European Review for Medical and Pharmacological Sciences 2015; 19: 225-234

Screening of lung cancer related SNPs and CNVs with SNP microarrays

T.-Y. LI, F. ZHANG[1]

Department of Respiratory Medicine, General Hospital of Benxi Iron And Steel Co., LTD, Benxi, China
[1]Department of Respiratory Medicine, the First Affiliated Hospital of China Medical University, Shenyang, China

Abstract. – OBJECTIVE: The objective of this study was to analyze the lung cancer related SNPs (Single Nucleotide Polymorphisms) and CNVs (copy number variations) with SNP microarrays, as well as to identify the CNV related genes and functions.

MATERIALS AND METHODS: The GSE29172 SNP array data were downloaded from Gene Expression Omnibus, including 100%, 30%, 50%, 70% cancer samples, and the mixture is matched blood. The PennCNV software was applied to analyze SNP arrays, and then the related SNPs and CNVs in different degrees of cancer samples were identified. Furthermore, according to the CNVs related chromosome loci, their corresponding genes were selected from University of California Santa genome browser database. Finally, the corresponding genes were performed Gene Ontology and pathway enrichment analysis using DAVID.

RESULTS: The numbers of SNPs in four different degrees of cancer samples were 4299, 1108, 483 and 417, respectively. Meanwhile, most of these SNPs distributed on chromosome 1, 3, 7, 11, 15, 17 and 21. Compared with the known SNPs, all the SNPs identified in our research were novel SNPs associated with lung cancer. The CNVs related genes, ABCG8 and ABCG5 were identified in clusters; meanwhile, the KO05223 pathway and Forkhead gene were screened out in pathway enrichment analysis.

CONCLUSIONS: The lung cancer related SNPs and CNVs were analyzed and their corresponding genes were identified. Furthermore, by analyzing the functions of these genes, the researchers may explore the relationship between genetic mutation and the initiation and development of lung cancer.

Key Words:
Lung cancer, SNPs, CNVs, Function annotation.

Abbreviations

SNPs = Single Nucleotide Polymorphisms; CNVs: copy number variation; GEO = Gene Expression Omnibus; NSCLC = non-small cell lung cancer; TP53 = tumor protein p53; RB1 = retinoblastoma 1; Foxm1 = Forkhead Box m1; Cdk = cyclin-dependent kinases.

Introduction

Lung cancer has the highest cancer mortality rate, with an estimated 160,340 individuals expected to succumb to the disease in 2012 in the US[1]. Tobacco smoking is the most causal factor for lung cancer; however, fewer than 20% of cigarette smokers develop this disease[2], suggesting that inherited genetic factors may also be important risk determinants. Genetic variations at tobacco carcinogen metabolizing enzymes may lead to inter-individual differences in the level of internal carcinogenic dose and to differential risk for individuals with similar exposures[3]. For this reason, genes that encode enzymes give rise to harmful chemicals are suitable candidates for lung cancer susceptibility studies and have been intensively studied[4]. Nevertheless, the published data generally offer inconsistent results[5], due to population heterogeneity, low sample size, poor characterization of the exposure, and a few polymorphisms tested with low power to address the presence of their joint effects.

The power of whole-genome analysis is to analyze a large number of markers and abundant of samples at once, which can reveal weak genetic associations in complex diseases, such as lung cancer[6]. The DNA variations in mutations and single nucleotide polymorphisms (SNPs) can be beneficial, harmful, or have no obvious effect. The previous researches have reported that genetic

Corresponding Author: Fang Zhang, MD; e-mail: fangzhangffzhzh@hotmail.com

mutations occur at 15q24-25, 5p15.33, and 6p21.00 may associate with the risk of lung cancer[7]. Meanwhile, the SNPs of the TERT (telomerase reverse transcriptase) and CLPTM1L (cleft lip and palate transmembrane 1-like) genes, which are located in these regions, are independently and significantly associated with lung cancer risk[8]. In addition, two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese may contribute to the disease[9]. Although, many studies have been conducted and reported to reveal the relationship between genetic mutations and lung cancer, their underlying associations are still not fully explored.

In this study, to screen for common DNA variations, or SNPs, which were associated with lung cancer risk, the lung cancer related SNPs and CNVs (copy number variations) were identified using SNP microarrays. Finally, several genes were screened out after functional analysis. Our findings may replenish the pre-existing knowledge about the association between genetic mutations and lung cancer and have the potential to develop new therapies for this disease.

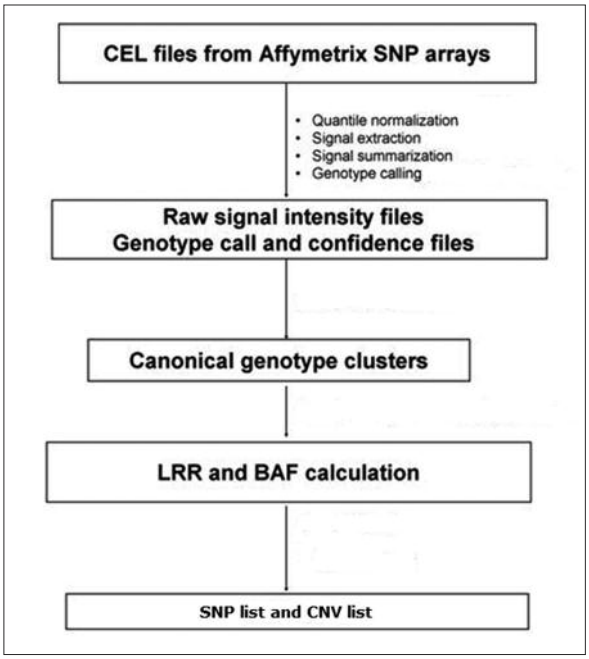

Figure 1. The protocol of PennCNV calculation.

