A Decade of GWAS Results in Lung Cancer

Yohan Bossé[1,2] and Christopher I. Amos[3]

Abstract

Genome-wide association studies (GWAS) were successful to identify genetic factors robustly associated with lung cancer. This review aims to synthesize the literature in this field and accelerate the translation of GWAS discoveries into results that are closer to clinical applications. A chronologic presentation of published GWAS on lung cancer susceptibility, survival, and response to treatment is presented. The most important results are tabulated to provide a concise overview in one read. GWAS have reported 45 lung cancer susceptibility loci with varying strength of evidence and highlighted suspected causal genes at each locus. Some genetic risk loci have been refined to more homogeneous subgroups of lung cancer patients in terms of histologic subtypes, smoking status, gender, and ethnicity. Overall, these discoveries are an important step for future development of new therapeutic targets and biomarkers to personalize and improve the quality of care for patients. GWAS results are on the edge of offering new tools for targeted screening in high-risk individuals, but more research is needed if GWAS are to pay off the investment. Complementary genomic datasets and functional studies are needed to refine the underlying molecular mechanisms of lung cancer preliminarily revealed by GWAS and reach results that are medically actionable. *Cancer Epidemiol Biomarkers Prev; 27(4); 363–79. ©2017 AACR.*

See all articles in this *CEBP Focus* section, "Genome-Wide Association Studies in Cancer."

Introduction

Lung cancer is the leading cause of cancer-related deaths worldwide in both men and women (1, 2). Our molecular understanding of this disease is in progress. Although it has long been recognized that lung cancer runs strongly in families (3–5), the specific genes that are responsible for enhanced risk are just starting to be revealed. Identifying genes responsible for lung cancer before the era of genome-wide association studies (GWAS) has been limited. Candidate susceptibility genes coding for enzymes involved in the activation, detoxification, and repair of damages caused by tobacco smoke as well as genes in inflammatory and cell-cycle pathways have been extensively studied (6, 7). Many of these candidate gene studies are either preliminary or controversial (8). Rare germline mutations in *TP53*, *RB1*, and *EGFR* have been shown to confer inherited predisposition to lung cancer (9–11). Fine mapping of genome-wide linkage peak on 6q23-25 also identified *RGS17* as a predisposing gene (12). With the arrival of GWAS approximately 10 years ago, it became possible to interrogate the human genome more comprehensively for lung cancer susceptibility genes.

[1]Institut Universitaire de Cardiologie et de Pneumologie de Québec, Quebec, Canada. [2]Department of Molecular Medicine, Laval University, Quebec, Canada. [3]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Hanover, New Hampshire.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Yohan Bossé, Institut Universitaire de Cardiologie et de Pneumologie de Québec, 2725, chemin Ste-Foy, Québec G1V 4G5, Canada. Phone: 418-656-8711, ext. 3725; Fax: 418-656-4602; E-mail: yohan.bosse@criucpq.ulaval.ca

doi: 10.1158/1055-9965.EPI-16-0794

Chronologic Presentation of Published GWAS on Lung Cancer

GWAS have identified genetic factors robustly associated with lung cancer. Tables 1 and 2 provide a chronologic presentation of published GWAS on lung cancer in European and Asian populations, respectively, and summarize susceptibility loci identified. During the last decade, GWAS have evolved from finding lung cancer loci *per se*, to a more refined search strategy focused on specific subgroups of lung cancer patients. Advances in analysis strategies were also achieved moving from single marker analyses to pathway-based and variant prioritization approaches. GWAS have also been performed to find genetic loci associated with lung cancer survival, response to conventional therapies, and multiple sites of cancer. GWAS by environmental exposures and genome-wide epistasis analyses are also emerging. This review aims to provide a concise overview of this literature. All GWAS on lung cancer susceptibility, survival, and response to treatment reported in the literature were cumulated and manually curated by the authors. At the time of writing this review, a search on PubMed with the keywords "GWAS" and "lung cancer" was performed to identify any missing GWAS in the field. Finally, GWAS and susceptibility loci on lung cancer were further refined with the GWAS catalog (13). Please note that we have attempted to include all loci reported in the literature without quality assessment or exclusion criteria based on the magnitude of effects, sample size, or other criteria. Lung cancer susceptibility loci were reported on the basis of the interpretation of the authors in the original articles.

GWAS on Lung Cancer Susceptibility

GWAS in European populations

The first GWAS on lung cancer were reported in 2008. Three independent studies identified a susceptibility locus on

Table 1. Susceptibility loci for lung cancer identified by GWAS in European ancestry populations

