., Zhao Y., Ma R., Wang Q., (2014): Genetic polymorphisms of TERT and CLPTM1L, cooking oil fume exposure, and risk of lung cancer: a case-control study in a Chinese non-smoking female population. Med Oncol., 31(8):114.

Yoon KA., Park JH., Han J., Park S., Lee GK., (2010): A genome-wide association study reveals susceptibility variants for non-small cell lung cancer in the Korean population. Hum Mol Genet, 19: 4948–4954.

Youlden DR., Cramb SM., Baade PD., (2008): The International Epidemiology of Lung Cancer: Geographical Distribution and Secular Trends. J Thoracic Oncol, 3(8): 819-31.

Yuan Y., Lu C., Xue L., Ge D., (2014): Association between TERT rs2736100 polymorphism and lung cancer susceptibility: evidence from 22 case-control studies. Tumour Biol., 35(5):4435-42.

Zhan P., Suo L-j., Qian Q., Shen XK., Qiu LX., Yu LK., Song Y., (2011): Chlamydia pneumoniae infection and lung cancer risk: A meta-analysis. Eur J Cancer, 47(5):742-47.

Zhang A., Zheng C., Lindvall C., Hou M., Ekedahl J., Lewensohn R., Yan Z., Yang X., Henriksson M., Blennow E., Nordenskjold M., Zetterberg A., Bjorkholm M., Gruber A., Xu D., (2000): Frequent amplification of the telomerase reverse transcriptase gene in human tumors. Cancer Res., 60:6230-6235.

Zhang M., Hu L., Shen H., Dong J., Shu Y., Xu L., Jin G., Tian T., Hu Z. and Shen H., (2010): Candidate variants at 6p21.33 and 6p22.1 and risk of non-small cell lung cancer in a Chinese population. Int J Mol Epidemiol Genet, 1:11-18.

Zhao DP., Yang CL., Zhou X., Ding JA., Jiang GN., (2014). Association between CLPTM1L polymorphisms (rs402710 and rs401681) and lung cancer susceptibility: evidence from 27 case-control studies. Mol Genet Genomics, 289(5):1001-12.

Ziegler A., Konig IR., and Thompson JR., (2008). Biostatistical aspects of genome-wide association studies. Biometrical, 50:8-28.

Zienolddiny S., Skaug V., Landvik NE., Phillips DH., Houlston R., (2009): The TERT-CLPTM1L lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. Carcinogenesis, 30(8): 1368–1371.

Zou P., Gu A., Ji G., Zhao L., Zhao P., Lu A., (2012): The TERT rs2736100 Polymorphism and Cancer Risk: A Meta-analysis Based on 25 Case-Control Studies. BMC Cancer, 12:7

APPENDIX

المجلس الفلسطيني للبحث الصحي

Palestinian Health Research Council

تعزيز النظام الصحي الفلسطيني من خلال مأسسة استخدام المعلومات البحثية في صنع القرار

Developing the Palestinian health system through institutionalizing the use of information in decision making

Helsinki Committee
For Ethical Approval

Date: 28\10\2013 **Number:** PHRC/HC/51 /13

Name: Rami Al-Masri الاسم: رامي المصري

We would like to inform you that the committee had discussed the proposal of your study about: نفيدكم علماً بأن اللجنة قد ناقشت مقترح دراستكم حول:-

"Genetic Polymorphisms and Lung Cancer Susceptibility between Palestinian Population (Gaza Strip)".

The committee has decided to approve the above mentioned research. Approval number PHRC/HC/51/13 in its meeting on 28/10/2013 و قد قررت الموافقة على البحث المذكور عاليه بالرقم والتاريخ المذكوران عاليه

Signature

Member **Member**

Chairman

28 10 2013

Genral Conditions:-
1. Valid for 2 years from the date of approval.
2. It is necessary to notify the committee of any change in the approved study protocol.
3. The committee appreciates receiving a copy of your final research when completed.