Materials and Methods

SNP chip data of endothelial progenitor cells GSE29172[10] was downloaded from GEO (Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/) database. These experiments packages contains 4 samples, dividedly diluted into different concentration of tumor samples (GSEM721892, GSEM721893, GSEM721894 and GSEM721894 represents 100%, 30%, 50% and 70% tumor samples, respectively. Meanwhile, the mixtures are the blood samples associated with tumor samples).

PennCNV software was utilized to analyze the chip data. This software integrates the information, including the allele B frequency, Log R ratio, probe signal strength, the human feature with multiple kinds of SNP sites and family history to identify CNVs by integrated HMM Algorithms[11]. The protocol of PennCNV calculation was shown in Figure 1. Firstly, the downloaded CEL file was standardized; secondly, the probes were extracted; finally, the SNPs and CNVs results were listed. The result of PennCNV software contains different signal values and probes with various types of SNPs.

Furthermore, the identified SNPs and CNVs were performed statistical analysis and compared with the known lung cancer related SNPs and CNVs. The probes were identified with log Ratio more or less than zero and the SNP frequency more than 0.95. Then, the reported SNPs which were associated with lung cancer were downloaded from NCBI (the National Center for Biotechnology Information) dbSNP database (database for single nucleotide polymorphisms)[12]. Then, the identified SNPs were compared with the published SNPs. Removing the known ones, the remaining SNPs were the novel SNPs associated with lung cancer. In addition, the various types of CNVs in the samples were shown in CNV list.

The corresponding genes in CNV list were searched and the functional analysis was performed for them. According to the chromosome loci in CNV list, the related genes in UCSC (University of California Santa)[13] genome browser database were identified. The UCSC database is an up to date source for genome sequence data integrated with a large number of related annotations[14].

The genes, were conducted Gene ontology (GO) and KEGG (Kyoto encyclopedia of genes and genomes) pathway enrichment analysis by DAVID software with p values less than 0.05.

The GO project convers several domains of molecular and cellular biology and are freely available for community use in the annotation of genes, gene products and sequences[15]. Additionally, as a database resource, the KEGG integrates genomic, chemical and systemic functional information[16].

Results

SNP Analysis

A total of 16876 SNP probes were identified in each samples, including the probes whose log Ratio were zero. The distribution of SNPs in different samples was shown in Figure 2. It was obvious that the signal values in GSEM721892 were stronger than those in the other samples. In Figure 3, the numbers of significative SNPs the number of SNPs in GSEM721892 was larger than those in GSEM721893, GSEM721894 and GSEM721895.

In Table I, the significative SNPs in the four groups were distributed in chromosome 1 to 22. Although, the number of SNPs in each chromosome was different for the same sample, a large number of SNPs were observed mainly located in chromosome 1, 3, 7, 11, 15, 17 and 21 for all the samples (Figure 4). In addition, the number of SNPs in GSEM721892 was larger than those in GSEM721893, GSEM721894 and GSEM721895.

CNV Analysis

As shown in Figure 5, the CNVs were only identified in the GSEM721892 and GSEM721893 samples. It was obvious that duplication was dominant when the samples were 100% lung cancer cells. When the mixtures were 30% lung cancer cells, the deletion and duplication were becoming similar, and the duplication had a little advantage.

Comparing with the Known SNPs

A total of 17 reported SNPs which associated with lung cancer were listed in Table II. Compared with the 17 known SNPs, the significative SNPs were all novel SNPs associated with lung cancer.

Identifying the CNVs Corresponding Genes and Performing GO Functional Analysis

Based on the chromosome loci as well as the hg18 genome annotation information in UCSC, 83 genes were identified. Then DAVID software was applied to perform GO functional analysis. In A total of 16 functional clusters were screened out with p values less than 0.05, and the most significant one was the negative regulation of intestinal cholesterol absorption (p value = 0.005462857)