Reference	Study[a]	Sample size (cases/controls)	Disease/trait	Platform (# SNPs)	Region (size)	Gene	Key SNPs
Hung et al. (14)	IARC	1,926/2,522	Lung cancer	Illumina HumanHap300 (310,023)	15q25 (182 kb)	*CHRNA5*	rs1051730
	Toronto	330/453				*CHRNA3*	rs8034191
	EPIC	781/1,578				*CHRNB4*	rs16969968
	CARET	764/1,515				*IREB2*	(D398N)
	Liverpool	403/814				*PSMA4*	
	HUNT/Tromsø	235/392				*HYKK (AGPHD1)*	
Thorgeirsson et al. (15)	Icelandic smokers	10,995	Smoking quantity	Illumina HumanHap300 (306,207)	15q24	*CHRNA5*	rs1051730
	Icelandic smokers	2,950	Nicotine dependence			*CHRNA3*	
	Iceland/Spain/The Netherlands	1,024/32,244	Lung cancer			*CHRNB4*	
Amos et al. (16)	Texas	1,154/1,137	Non-small cell lung cancer	Illumina HumanHap300 (315,450)	15q25.1 (88 kb)	*CHRNA5*	rs1051730
	Texas replication	711/632				*CHRNA3*	rs8034191
	UK	2,013/3,062				*PSMA4*	rs931794
						HYKK	
Liu et al. (17)	GELCC	194/219	Familial lung cancer	Affymetrix 500K (399,377) or 6.0 (722,376)	15q24-25.1 (160 kb)	*CHRNA5*	rs8034191
						CHRNA3	rs1051730
						CHRNB4	rs16969968
						IREB2	(D398N)
						PSMA4	rs578776
						HYKK	
McKay et al. (19)	Central Europe/Toronto/ HUNT2-Tromso/CARET	2,971/3,746	Lung cancer	Illumina HumanHap300 (315,194)	15q25.1		rs1051730
	EPIC/Szczecin/CARET2/ Liverpool	2,899/5,573			5p15.33	*TERT*	rs402710
						CLPTM1L	rs2736100
Wang et al. (18)	British cohort	1,952/1,438	Lung cancer	Illumina HumanHap550 (511,919)	15q25.1	*BAG6*	rs8042374
	IARC	1,989/2,625			6p21.33 (627 kb)	*MSH5*	rs3117582
	Texas	1,154/1,137			5p15.33 (60 kb)	*CLPTM1L*	rs3131379
	UK replication	2,448/2,983					rs401681
Broderick et al. (20)	GELCAPS phase 1	1,952/1,438	Lung cancer	Illumina HumanHap550 (511,919)	15q25.1 (248 kb)	*CHRNA3*	rs12914385
	GELCAPS phase 2	2,465/3,005					rs938682
							rs8042374
							rs8034191
	Meta-analysis (GELCAPS, IARC, Texas)	7,560/8,205			5p15.33 (60 kb)	*CLPTM1L*	rs4975616
						TERT	
					6p21.33	*BAG6*	rs3117582
						TNXB	rs1150752
Landi et al. (21)	NCI (EAGLE, ATBC, PLCO, CPS-II)	5,739/5,848	Lung cancer Adenocarcinoma Squamous cell	Illumina (515,922)	15q25	*CHRNA3*	rs12914385
						CHRNA5	rs1051730
						HYKK	rs8034191
	Meta-analysis (UK, Central Europe, Texas, DeCODE Genetics, HGF Germany, CARET, HUNT2/Tromso, Canada, France, Estonia)	13,300/19,666	Small cell		5p15	*TERT*	rs2736100
						CLPTM1L	rs4635969
							rs31489
					6p21	*BAG6*	rs3117582
						APOM	

(Continued on the following page)

Table 1. Susceptibility loci for lung cancer identified by GWAS in European ancestry populations (Cont'd)

Reference	Study[a]	Sample size (cases/controls)	Disease/trait	Platform (# SNPs)	Region (size)	Gene	Key SNPs
Timofeeva et al. (22)	TRICL[b]	14,900/29,485	Lung cancer	Illumina HumanHap300 (318,094) + HumanHap550 or 610Quad (217,914)	15q25	*CHRNA5*	rs1051730
	Han Chinese	2,338/3,077				*CHRNA3*	rs8034191
						CHRNB4	rs6495309
						HYKK	rs680244
							rs6495306
							rs951266
					5p15.33	*TERT*	rs2736100
						CLPTM1L	rs401681
							rs2853677
							rs465498
					6p21-22	*BAG6*	rs3117582
						MSH5	rs2523546
							rs2523571
			Squamous cell carcinoma		12p13.33	*RAD52*	rs10849605
							rs3748522
					9p21.3	*CDKN2A*	rs1333040
						CDKN2B	rs1537372
						ANRIL (*CDKN2B-AS1*)	
					2q32.1	*NUP35*	rs11683501
Wang et al. (23)	MDACC	11,348/15,861	Lung cancer	Illumina 317, 317+240S, 370Duo, 550, 610 or 1M	13q13.1	*BRCA2*	rs11571833 (K3326X)
	ICR		Adenocarcinoma				
	NCI		Squamous cell carcinoma				
	IARC						
	EPIC	*10,246/3,8295*				*FRY*	rs56084662
	ICR						
	IARC						
	Toronto						
					22q12.1	*CHEK2*	rs17879961 (I157T)
					3q28	*TP63*	rs13314271
							rs4488809
McKay, Hung et al. (25)	OncoArray[c]	29,863/55,586	Lung cancer	Oncoarray (10,439,017)	1p31.1	*FUBP1*	rs71658797
					6q27	*RNASET2*	rs6920364
					8p21.1	*EPHX2 CHRNA2*	rs11780471
					13q13.1	*BRCA2*	rs11571833
					15q21.1	*SEMA6D*	rs66759488
					15q25.1	*CHRNA5*	rs55781567
					19q13.2	*CYP2A6*	rs56113850
		11,245/54,619	Adenocarcinoma		3q28	*TP63*	rs13080835
					5p15.33	*TERT*	rs7705526
					8p12	*NRG1*	rs4236709
					9p21.3	*MTAP CDKN2A*	rs885518
					10q24.3	*OBFC1*	rs11591710
					11q23.3	*MPZL3 AMICA1*	rs1056562
					15q21.1	*SECISBP2L*	rs77468143
					20q13.33	*RTEL1*	rs41309931
		7,704/54,763	Squamous cell carcinoma		6p21.33	*MHC*	rs116822326
					12p13.33	*RAD52*	rs7953330
					22q12.1	*CHEK2*	rs17879961

(Continued on the following page)

Table 1. Susceptibility loci for lung cancer identified by GWAS in European ancestry populations (Cont'd)