Specific Conditions:-
1. Inform consent from patients.
2. Authorized person to collect blood samples

The subject was approved following the World Medical Association Declaration of Helsinki-Ethical principles for medical research involving human subjects, adopted by the 18th World Medical Association General Assembly, Helsinki, Finland, June 1964 and amended by the 59th WMA General Assembly, Seoul, Korea, October 2008.

E-Mail:pal.phrc@gmail.com

Gaza - Palestine غزة - فلسطين
شارع النصر - مفترق العيون

FROM : FAX NO. :

The Palestinian National Authority
Ministry of Health
Directorate General of Human Resources Development

ة القـوى البشـريـة

الرقم .....

التاريخ: 2014/02/17

الأخ / د. يوسف أبو الريش المحترم،،،

مدير عام المستشفيات

السلام عليكم ورحمة الله وبركاته،،،

الموضوع/ تسهيل مهمة باحث

بخصوص الموضوع أعلاه، يرجى تسهيل مهمة الباحث/ رامي عبد الرحيم المصري الملتحق ببرنامج ماجستير العلوم الحياتية تخصص تكنولوجيا حيوية- كلية العلوم- الجامعة الإسلامية بغزة ، في إجراء بحث بعنوان :-

"Genetic Polymorphisms in Lung Cancer Susceptibility between Palestinian Populations (Gaza Strip)"

حيث الباحث بحاجة لتعبئة استبانه وجزء من عينة دم سحبت لأغراض تشخيصية من عدد من مرضى سرطان الرئة وأخرى ممن لا يعانون من هذه المشكلة من المراجعين لأقسام الأورام في مجمع الشفاء الطبي ومستشفى غزة الأوربي.

كما نأمل توجيهاتكم لذوي الاختصاص بعدم السماح للباحث بالتطبيق إلا بعد الحصول على الموافقة المستصدرة من المشاركين في البحث وبإشراف العاملين في أقسام المختبرات ووفق الأسس التي يتم بها التعامل مع هذا النوع من العينات في الوزارة وعلى مسئولية الباحث، بما لا يتعارض مع مصلحة العمل وضمن أخلاقيات البحث العلمي، و دون تحمل الوزارة أي أعباء أو مسئولية.

وتفضلوا بقبول التحية والتقدير،،،

صادر: ...
وارد: ...
تاريخ: ...

د. ناصر رأفت أبو شعبان
مدير عام تنمية القوى البشرية

الإدارة العامة للمستشفيات
وارد
3363
2014/2

دكتور: نصر هاشم الدكتور
Tel / 08-2827298 Fax / 08-2868809 Hmail Ahrd@moh.gov

استشاري امراض باطنة وقلب
ترخيص ٦/٨

The Palestinian National Authority
Ministry of Health
Directorate General of Human Resources Development

السلطة الوطنية الفلسطينية
وزارة الصحة
الإدارة العامة لتنمية القوى البشرية

الرقم: .....

التاريخ: 2014/05/12م

وزارة الصحة
الإدارة العامة للرعاية الأولية
الرقم: 2890
التاريخ: 2014/5/13

الأخ / د. فؤاد العيسوي المحترم،،،

مدير عام الرعاية الأولية

السلام عليكم ورحمة الله وبركاته،،،

الموضوع/ تسهيل مهمة باحث

بخصوص الموضوع أعلاه، يرجى تسهيل مهمة الباحث/ راسي عبد الرحيم المصري الملتحق ببرنامج ماجستير العلوم الحياتية تخصص تكنولوجيا حيوية- كلية العلوم- الجامعة الإسلامية بغزة ، في إجراء بحث بعنوان :-

"Genetic Polymorphisms in Lung Cancer Susceptibility between Palestinian Populations (Gaza Strip)"

حيث الباحث بحاجة لتعبئة استبانه وجزء من عينة دم سحبت لأغراض تشخيصية من عدد من المراجعين لمركز شهداء الرمال الصحي (العينة الضابطة).