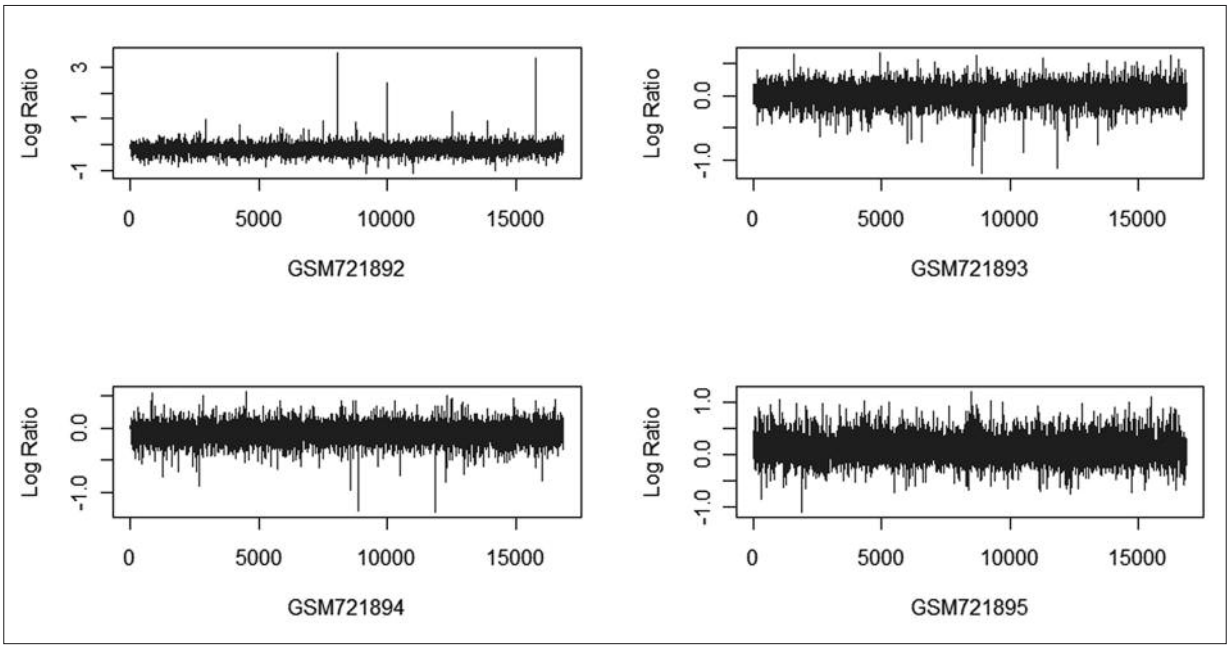

Figure 2. The signal values in four sample groups.

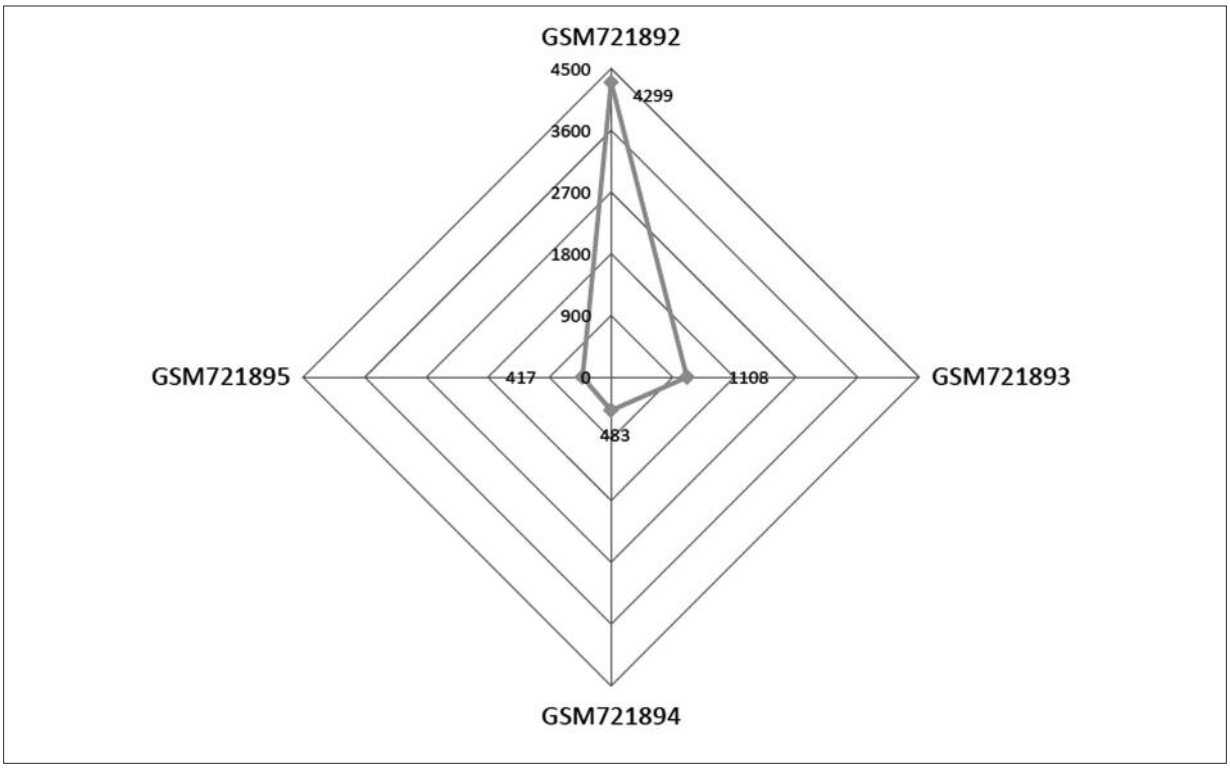

Figure 3. The significative SNPs numbers in four sample groups.

Table I. The distributions of SNPs in the chromosomes for the four sample groups.