Reference	Study[a]	Sample size (cases/controls)	Disease/trait	Platform (# SNPs)	Region (size)	Gene	Key SNPs
Never smokers							
Li et al. (34)	Mayo *MDACC* *Harvard* *UCLA*	377/377 328/407 92/161 91/439	Lung cancer in never smokers	Illumina HumanHap370 and HumanHap610 (331,918)	13q31.3	*GPC5*	rs2352028 rs2352029
Pathway-based GWAS							
Shi et al. (40)	NCI *UK1* *Texas* *UK2*	5,355/4,344 *592/2,699* *306/1,137* *1,038/933*	Lung cancer Squamous cell carcinoma Adenocarcinoma, small cell	Illumina (19,082) (pathway-based analysis)	12p13.33	*RAD52*	rs6489769
Spitz et al. (41)	Texas *Mayo*	451/508 *303/311*	NSCLC in never smokers	Illumina (11,737) (pathway-based analysis)	12q13	*ACVR1B* *NR4A1*	rs12809597 rs2701129 rs1882119
Wang et al. (42)	ICR MDACC IARC NCI Toronto HGF Germany	12,160/16,838	Lung cancer	Illumina (826 functional SNPs) (pathway-based analysis)	6p21.33 5q14.2	*MSH5* *GTF2H4* *XRCC4*	rs3115672 rs114596632 rs1056503 rs2035990
Variant prioritization approaches							
Li et al. (44)	Texas	1,154/1,137	Lung cancer + intermediate phenotype (cigarettes per day)	Illumina HumanHap300	15q24-25.1 19q13 3p26	*HYKK* *CHRNA3* *CHRNA5* *CHRNB4* *TGFB1* *B9D2*	rs12914385 rs1800469 rs1982072 rs2241714 rs1444056 rs1403124
Poirier et al. (45)	Toronto IARC MDACC HMGU *NCI* *MSH-PMH* *MEC* *Harvard*	331/499 1,964/2,610 1,154/1,137 504/484 *5,699/5,815* *1,073/939* *215/225* *523/497*	Lung cancer + family history	Illumina HumanHap300, 550	10q23.33 4p15.2	*FFAR4* *KCNIP4*	rs12415204 rs1158970
Brenner et al. (46)	TRICL[b] ILCCO	5,061 (SQ)/6,756 (AD)/2,216 (SCLC)/33,456 *625 (SQ)/1,417 (AD)/ 369 (SCLC)/2,966*	Squamous cell carcinoma Adenocarcinoma	Illumina HumanHap300, 550, 610	4p15.2 18q12.1	*KCNIP4* *GAREM*	rs6448050 rs9799795 rs11662168 rs3786309

(Continued on the following page)

Table 1. Susceptibility loci for lung cancer identified by GWAS in European ancestry populations (Cont'd)

Reference	Study[a]	Sample size (cases/controls)	Disease/trait	Platform (# SNPs)	Region (size)	Gene	Key SNPs
Cross-cancer loci							
Hung et al. (55)	GAME-ON/GECCO	64,591/74,467	Cross-cancer	Illumina, Affymetrix (12,370 (pathway-based analysis)	12q24	*SH2B3*	rs3184504 (W262R)
					5p15	*TERT*	rs2736100
Fehringer et al. (56)	GAME-ON/GECCO	61,851/61,820	Cross-cancer	Illumina, Affymetrix (9,916,564)	1q22	*MUC1*	rs1057941
	Europeans	*55,789/330,490*				*ADAM15*	rs4072037
	Others	*18,152/21,410*				*THBS3*	
					9p21.3	*CDKN2B-AS1*	rs62560775
					13q13.1	*BRCA2*	rs11571833

Abbreviations: ATBC, Alpha-Tocopherol, Beta-Carotene Cancer Prevention study; CARET, Beta-Carotene and Retinol Efficacy Trial; CPS-II, Cancer Prevention Study II Nutrition Cohort; EAGLE, Environment and Genetics in Lung Cancer Etiology; EPIC, European Prospective Investigation in Cancer and Nutrition; GAME-ON, The Genetic Associations and Mechanisms in Oncology; GECCO, The Genetic and Epidemiology of Colorectal Cancer Consortium; GELCAPS, Genetic Lung Cancer Predisposition Study; GELCC, Genetic Epidemiology of Lung Cancer Consortium; HGF Germany, Helmholtz-Gemeinschaft Deutscher Forschungszentren Lung Cancer GWAS; HMGU, Germany Study; HUNT, Health Study of North-Trondelag; IARC, International Agency for Research on Cancer; ICR, Institute of Cancer Research; ILCCO, International Lung Cancer Consortium; MDACC, M.D. Anderson Cancer Center; MEC, Multi-Ethnic Cohort; MSH-PMH, Mount Sinai Hospital and Princess Margaret Hospital in Toronto; NCI, National Cancer Institute; PLCO, Prostate, Lung, Colon, Ovary Screening Trial; TRICL, Transdisciplinary Research In Cancer of the Lung; UCLA, University of California in Los Angeles.

[a]Italic text indicates replication cohorts.

[b]TRICL includes the following cohorts: MDACC, Liverpool, ICR, Toronto, IARC (Central Europe, CARET, Estonia, France, HUNT2/Tromso), DeCODE Genetics, HGF Germany, Harvard, NCI (EAGLE, ATBC, PLCO, CPS-II).

[c]OncoArray includes the following cohorts: CARET, PLCO, MEC, NELCS (New England Lung Cancer Study), Harvard, MDACC, Tampa (Tampa Lung Cancer Study), BioVU (Vanderbilt 2), LCRI-DOD, TLC (Total Lung Cancer: Molecular Epidemiology of Lung Cancer Survival), Canada (Canadian Screening Study), MSH-PMH, CAPUA (Cancer de Pulmon en Asturias), ATBC, NSHDC (Northern Sweden Health and Disease Cohort), MDCS (The Malmö Diet and Cancer Study), EPIC, Liverpool, Norway (Norway Lung Cancer Study), EAGLE, Nijmegen (The Nijmegen Lung Cancer Study), NICCC-LCA (Clalit National Israeli Cancer Control Center Lung Cancer Study), L2 (The IARC L2 Study), Copenhagen (Copenhagen Lung Cancer Study), Germany, ReSoLuCENT.