كما نأمل توجيهاتكم لذوي الاختصاص بعدم السماح للباحث بالتطبيق إلا بعد الحصول على الموافقة المستبصرة من المشاركين في البحث وبإشراف العاملين في أقسام المختبرات ووفق الأسس التي يتم بها التعامل مع هذا النوع من العينات في الوزارة وعلى مسئولية الباحث، بما لا يتعارض مع مصلحة العمل وضمن أخلاقيات البحث العلمي، و دون تحمل الوزارة أي أعباء أو مسئولية.

وتفضلوا بقبول التحية والتقدير،،،

د. ناصر رافت أبو شعبان
مدير عام تنمية القوى البشرية

وزارة الصحة
تنمية القوى البشرية
التاريخ 2014/5/12

وزارة الصحة
الإدارة العامة للرعاية الأولية
MINISTRY OF HEALTH
Training & Education Dept - PHC
قسم التدريب والتعليم

الإدارة العامة للرعاية الأولية

صاحب العلاقة

Email / hrd@moh.gov.ps

Fax / 08-2868109

FAX NO. :

MAY. 12 2014 03:38PM P1

Questioner for Lung Cancer Susceptibility in Palestinian Population

1. No. of questioner ---------- Place: ----------
2. What is your sex? ☐ Male ☐ Female
3. What is your age? ---------- Tall: ---------- Weight: ----------
4. What is the type of lung cancer in patient? ☐ SCLC ☐ NSCLC
5. Have you ever been diagnosed with recurrent pneumonia? ☐ Yes ☐ No
6. Have any of your first degree relatives (including parents) had a diagnosis of lung cancer?

   ☐ No ☐ Yes If Yes, who's: ----------
7. Have you ever had any type of cancer?

   ☐ Yes ☐ No If Yes, what is the type?: ----------
8. Do you smoke cigarettes?

   ☐ Yes ☐ No ☐ I used to smoke cigarettes, but I quit
9. If you are smoking, at what age did you start smoking? ----------
10. If you are former smoker, what is the period of smoking? ----------
11. How many cigarettes did you used to smoke per day? ----------
12. What is the type of cigarettes? ----------
13. When did you quit smoking cigarettes? ----------
14. Have you been living with a smoker most of your life? ☐ Yes ☐ No
15. Have you smoked Nargilah? ☐ Yes ☐ No
16. Have you smoked water pipe? ☐ Yes ☐ No
17. How long you smoke Nargilah?

    ----
18. How many times you smoke Nargilah in day?

A. The genomic sequences containing the SNPs under investigation (the SNP and the region of the primer annealing are colored and underlined).

1. rs16969968

TCGTTATTGAAGAGATCATACCATCATCTTCAAAAGTCATACCTCTAATTG
GAGAGTATCTGGTATTTACCATGATTTTTGTGACACTGTCAATTATGGTAA
CCGTCTTCGCTATCAACATTCATCATCGTCTTCCTCAACACATAATGCCAT
GGCGCCTTTGGTCCGCAAGATATTTCTTCACACGCTTCCCAAACTGCTTT
GCATGAGAAGTCATGTAGACAGGTACTTCACTCAGAAAGAGGAAACTGA
GAGTGGTATGGACCAAAATCTTCTAGAAACACATTGGAAGCTGCGCTC[A/
G]ATTCTATTCGCTACATTACAAGACACATCATGAAGGAAAATGATGTCCG
TGAGGTCTGTGATGTGTATTTACAAATGCAGATCTCTTCCATTTTAAGTTC
AGAAGTTACTTTCATTAATTTTGGCAGAGTAAACAGCATGACCCTTAAGTA
AGACTAAGCATAGATTGAGGGCCAGAATTGTTGACATATTTTCTATAAAA
GATCTTTACTAAGCTTGTTTCAGTTAAAGCACCTGCAAAATGGGGCATTT
ACACAAATCTCACTTCTCCACTTCCCCCATCAGCATCTTGGATAACT