RANGE_GB	Chr	GSM721892	GSM721893	GSM721894	GSM721895
NC_000001.10	1	890	247	85	90
NC_000002.11	2	105	26	15	14
NC_000003.11	3	816	192	87	85
NC_000004.11	4	25	10	10	1
NC_000005.9	5	35	6	6	3
NC_000006.11	6	325	93	56	20
NC_000007.13	7	794	218	74	79
NC_000008.10	8	257	45	27	28
NC_000009.11	9	7	14	8	1
NC_000010.10	10	14	11	12	4
NC_000011.9	11	287	58	27	22
NC_000012.11	12	26	4	8	0
NC_000013.10	13	17	4	1	1
NC_000014.8	14	13	5	2	0
NC_000015.9	15	220	62	21	22
NC_000016.9	16	18	3	4	2
NC_000017.10	17	73	28	6	13
NC_000018.9	18	93	36	14	10
NC_000019.9	19	95	14	7	11
NC_000020.10	20	12	3	2	1
NC_000021.8	21	170	28	11	9
NC_000022.10	22	7	1	0	1
In total		4299	1108	483	417

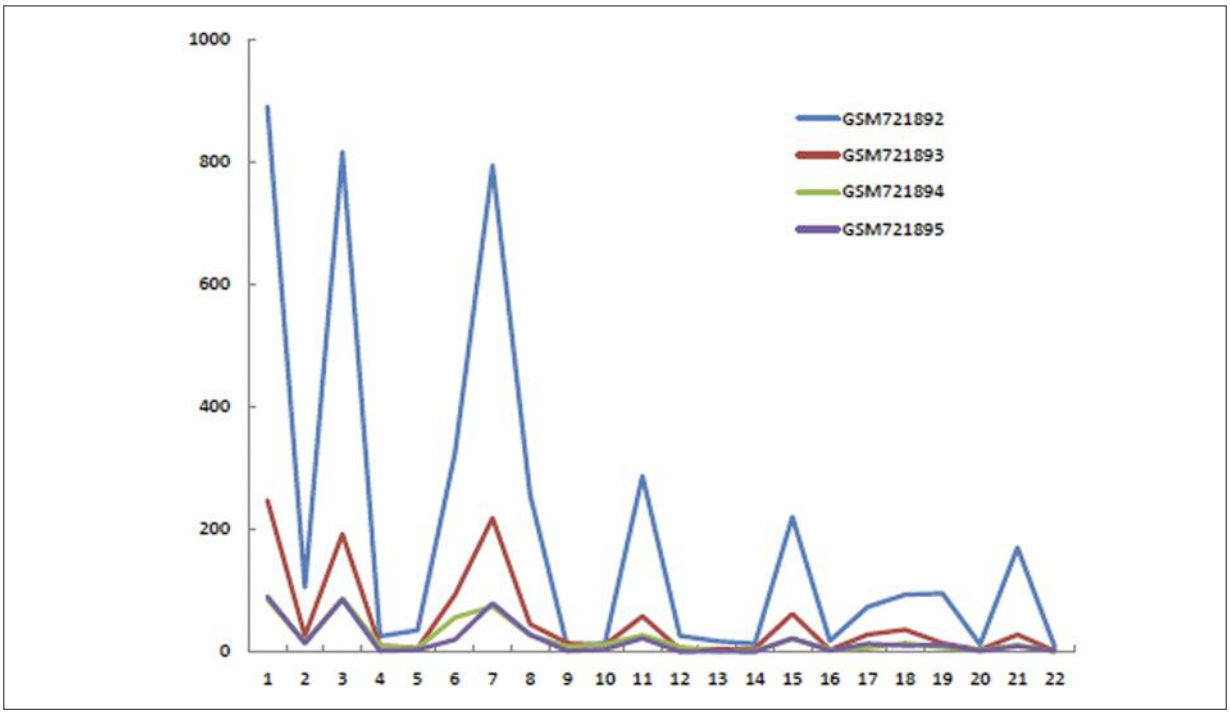

Figure 4. The SNPs numbers in each chromosome for all samples.