chromosome 15q. Hung and colleagues (14) found two SNPs strongly associated with lung cancer on chromosome 15q25. Further genotyping in this region revealed many SNPs in tight linkage disequilibrium (LD) showing evidence of association. Six genes are located in this region including three nicotinic acetylcholine receptor subunits (*CHRNA5*, *CHRNA3*, and *CHRNB4*). Interestingly, no appreciable variation in the risk was found across smoking categories or histologic subtypes of lung cancer. In a second GWAS, a SNP within the *CHRNA3* gene was strongly associated with smoking quantity and nicotine dependence (15). The same SNP was also strongly associated with lung cancer. The results suggest that the variant on chromosome 15q25 confers risk of lung cancer through its effect on tobacco addiction. In contrast, a third study showed weak evidence that the 15q25 locus influences smoking behavior and is mostly directly associated with lung cancer (16). However, it should be emphasized that the later GWAS was conducted in cases and controls matched on smoking status, thus limiting variation between the two groups and the power to detect any smoking association. Further analyses from the same study suggest that SNPs and smoking have independent effects on risk. Together, these three studies unequivocally support the 15q25 locus as harboring susceptibility variants for lung cancer or smoking behavior.

A GWAS performed in familial lung cancer confirmed the susceptibility locus on 15q24-25.1 (17). A subsequent GWAS identified two newly associated risk loci for lung cancer (18). In this study, 15q25 was again the most strongly associated locus. However, by pooling the results with other studies (14, 16), new cancer risk loci were found. Two intronic SNPs located in different genes (*BAG6*, previously known as *BAT3*, and *MSH5*) and separated by more than 600 kilobases on chromosome 6p21 were significantly associated with lung cancer. The strongest association, aside 15q25 and 6p21, was found on chromosome 5p15 within the *CLPTM1L* gene. The 5p15 locus was further supported by an expanded GWAS from previous populations (19). Two uncorrelated SNPs in that region were strongly associated with lung cancer. These SNPs are located within or in proximity to two biologically relevant genes namely *CLPTM1L* and *TERT*. Together, by the end of 2008, three susceptibility loci for lung cancer were identified, that is, 15q25, 6p21, and 5p15.

A more extensive follow-up on a previous GWAS (18) further supports the contribution of the three loci (20). The latter study supports the possibility that two independent loci are acting on 15q25. The latter locus was also associated with smoking behavior, with risk alleles correlated with higher tobacco consumption. In contrast, the 5p15 and 6p21 loci were not associated with smoking behavior. However, DNA variants at 5p15 were associated with histologic subtypes of lung cancer, with an increased frequency of the risk allele in cases with adenocarcinoma. This observation was subsequently confirmed in a meta-analysis published in 2009 (21). This study provides compelling evidence that the 5p15 susceptibility locus for lung cancer is confined to a more specific subtype of lung cancer, that is, adenocarcinoma. Particularly intriguing in that study of more than 30,000 subjects is the absence of new genomic regions associated with lung cancer. In 2010, a meta-analysis of 16 GWAS confirmed lung cancer loci on 15q25, 5p15, and 6p21 (22). Again, the association at 5p15 was confined to adenocarcinoma, whereas the 6p21 locus was more strongly

Table 2. Susceptibility loci for lung cancer identified by GWAS in Asian ancestry populations

Reference	Study[a]	Sample size (cases/controls)	Disease/trait	Platform (# SNPs)	Region (size)	Gene	Key SNPs
Yoon et al. (26)	Korea	621/1,541	NSCLC	Affymetrix 5.0 (246,758)	3q29	*C3orf21*	rs2131877
	Korea replication	*804/1,470*					rs10433328
							rs952481
							rs4677657
					5p15	*TERT*	rs2736100
						CLPTM1L	rs402710
							rs401681
Miki et al. (27)	Japanese	1,004/1,900	Adenocarcinoma	Illumina HumanHap610-Quad and HumanHap550 (432,024)	5p15	*TERT*	rs2736100
	Japanese	*525/7,678*			3q28	*TP63*	rs10937405
	Korean	*569/1,470*					rs4488809
							rs9816619
							rs4600802
Hu et al. (28)	Han Chinese (Nanjing, Beijing, Shanghai)	2,331/3,077	Lung cancer	Affymetrix 6.0 (591,370)	3q28	*TP63*	rs4488809
							rs10937405
	1st stage replication	*2,283/2,243*					
	2nd stage replication	*4,030/4,166*			5p15.33	*TERT*	rs465498
						CLPTM1L	rs2736100
					13q12.12	*MIPEP*	rs753955
						TNFRSF19	
					22q12.2	*MTMR3*	rs17728461
						HORMAD2	rs36600
						LIF	
Dong et al. (29)	Chinese (GWAS)	2,331/3,077	Lung cancer	Affymetrix 6.0 (591,370)	10p14	*GATA3*	rs1663689
	Chinese (stage 1)	*2,283/2,243*					
	Chinese (stage 2)	*5,153/5,240*					
					5q32	*PPP2R2B*	rs2895680
						STK32A	
						DPYSL3	
					20q13.2	*CYP24A1*	rs4809957
							rs2296239
					5q31.1	*IL3*	rs247008
						CSF2	
						P4HA2	
						SLC22A5	
						ACSL6	
					1p36.32	*AJAP1*	rs9439519
						NPHP4	
Shiraishi et al. (32)	Japanese	1,695/5,333	Adenocarcinoma	Illumina OmniExpress & Omni1-Quad (538,166)	5p15.33	*TERT*	rs2736100
	1st validation	*2,955/7,036*					rs2853677
	2nd validation	*1,379/1,166*					
					3q28	*TP63*	rs10937405
					17q24.3	*BPTF*	rs7216064
					6p21.3	*BTNL2*	rs3817963
Dong et al. (30)	Han Chinese	833/3,094	Squamous cell carcinoma	Affymetrix 6.0 (570,009)	12q23.1	*NR1H4*	rs12296850
	Replication 1	*822/2,243*				*SLC17A8*	
	Replication 2	*1,401/4,166*					
Jin et al. (31)	Han Chinese	1,341/1,982	Lung cancer	Illumina HumanExome (72,423)	6p21.33	*PRRC2A* (*BAT2*)	rs9469031 (P515L)
	Replication 1	*1,115/1,246*				*FKBPL*	rs200847762 (P137L)
	Replication 2	*3,584/3,669*			20q11.21	*BPIFB1*	rs6141383 (V284M)
					6p22.2	*HIST1H1E*	rs2298090 (L152R)
Never smokers							
Hsiung et al. (35)	GELAC (Han Chinese)	584/585	Lung adenocarcinoma in never-smoking Asian females	Illumina HumanCNV370-Duo and HumanHap610 Quad (457,504)	5p15.33	*CLPTM1L*	rs2736100
	GELAC (replication)	610/560				*TERT*	
	CAMSCH	287/287					
	SNU	259/293					
	SWHS	209/213					
	WHLCS	207/207					
	KNUH	121/119					
	KUMC	95/87					
	GEL-S	193/546					
	NJLCS	203/203					