Sequence Size: 599

2. rs1051730

GTGTGTCACTGTGCCCCCTCTTTGTCTTTGCAGTGCTGTTGGGGATTTCC
AGGTGGACGACAAGACCAAAGCCTTACTCAAGTACACTGGGGAGGTGAC
TTGGATACCTCCGGCCATCTTTAAGAGCTCTGTAAAATCGACGTGACCTA
CTTCCCGTTTGATTACCAAAACTGTACCATGAAGTTCGGTTCCTGGTCCT
ACGATAAGGCGAAAATCGATCTGGTCCTGATCGGCTCTTCCATGAACCT
CAAGGACTATGGGAGAGCGGCGAGTGGGCCATCATCAAAGCCCCAGGC
TA[C/T]AAACACGACATCAAGTACAACTGCTGCGAGGAGATCTACCCCGA
CATCACATACTCGCTGTACATCCGGCGCCTGCCCTTGTTCACACCATCAA
CCTCATCATCCCCTGCCTGCTCATCTCCTTCCTCACTGTGCTCGTCTTCT
ACCTGCCCTCCGACTGCGGTGAGAAGGTGACCCTGTGCATTTCTGTCCT
CCTCTCCCTGACGGTGTTTCCCTGGTGATCACTGAGACCATCCCTTCCA
CCTCGCTGGTCATCCCCCTGATTGGAGAGTACCTCCTGTTCACCATGATT
TTTGTA

Sequence Size: 597

3. **rs402710**

TCTGACTGTGAATAGGTAAATAAATCGCTTAAGGAGAGACATTTGCTTTC
AGTGGCTCATCAATAATTCACATAGAAACGAGCATTTCTAAAGCACAGTG
AGGAGACAGAGCTGGAACAGTTTCTTGCCAGACATATCATGGGCAACTG
GAACCCAAGTTTAGGCAAGACAGGAAAAACCACCACCTGCAAATTATCTT
TTCCCTCAAATGGATAAACAGGCGCAGGGTGCGGTGAAAGCCGTCATTC
CGTTCAGCAGCACCACGCCGCTGAGACGGAGCAACGGCCGAGCATACG
CAGC[C/T]GCACTCACCACCGCTGGTACAGGTAGACCAGAAACACCACG
TCGTCCCGGAAGCAGGCCAGCCGGTGAGACGTGGGCATGGTGAGATGA
AGGCAAAGACGTCATCAATGAAGGTGTTGAAAGCCTGCAGGGCCAGAC
GGGAGGAGGGTGAACCCCAGTTGCTGGGGCTGGAATCCTACTGTTTTTG
GTAACCTAACCAAGCCAACGGCTTTTGCAGATGCTTGGAACTAACTGGA
ACTCCTCACAGCAACAACAAAAAGAGCAGAAAGCCGGCAAGTGGAGATA
CGGAGCTCTGTTCC

Sequence Size: 599

4. **rs2736100**

TTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCTG
CCCGCCTTGGCCTCCCACAGTGCTGGGATTACAGGTGCAAGCCACCGT
GCCCGGCATACCTTGATCTTTTAAAATGAAGTCTGAAACATTGCTACCCT
TGTCCTGAGCAATAAGACCCTTAGTGTATTTTAGCTCTGGCCACCCCCCA
GCCTGTGTGCTGTTTTCCCTGCTGACTTAGTTCTATCTCAGGCATCTTGA
CACCCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTTTC
T[G/T]TAGCTTTGCCCCCGCCCTGCTTTTCCTCCTTTGTTCCCCGTCTGTC
TTCTGTCTCAGGCCCGCCGTCTGGGGTCCCCTTCCTTGTCCTTTGCGTG
GTTCTTCTGTCTTGTTATTGCTGGTAAACCCCAGCTTTACCTGTGCTGGC
CTCCATGGCATCTAGCGACGTCCGGGGACCTCTGCTTATGATGCACAGA
TGAAGATGTGGAGACTCACGAGGAGGGCGGTCATCTTGGCCCGTGAGT
GTCTGGAGCACCACGTGGCCAGCGTTCCTTAGCCAGTGAGTGACAGCAA
CGTCCGCTC