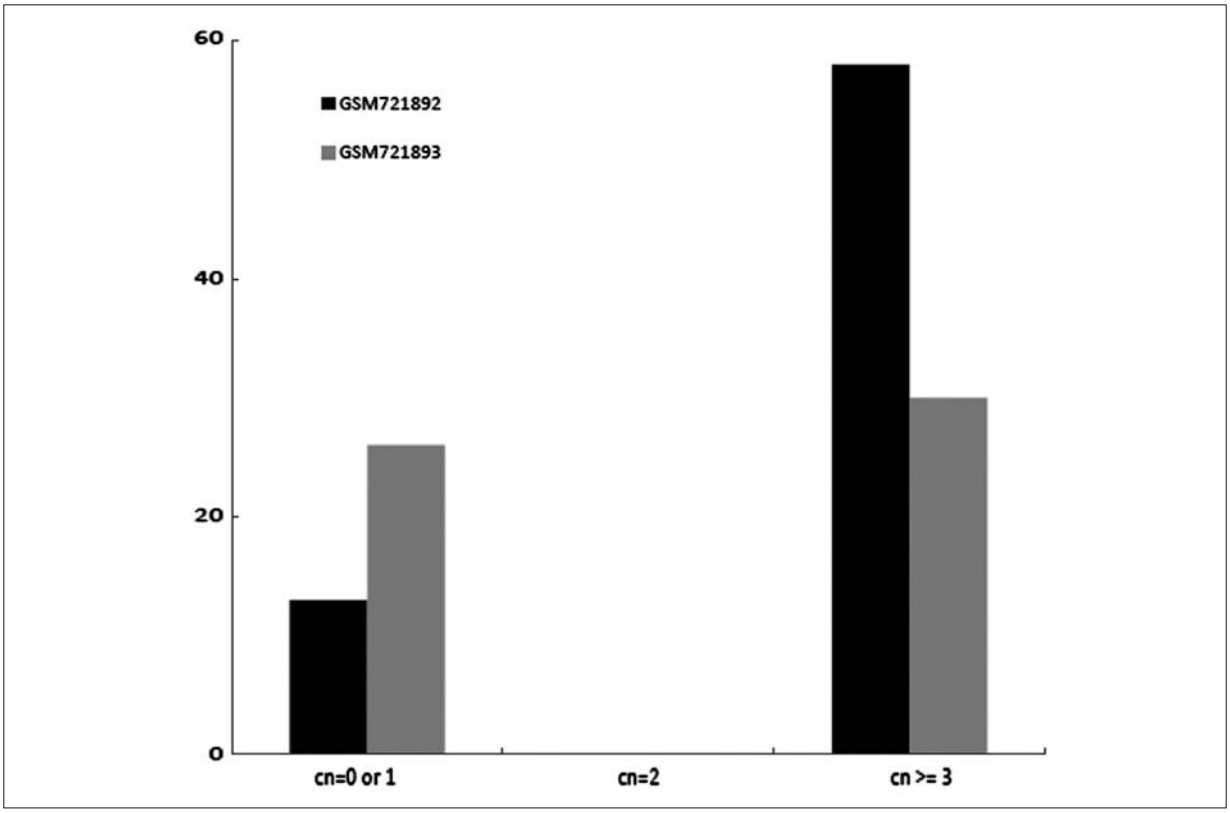

Figure 5. The distribution of each types of CNV in two samples. cn represents the type of copy number, cn = 0 or 1 represents deletion (CNV of deletion), cn = 2 represents normal (normal CN), cn ≥ 3 represents duplication (CNV of replication).

Table II. The known SNPs associated with lung cancer.

SNP_ID	Type
rs17883862	C/T
rs17879961	C/T
rs121913469	CC/TT
rs16969968	A/G
rs1051730	C/T
rs121913297	G/T
rs121913530	A/C/G/T
rs121909071	C/T
rs1805076	A/G
rs121917738	C/T
rs121917737	G/T
rs121434569	C/T
rs121434568	G/T
rs28929495	A/G/T
rs121917702	A/G
rs1801270	A/C
rs3834129	-/CTTACT

(Table III). In addition, the ABCG8 and ABCG5 genes were discovered frequently present in 13 clusters of the 16 function clusters.

KEGG Pathway Enrichment Analysis

The KEGG pathway analysis was performed to the CNVs corresponding genes, and the KO05223 pathway was identified (Figure 6). Additionally, Forkhead gene, whose copy number had varied, was participated in the KO05223 pathway.

Discussion

Research has shown that genetic factors play important roles in the development of lung cancer. In addition, individuals with inherited variations in certain genes have an increased risk of lung cancer. In this study, the SNP data GSE29172 were downloaded, and then were analyzed by PennCNV software. Furthermore, the identified SNPs in our study were compared with the published SNPs associated with lung cancer. In addition, the CNVs corresponding genes were selected and GO and KEGG pathway enrichment analyses were performed for these genes. Finally, ABCG8, ABCG5 and Forkhead genes, and KO05223 pathway, which may have close relationship with the development of lung cancer,were identified in this study.

The numbers of SNPs in the GSEM721892, GSEM721893, GSEM721894 and GSEM721895 were 4299, 1108, 483 and 417, respectively. These SNPs were mainly located on chromosome 1, 3, 7, 11, 15, 17 and 21. Meanwhile, the GSEM721892 sample had more SNPs than the other samples and the dominant CNV was duplication. These findings suggest that 100% lung

Table III. GO enrichment analysis with p values < 0.05.

IDs and names	*p*-value	Gene symbols
GO:0045796~negative regulation of intestinal cholesterol absorption	0.005462857	ABCG8, ABCG5
GO:0060457~negative regulation of digestive system process	0.005462857	ABCG8, ABCG5
GO:0010949~negative regulation of intestinal phytosterol absorption	0.005462857	ABCG8, ABCG5
GO:0010647~positive regulation of cell communication	0.012058414	COL4A4, GPR89C, GPR89B, GRIK2, SOS1, IRS1, GPR89A
GO:0030300~regulation of intestinal cholesterol absorption	0.0163016	ABCG8, ABCG5
GO:0019899~enzyme binding	0.017431144	CUL3, GRIK2, FOXO3, DOCK10, IRS1, DLG1
GO:0032375~negative regulation of cholesterol transport	0.018993276	ABCG8, ABCG5
GO:0032372~negative regulation of sterol transport	0.018993276	ABCG8, ABCG5
GO:0030299~intestinal cholesterol absorption	0.021677785	ABCG8, ABCG5
GO:0044241~lipid digestion	0.027025377	ABCG8, ABCG5
GO:0044058~regulation of digestive system process	0.029688498	ABCG8, ABCG5
GO:0044456~synapse part	0.03153099	GRIK2, SOS1, NRXN1, DLG1
GO:0017127~cholesterol transporter activity	0.034577342	ABCG8, ABCG5
GO:0015248~sterol transporter activity	0.040225901	ABCG8, ABCG5
GO:0032369~negative regulation of lipid transport	0.040270235	ABCG8, ABCG5
GO:0050892~intestinal absorption	0.042898075	ABCG8, ABCG5