(Continued on the following page)

Table 2. Susceptibility loci for lung cancer identified by GWAS in Asian ancestry populations (Cont'd)

Reference	Study[a]	Sample size (cases/controls)	Disease/trait	Platform (# SNPs)	Region (size)	Gene	Key SNPs
Lan et al. (36)	Female Lung Cancer Consortium in Asia	5,510/4,544 *1,099/2,913*	Lung cancer in never smokers	Illumina (512,226)	10q25.2 6q22.2 6p21.32 5p15.33 3q28 17q24.3	*VTI1A* *ROS1, DCBLD1* HLA class II region *TERT* *TP63* *BPTF*	rs7086803 rs11196080 rs9387478 rs2395185 rs2736100 rs4488809 rs7216064
Wang et al. (37)	Female Lung cancer Consortium in Asia	6,877/6,277 *5,878/7,046*	Lung cancer in never smokers	Illumina (7,564,751)	6p21.1 9p21.3 12q13.13	*FOXP4* *FOXP4-AS1* *CDKN2B* *CDKN2B-AS1* *ACVR1B*	rs7741164 rs72658409 rs116101143
Ahn et al. (38)	Korean	446/497 434/1,000	NSCLC in never smokers	Affymetrix 6.0 (474,503)	18p11.22	*FAM38B (PIEZO2)* *APCDD1* *NAPG*	rs11080466 rs11663246
Kim et al. (39)	Korean *Replication 1* *Replication 2*	285/1,455 *293/495* *546/744*	Lung cancer in never smoker women	Affymetrix 5.0 (331,088)	2p16.3	*NRXN1*	rs10187911
Genome-wide epistasis							
Chu et al. (52)	Han Chinese *Replication 1* *Replication 2*	2,331/3,077 *1,534/1,489* *2,512/2,449*	Lung cancer	Affymetrix 6.0 (591,370) (epistatis)	2q32.2	*HIBCH* *INPP1* *PMS1* *STAT1*	rs2562796 rs16832404
Cross-cancer loci							
Jin et al. (54)	Han Chinese *Han Chinese*	5,368/4,006 *9,001/11,436*	Cross-cancer	Affymetrix 6.0	6p21.1 7p15.3	*LRFN2* *SP4* *DNAH11*	rs2494938 rs2285947

Abbreviations: CAMSCH, Chinese Academy of Medical Sciences Cancer Hospital Study; GELAC, Genetic Epidemiological Study of Lung Adenocarcinoma; GEL-S, Genes and Environment in Lung Cancer, Singapore study; KNUH, Kyungpook National University Hospital Study; KUMC, Korea University Medical Center Study; NJLCS, Nanjing Lung Cancer Study; SNU, Seoul National University Study; SWHS, Shanghai Women's Health Cohort Study; WHLCS, Wuhan Lung Cancer Study.
[a]Italic text indicates replication cohorts.

associated with squamous cell carcinoma. Stratification by histology identified three loci for squamous cell carcinoma including 12q13.33, 9p21.3, and 2q32.1. In 2014, another GWAS meta-analysis taking advantage of the imputation based on the 1000 Genomes Project was performed (23), which allowed testing for less frequent SNPs not measured in earlier studies. The top nine signals were followed-up and rare genetic variants were associated with squamous cell carcinoma in the *BRCA2* gene on 13q13.1 and in *CHEK2* on 22q12.1. *CHEK2* was previously associated with lung cancer (24), but this was the first time using a GWAS approach. The 3q28 locus was associated with lung adenocarcinoma, which has been previously found in Asian populations (see the following section). Finally, the latest and largest lung cancer GWAS in individuals of European ancestry was performed in 29,266 cases and 56,450 controls (25). This GWAS highlighted the genetic heterogeneity across histologic subtypes of lung cancer and reported novel loci for lung cancer *per se* (1p31.1, 6q27, 8p21.1, and 15q21.1) and adenocarcinoma (8p12, 10q24.3, 11q23.3, and 20q13.33). Previously reported lung cancer loci were more specifically associated with squamous cell carcinoma in this study including 6p21.33, 12p13.33, and 22q12.1.

GWAS in Asian populations

Genetic heterogeneity in lung cancer susceptibility is observed between populations of European and Asian descent. For example, the strongest lung cancer susceptibility variants on 15q25 have very low allele frequencies in Asian populations. Similarly, variants on 6p21 found in Europeans are not polymorphic in Asians. Accordingly, GWAS specific for Asian populations were required.