Sequence Size: 601

5. **rs4324798**

ACCCGGGATTGAACCAGGGACCTTTAGATCTTCAGTCTAACGCTCTCCC
AACTGAGCTATCTCGGCCACCGTGATCCTACTGCTTTTGTCATTTCTTCA
AAATACAGAAACTGCCATTTGTAGGGTCAGGTATCTTCCAACGCCTAATT

CTGTTGTCTTCAATATCACCCGTCATTCACTCACCTCCCCTCCACCCAAG
AAATATAAGTTCTGCTGCAATTTATGTGTGAAATAGGATCCAATTTTCCCC
AGCAAAAGTGGGAAAGAAAAGGCGAGGAATAGGTCAAATGAGGAAGATA
CTCCCATGCTTGGTCACCGTATAAAACACTGCTCAGAAAACTAAGGAATT
CAAAATGAAATTATGTAGGCATTTCCTTTTCTCTTTTTCGGATTTTCTTTTT
CTGGCTTGCTCTTCAATGGCATGTCATAAAGGAACAGAAGATTAGTGGAC
ACTTTAACACGGTAGTGGGCTTATAGCTTCCGAAAAAAGACATCCT[A/G]A
GCGAGGTAGTTCTTTTTTCTATTTTCTTCCTTTTACCAGTCTTGTGCTCAC
ACATCCACCTTGGGTGGTACGGAGACCCAGGGAGTGAAAATGGAAAGTA
TAATATGTTTGTTTGTTTGTTTCTTTGTTTCTTTGTTTTGAGTGGAGTCCCG
CTCTGTCTCCCAGGCTGGAGTGCAGTGGCACGATCTGGACTTAGTGCAA
CCTCCGTCTTTCAGGTTCAAGCGATTCTCCTGACTCAGTCTCTTCCAGTA
GGTGGGATTACAGGCGCGCCCACCACGCCCAGCTAATTTTTTTGTATTAT
TAGTAGAGACGAAGTTTCACCATGTTGATCAGTCTGGTCTCGCCTCGGC
CTCCCAAAGTGCTAGGATTACAGGCTTGAGCCACCGTTCCCGGCCTATT
CTTGGAGTTCAGAGAATTGTGGTCTGCACATTGATGCATAAGAATTGTTT
TTTTTTTTCCAGCTGGGTGCAGTGGCTCACGCCTGTAATCCCAGCA

Sequence Size: 994

6. **rs9295740**

AGAAAAACAAGGGAGCGGTTGGAGGGGAAGGGTGGAGAGATGAGGGG
AGGGAGTGCCCTAGTGGAAACCACAGCATTAACACCACCTACTACTTCT
ACTTCACTAAAGGCACTGTCCCGATTTTTTCTTCAGAGATCACTGTTTTGC
CTGCTGAATTCAAACCTCCACCCCAGACACACTGATGTCATTGGAGGCAT
CAGGACTGGGGGCCCAAGTTTTATTATATTAAAACGAGTTCATGCTGGG
GTAAATTTTAAGATCTTTAGTGGACAGAAAGGCAGTTCAAATTCTTTGATT
TTA[A/G]TGACAAAATGCTTTAAACTGACAATGCAACCTATCAACAAAAGG
ACCATATTGAGCTGTGTGTGGGCTGCACAGAAATACGCCGCCCCAGAAC
TCTAAGTGCTCCCGGAAAAGCTCGCAATTGTTACAACAGAGAATCCAATT
CTTGTGGCTAAAGTATCTCCTGGGGGACTTATTACAAATGTGGATTCTAG
AACCCCTCGTGCAATGATCCTGACTAAAGCGATTTACTATGGGTCCCAAG
AATCTTAATCTTACCAAGCACCCTCAAGTAATTCTAACTTCAATAGTCTGC
TAAACG

Sequence Size: 601