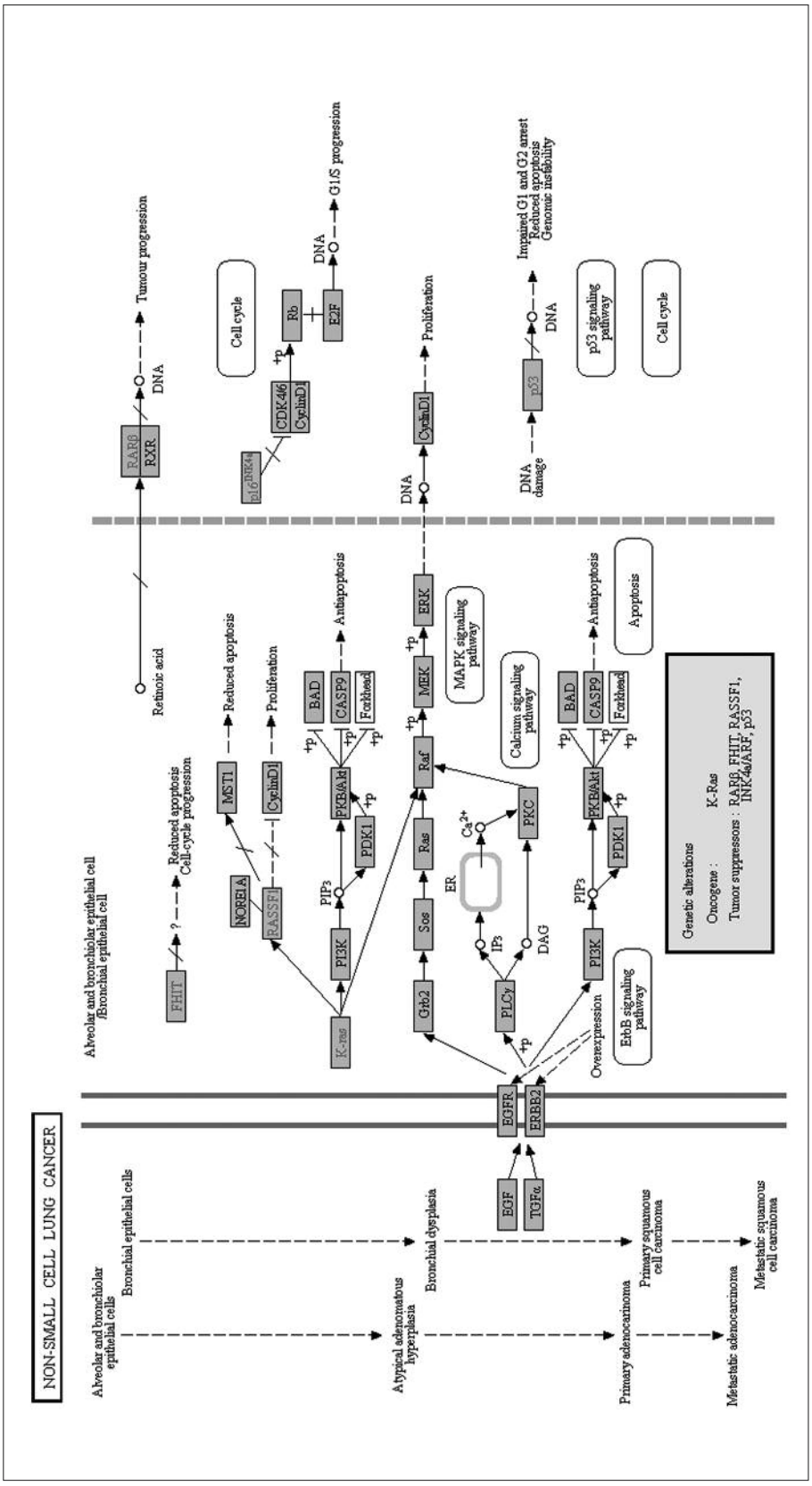

Figure 6. The lung cancer related pathway in which the CNV genes participated. The pink text represents suppressor genes and oncogenes, the genes in the yellow box are the discovered CNVs corresponding genes.

cancer takes more gene mutations, such as duplication, than the 30%, 50% and 70% lung cancers, which is agree with the previous study[6]. Additionally, the significative SNPs with mutation frequency more than 0.95 and log Ratio more or less than zero that had not been reported were identified for the first time in this study. Therefore, these SNPs are novel SNPs associated with lung cancer, which may be the target to explore the mechanism of lung cancer and develop new drugs to combat this disease.

Functional analyses were performed to the CNVs related genes, then ABCG5 and ABCG8 were identified which may play important roles in the initiation and development of lung cancer. ABCG5 is one member of the ATP-binding cassette subfamily G and plays a role in the efflux transport of cholesterol[20,21]. Its expression has been correlated with melanoma progression and it is hypothesized to contribute to the refractoriness of metastatic cancer to chemotherapy[22]. Indeed, specific target of ABCG5 with monoclonal antibodies appears to significantly inhibit cell growth. To date, ABCG5-positivity in tumor buds has been proved to be an indicator of poor prognosis in node-negative colorectal cancer patients[23]. Additionally, SNP in the ABCG8 transporter gene is found to be associated with gallbladder cancer susceptibility[24]. Genetic variation at the ABCG5/8 locus has been investigated and discovered to be associated with markers of cholesterol homeostasis[25]. In our study, both ABCG5 and ABCG8 genes were proved to be associated with CNVs in lung cancer samples and mutations in both the genes have the potential to lead to lung cancer.