In a Korean population, a GWAS on NSCLC revealed a new locus on chromosome 3q29 (26). This study also confirmed the 5p15 susceptibility locus in Koreans. Other GWAS in Asian populations have followed. Two susceptibility loci were identified in Japanese and Korean populations confirming 5p15 and elucidating a new locus on 3q28 (27). The 5p15 and 3q28 were subsequently confirmed in a larger GWAS in Han Chinese (28). In addition, two new loci on 13q12.12 and 12q12.2 were identified. In the same GWAS, but with an extended validation sample size, five new lung cancer loci were identified including 10p14, 5q32, 20q13.2, 5q31.1, and 1p36.32 (29). A subsequent GWAS specifically for lung squamous cell carcinoma in Han Chinese revealed a new locus on 12q23.1 (30). Using the exome genotyping chip, rare variants on 6p21.33, 20q11.21, and 6p22.2 were also found in the Chinese populations (31). On 6p21.33, two missense variants, one in *PRRC1A* (also known as *BAT2*) and the other in *FKBPL*, were independently associated with the risk of lung cancer, suggesting more than one genetic signal in this region. This study also demonstrated that 6p21.33 is also a susceptibility locus for Asian populations, but with different risk variants. The aforementioned GWAS in Japanese population (27)

was later expanded in terms of sample size and SNP coverage to identify a new locus on 17q24.3 (32). The association with lung adenocarcinoma was also confirmed for 5p15, 3q28, and 6p21, but not for 13q12.12 and 22q12.2.

GWAS in never smokers

Lung cancer in never smokers is known to be a distinct entity (33). The first GWAS on lung cancer in never smokers was reported in 2010 (34). A single locus on chromosome 13q31.3 was identified. The lung cancer–associated SNPs were located in the *GPC5* gene and were also associated with mRNA expression levels of this gene in human lung tissues. A subsequent GWAS was performed in never-smoking females from Asia (35). The 5p15 locus was confirmed with an effect size greater than the estimates reported in populations of European background. Interestingly, the 15q25 and 6p21 loci were not associated with lung cancer in this study and no new loci were identified. In a larger GWAS of the same population forming the Female Lung Cancer Consortium in Asia, new susceptibility loci were revealed at 10q25.2 and 6q22.2 (36). The 6p21 was also associated with lung cancer in this study, but significant markers were not in LD with those previously reported, suggesting again more than one independent genetic signal at 6p21. This study in Asian females also confirmed other loci reported before including 5p15, 3q28, and 17q24.3. A recent meta-analysis was reported with an extended sample size of the Female Lung Cancer Consortium in Asia (37). A new locus on 12q13.13 was identified as well as genetic variants not correlated with lung cancer-SNPs previously associated with lung cancer on 6p21.1 and 9p21.3. In never smokers from Korea, a new NSCLC locus on 18p11 was identified (38). The 2p16.3 locus was also suggested in nonsmoking Korean women (39). However, previous loci identified in never smoker populations were not replicated in these Korean studies including 5p15 and 13q31.3.

Pathway-based GWAS

Pathway-based analyses have been used to identify lung cancer loci. Using GWAS data, genes listed under the category of inflammation were evaluated by lung cancer histologic subtypes (40). This analysis identified a risk locus on chromosome 12p13.33 harboring the *RAD12* gene. A similar approach was used to evaluate SNPs in inflammatory pathway genes in lifetime never smokers (41). SNPs on chromosome 12q13 in the *ACVR1B* and *NR4A1* genes were associated with lung cancer, particularly in women and those who reported environmental tobacco smoke exposure. Focused on DNA repair genes, a recent study revealed variants in *GTF2H4* on 6p21 and in *XRCC4* on 5q14.2 associated with lung cancer risk (42). Pathway-based analyses of GWAS data have also identified groups of genes linked by known biological pathways (ABC transporters, VEGF signaling, G_1–S check point, and NRAGE signals death through JNK) that were modestly, but coordinately associated with the risk of developing lung cancer (43).

Variant prioritization approaches

New lung cancer loci were also revealed by incorporating an intermediate phenotype, that is, smoked cigarettes per day, into the analyses (44). By combining the estimates derived from the case–control analysis and the intermediate phenotype, a stronger signal was observed on 15q25 locus compared with the case–control study alone. Genetic associations with lung cancer were also detected on 19q13 and 3p26, which demonstrated improved power to identify genetic loci by combining different types of data from a single population. Studying cohorts of patients well-characterized for lung cancer may thus be very promising using this approach. A similar approach was used by assigning higher priors to SNPs associated with family history of lung cancer (45). By focusing on SNPs missed by traditional GWAS, this study identified 30 variants that showed evidence of association with lung cancer risk. The strongest associations were found on 10q23.33 and 4p15.2. Biological priors within a Bayesian framework were also applied to histology-specific analyses (46). In this study, the 4p15.2 locus was assigned more specifically to squamous cell carcinoma and a new adenocarcinoma locus was identified on 18q12.1.

GWAS-by-exposure interaction

Accounting for environmental exposure is challenging owing to the large number of possible factors as well as the level of measurement accuracy that can be achieved for each exposure. Despite these challenges, some genome-wide gene–environment interaction studies are starting to emerge in the field of lung cancer. The first attempt of a genome-wide gene–smoking interaction study identified two SNPs on 14q22.1 and 15q22.32 influencing the risk of lung cancer (47). For asbestos exposure, interacting loci were suggested on 2q34, 7q32.1, and 11q13 (48). A risk locus for asbestos-associated lung cancer was also discovered on 22q13.31 (49). Interacting loci with household air pollution caused by solid fuel burning for heating and cooking were also evaluated in never smoker women from Asia (50). Interestingly, interactions were reported for GWAS-nominated loci previously identified in this population (36), but no new loci reached significance at the genome-wide scale level. Exploratory analyses of gene–occupation interactions in determining lung cancer susceptibility were also performed for 17 established or suspected lung carcinogens and 49 additional occupational agents (51). A large number of gene–environment interactions were reported in that study. However, the results could not be validated in an independent population because of the uniqueness of the dataset with detailed occupational exposure data. So far, results from genome-wide gene–environment studies in lung cancer have been more hypothesis-generating owing to limited sample size and power as well as the lack of appropriate replication sets. To make further progress, extra care will be needed to build large cohorts that are well-characterized for environmental exposures.