Furthermore, the KEGG pathway enrichment analysis was conducted and the KO05223 pathway was identified. In this pathway, the CNVs corresponding gene, Forkhead (FOX gene family), was screened out. The FOX gene family provides instructions for producing proteins that play a critical role in the formation of many organs and tissues before birth[26]. Mutations in some FOX genes may lead to the initiation and development of tumors. FoxF1-null mutants show early embryonic lethality, and haploinsufficiency causes multiple defects in trachea, esophagus, lungs, and gallbladder[27,28]. The importance of FoxF1 in human lung development is also highlighted by the recent findings that inactivating mutations of FoxF1 cause congenital malformations of the lungs[29].

In addition, FoxM1 is known to activate the transcription of genes essential for progression of DNA replication and mitosis[30]. Increased expression of FoxM1 has been found in human basal cell carcinomas[31], intrahepatic cholangio carcinomas[32], infiltrating ductal breast carcinomas[33], and in many other solid tumors[34]. These findings suggest that FoxM1 is required for cellular proliferation in various human cancers. Although the FoxM1 protein is essential for hepatocyte proliferation during progression of hepatocellular carcinoma[35], the role of FoxM1 in lung cancer remains to be determined. In this study, the finding that Forkhead gene was a critical component in KO05223 pathway may fulfill the knowledge on functions of forkhead family and assist the researchers to understand the mechanism of lung cancer profoundly.

Conclusions

The genetic basis of inherited susceptibility to lung cancer out of the context of these disorders is at present undefined, but a model in which high-risk alleles account for all of the excess familial risk seems unlikely. This hypothesis implies that test for allelic association could be a powerful strategy for identifying alleles that predispose to lung cancer.

Conflict of Interest

The Authors declare that there are no conflicts of interest.

References

1) Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer J Clin 2012; 62: 10-29.

2) Wright GS, Gruidl ME. Early detection and prevention of lung cancer. Curr Opin Oncol 2000; 12: 143-148.

3) Schwartz AG, Prysak GM, Bock CH, Cote ML. The molecular epidemiology of lung cancer. Carcinogenesis 2007; 28: 507-518.

4) Gresner P, Gromadzinska J, Wasowicz W. Polymorphism of selected enzymes involved in detoxification and biotransformation in relation to lung cancer. Lung Cancer 2007; 57: 1-25.

5) Sugimura H, Tao H, Suzuki M, Mori H, Tsuboi M, Matsuura S, Goto M, Shinmura K, Ozawa T, Tanioka F, Sato N, Matsushima Y, Kageyama S, Funai K, Chou PH, Matsuda T. Genetic susceptibility to lung cancer. Front Biosci (Schol Ed) 2011; 3: 1463-1477.

6) Lee W, Jiang Z, Liu J, Haverty PM, Guan Y, Stinson J, Yue P, Zhang Y, Pant KP, Bhatt D, Ha C, Johnson S, Kennemer MI, Mohan S, Nazarenko I, Watanabe C, Sparks AB, Shames DS, Gentleman R, de Sauvage FJ, Stern H, Pandita A, Ballinger DG, Drmanac R, Modrusan Z, Seshagiri S, Zhang Z. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature 2010; 465: 473-477.

7) Rafnar T, Sulem P, Besenbacher S, Gudbjartsson DF, Zanon C, Gudmundsson J, Stacey SN, Kostic JP, Thorgeirsson TE, Thorleifsson G, Bjarnason H, Skuladottir H, Gudbjartsson T, Isaksson HJ, Isla D, Murillo L, García-Prats MD, Panadero A, Aben KK, Vermeulen SH, van der Heijden HF, Feser WJ, Miller YE, Bunn PA, Kong A, Wolf HJ, Franklin WA, Mayordomo JI, Kiemeney LA, Jonsson S, Thorsteinsdottir U, Stefansson K. Genome-wide significant association between a sequence variant at 15q15. 2 and lung cancer risk. Cancer Res 2011; 71: 1356-1361.

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

9) Hu Z, Wu C, Shi Y, Guo H, Zhao X, Yin Z, Yang L, Dai J, Hu L, Tan W, Li Z, Deng Q, Wang J, Wu W, Jin G, Jiang Y, Yu D, Zhou G, Chen H, Guan P, Chen Y, Shu Y, Xu L, Liu X, Liu L, Xu P, Han B, Bai C, Zhao Y, Zhang H, Yan Y, Ma H, Chen J, Chu M, Lu F, Zhang Z, Chen F, Wang X, Jin L, Lu J, Zhou B, Lu D, Wu T, Lin D, Shen H. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12. 12 and 22q12. 2 in Han Chinese. Nat Gen 2011; 43: 792-796.

10) Rasmussen M, Sundstrom M, Goransson Kultima H, Botling J, Micke P, Birgisson H, Glimelius B, Isaksson A. Allele-specific copy number analysis of tumor samples with aneuploidy and tumor heterogeneity. Genome Biol 2011; 12: R108.

11) Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res 2007; 17: 1665-1674.

12) Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: The NCBI database of genetic variation. Nucleic Acids Res 2001; 29: 308-311.

13) Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using david bioinformatics resources. Nat Protoc 2009; 4: 44-57.

14) Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A, Diekhans M, Dreszer TR, Giardine BM, Harte RA, Hillman-Jackson J, Hsu F, Kirkup V, Kuhn RM, Learned K, Li CH, Meyer LR, Pohl A, Raney BJ, Rosenbloom KR, Smith KE, Haussler D, Kent WJ. The UCSC Genome Browser database: update 2011. Nucleic Acids Res 2011; 39: D876-D882.

15) Young MD, Wakefield MJ, Smyth GK, Oshlack A. Method gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol 2010; 11: R14.

16) Zhang JD, Wiemann S. Kegggraph: a graph approach to kegg pathway in r and bioconductor. Bioinformatics 2009; 25: 1470-1471.

17) Wu Q, Min J, Jiang L, Li M, Yao K. Screening of biomarkers for lung cancer with gene expression profiling data. Eur Rev Med Pharmacol Sci 2013; 17(23): 3221-3228.

18) Wu X, Zang W, Cui S, Wang M. Bioinformatics analysis of two microarray gene-expression data sets to select lung adenocarcinoma marker genes. Eur Rev Med Pharmacol Sci 2012; 16(11): 1582-1587.

19) Li J, Shi R, Yu P. Effects of chemotherapy on global gene expression in non-small cell lung cancer. Eur Rev Med Pharmacol Sci 2014; 18(1): 126-131.

20) Hirata T, Okabe M, Kobayashi A, Ueda K, Matsuo M. Molecular mechanisms of subcellular localization of ABCG5 and ABCG8.Biosci Biotechnol Biochem 2009; 73: 619-626.

21) Kusuhara H, Sugiyama Y. ATP-binding cassette, subfamilyG (ABCG family). Pflugers Arch 2007; 453: 735-744.

22) Schatton T, Murphy GG, Frank NY, Yamaura K, Waaga-Gasser AM, Gasser M, Zhan Q, Jordan S, Duncan Lm, Weishaupt C, Fuhlbrigge RC, Kupper TS, Sayegh MH, Frank MH. Identification of cells initiating human melanomas. Nature 2008; 451: 345-349.

23) Hostettler L, Zlobec I, Terracciano L, Lugli A. ABCG5-positivity in tumor buds is an indicator of poor prognosis in node-negative colorectal cancer patients. World J Gastroenterol 2010; 16: 732-739.

24) Srivastava A, Tulsyan S, Pandey SN, Choudhuri G, MittaL B. Single nucleotide polymorphism in the ABCG8 transporter gene is associated with gallbladder cancer susceptibility. Liver Int 2009; 29: 831-837.

25) Jakulj L, Vissers MN, Tanck Mw, Hutten BA, Stellaard F, Kastelein JJ, Dallinga-Thie GM. ABCG5/G8 polymorphisms and markers of cholesterol metabolism: systematic review and meta-analysis. J Lipid Res 2010; 51: 3016-3023.

26) Jackson BC, Carpenter C, Nebert DW, Vasiliou V. Update of human and mouse forkhead box (fox) gene families. Hum Genomics 2010; 4: 345-352.

27) Kalinichenko VV, Lim L, Stolz DB, Shin B, Rausa Fm, Clark J, Whitsett JA, Watkins SC, Costa RH. Defects in pulmonary vasculature and perinatal lung hemorrhage in mice heterozygous null for the Forkhead Box f1 transcription factor. Dev Biol 2001; 235: 489-506.

28) Kalinichenko VV, Zhou Y, Bhattacharyya D, Kim W, Shin B, Bambal K, Costa RH. Haploinsufficiency of the mouse Forkhead Box f1 gene causes defects in gall bladder development. J Biol Chem 2002; 277: 12369-12374.

29) McLin VA. Summary of "genomic and genetic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations". J Pediatr Gastroenterol Nutr 2010; 50: 350-351.

30) Costa RH, Kalinichenko VV, Major Ml, Raychaudhuri P. New and unexpected: forkhead meets ARF. Curr Opin Genet Dev 2005; 15: 42-48.

31) Teh MT, Wong ST, Neill Gw, Ghali LT, Philpott MP, Quinn AG. FOXM1 is a downstream target of Gli1 in basal cell carcinomas. Cancer Res 2002; 62: 4773-4780.

32) Obama K, Ura K, Li M, Katagiri T, Tsunoda T, Nomura A, Satoh S, Nakamura Y, Furukawa Y. Genome-wide analysis of gene expression in human intrahepatic cholangiocarcinoma. Hepatology 2005; 41: 1339-1348.

33) Wonsey DR, Follettie MT. Loss of the forkhead transcription factor FoxM1 causes centrosome amplification and mitotic catastrophe. Cancer Res 2005; 65: 5181-5189.

34) Pilarsky C, Wenzig M, Specht T, Saeger HD, Grutzmann R. Identification and validation of commonly overexpressed genes in solid tumors by comparison of microarray data. Neoplasia 2004; 6: 744-750.

35) Kalinichenko VV, Major ML, Wang X, Petrovic V, Kuechle J, Yoder HM, Dennewitz MB, Shin B, Datta A, Raychaudhuri P, Costa RH. Foxm1b transcription factor is essential for development of hepatocellular carcinomas and is negatively regulated by the p19ARF tumor suppressor. Genes Dev 2004; 18: 830-850.