Genome-wide epistasis

The effects of genetic variants on lung cancer are likely to be amplified when multiple variants synergize together. Gene–gene interactions may identify genetic determinants of lung cancer. The first and only genome-wide two-locus interaction analysis performed so far revealed a significant interaction between two SNPs 60 kilobases apart on 2q32.2 (52). Individually, the two interacting SNPs were not significantly associated with the risk of lung cancer. Further investigations of gene–gene interactions will be needed to understand the genetic architecture of lung cancer.

Cross-cancer susceptibility loci

Large-scale GWAS across cancer sites have been conducted to identify pleiotropic loci. For lung cancer, the first

pleiotropic locus was identified on 5p15 (*TERT-CLPTM1L*; ref. 53). A novel pleiotropic association at 7p15.3 was found in Han Chinese involving lung cancer, non-cardia gastric cancer, and esophageal squamous cell carcinoma (54). The GAME-ON/GECCO Network on lung, ovary, breast, prostate and colorectal cancer then identified novel pleiotropic associations involving lung cancer on 12q24 (55) and 1q22 (56). Known lung cancer loci were also identified in cross-cancer analyses including 6p21 (54) and 5p15 (55) as well as 9p21.3 and 13q13.1 (56). These loci are particularly promising to reveal shared carcinogenesis mechanisms across multiple cancer sites.

Integration of GWAS on Lung Cancer Susceptibility

Excluding gene–environment loci that are more suggestive at this point, GWAS reported 45 loci associated with lung cancer. Figure 1 shows the chronologic and cumulative number of lung cancer susceptibility loci identified. Loci are also listed on the basis of chromosome number in Table 3. Note that these loci are an evolving list. The strength of evidence for association with lung cancer and effect size vary by loci (Fig. 2). Evidence supporting some loci is relatively modest and will require validation in independent studies. The magnitude of genetic associations

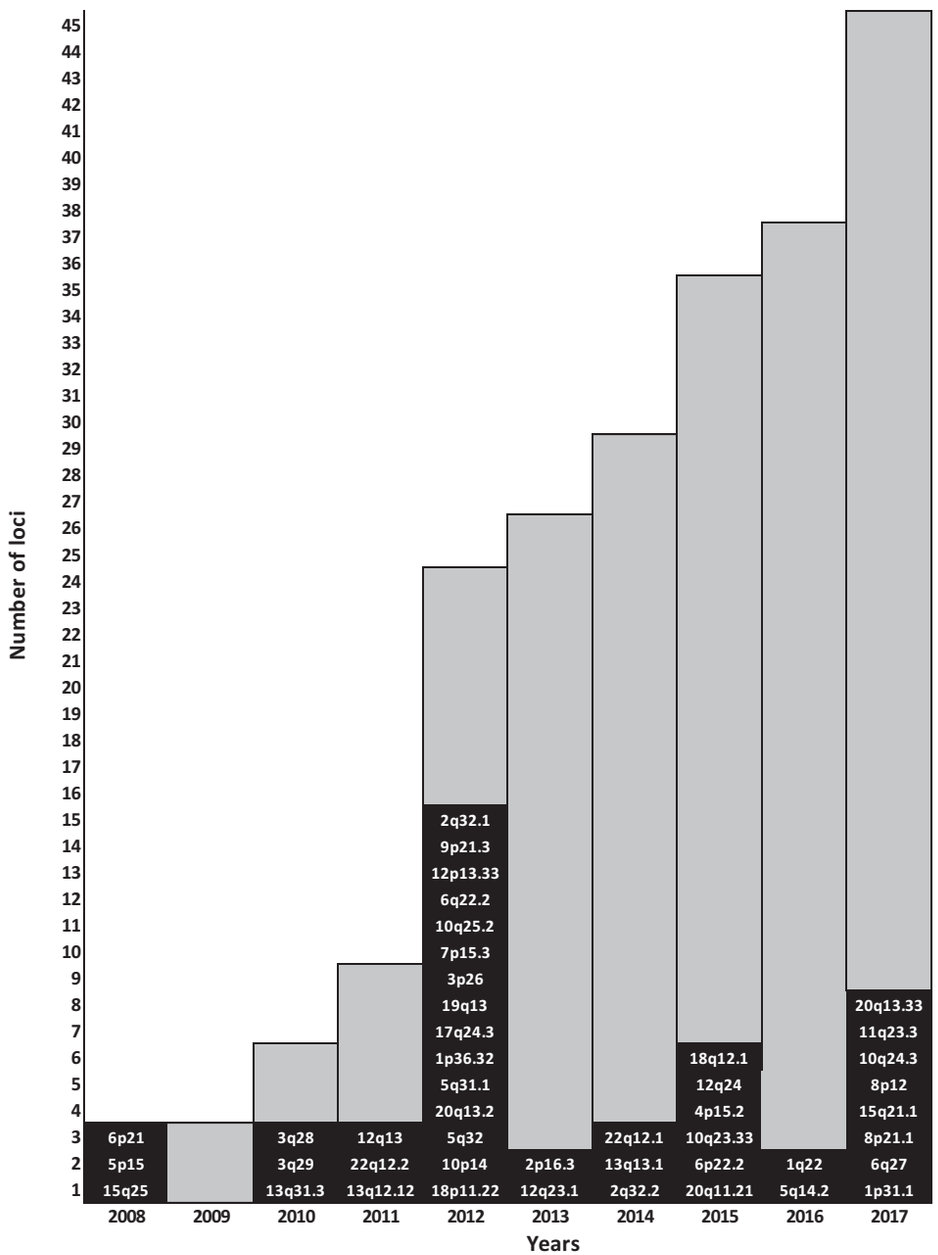

Figure 1.
Chronologic and cumulative lung cancer susceptibility loci identified by GWAS. New loci identified yearly are on a black background. The cumulative number of loci is shown with gray background.

Table 3. Lung cancer susceptibility loci derived from GWAS, genetic risk specificity by subgroups, and suspected causal genes

GWAS loci	Histology	Smoking	Gender	Ethnicity	Age of onset	>1 loci	Suspected causal genes
1p36.32 (29)		s	m	a			*AJAP1* (87), *NPHP4* (88)
1p31.1 (25)				e			*FUBP, DNAJB4*
1q22 (56)	SQ						*MUC1* (89, 90), *ADAM15* (91), *THBS3*
2p16.3 (39)	AD	n	w	a			*NRXN1* (92)
2q32 (22)	SQ			e			*NUP35* (93)
2q32.2 (52)	AD	s	m	a			*HIBCH, INPP1* (94), *PMS1, STAT1*
3p26 (44)				e			No genes. Deletions associated with cancer (95, 96)
3q28 (23, 25, 27, 28, 32, 36)	AD		w				*TP63* (97, 98)
3q29 (26)				a			*C3orf21* (26)
4p15.2 (45, 46)	SQ			e			*KCNIP4* (99)
5p15 (18–22, 25–28, 32, 35, 36, 55)	AD	n	w				*TERT* (100–105), *CLPTM1L* (106–110)
5q14.2 (42)				e			*XRCC4* (111)
5q31 (29)				a	o		*PAHA2* (112), *CSF2* (113), *IL3* (113), *SLC22A5* (29, 114), *ACSL6* (115)
5q32 (29)	AD	n	w	a			*STK32A* (29), *PPP2R2B* (116, 117), *DPYSL3* (118)
6p22.2 (31)				a			*HIST1H1E*
6p21 (18, 20–22, 25, 31, 32, 36, 37, 42, 54)	SQ			e			*BAG6* (119, 120), *APOM* (121, 122), *TNXB* (123), *MSH5* (124), *BTNL2, PRRC2A* (*BAT2*), *FKBPL* (125, 126), *HSPA1B* (127), *FOXP4, FOXP4-AS1, GTF2H4* (42), *LRFN2, HLA-A* (128), *HLA-DQB1* (128)
6q22 (36)	AD	n	w	a			*DCBLD1* (129, 130), *ROS1* (131–133)
6q27 (25)				e			*RNASET2*
7p15.3 (54)				a			*SP4, DNAH11*
8p21.1 (25)				e			*EPHX2, CHRNA2*
8p12 (25)	AD			e			*NRG1* (134)
9p21.3 (22, 25, 37, 56)							*CDKN2A* (135), *CDKN2B* (135), *CDKN2B-AS1* (136, 137), *MTAP*
10p14 (29)				a			*GATA3* (138–141)
10q23.33 (45)		n	w	e	y		*FFAR4* (142)
10q24.3 (25)	AD			e			*OBFC1*
10q25.2 (36)	SQ		w	a			*VTI1A* (143–145)
11q23.3 (25)	AD			e			*MPZL3, AMICA1*
12p13.33 (22, 25, 40)	SQ			e			*RAD52* (40, 146–151)
12q13.13 (37, 41)		n	w				*ACVR1B* (152, 153), *NR4A1*
12q23.1 (30)	SQ			a			*NR1H4* (154), *SLC17A8* (155, 156)
12q24 (55)				e			*SH2B3*
13q12.12 (28)				a	y		*MIPEP, TNFRSF19* (157)
13q13.1 (23, 25, 56)	SQ			e			*BRCA2* (158)
13q31.3 (34)		n		e			*GPC5* (159, 160)
15q21.1 (25)	AD			e			*SEMA6D, SECISBP2L* (161)
15q25 (14–22, 25, 44)		s		e			*CHRNA5, CHRNA3, CHRNB4, IREB2, PSMA4* (162), *HYKK*
17q24.3 (32, 36)				a			*BPTF* (32, 163, 164)
18p11.22 (38)		n		a			*FAM38B* (165), *APCDD1* (166, 167), *NAPG*
18q12.1 (46)	AD			e			*GAREM* (168)
19q13.2 (25, 44)				e			*TGFB1* (169), *CYP2A6*
20q11.21 (31)				a			*BPIFB1* (170)
20q13.2 (29)	AD			a			*CYP24A1* (171–174)
20q13.33 (25)	AD			e			*RTEL1* (175)
22q12.1 (23, 25)	SQ			e			*CHEK2* (24, 176)
22q12.2 (28)				a			*LIF* (177, 178), *HORMAD2, MTMR3*

NOTE: The color of the background illustrates the strength of evidence on a black-and-white scale, where black indicates convincing evidence and white indicates no evidence so far. The strength of evidence was assigned on the basis of the content of publications cited in the first column and our best possible judgment and comprehension of each locus considering the number of studies that replicated the associations, the level of statistically significance, and the quality of the studies, for example, sample size. No evidence (white) does not delineate the lack of associations that have been studied from those not yet examined and also highlights knowledge gaps.

Abbreviations: AD, adenocarcinoma; a, Asians; e, Europeans; m, men; n, never-smokers; o, older; s, smokers; SQ, squamous cell carcinoma; w, women; y, younger.

reported in publications also varies within loci. For example, the largest OR for chromosome 15q25 was 7.2 reported in familial form of lung cancer with a relatively modest *P* value of 1.03×10^{-3} (17). On the other hand, the same locus was reported highly significant ($P = 3.08 \times 10^{-103}$) with an OR of approximately 1.3 (25), which is an effect size more consistent with most studies on sporadic form of lung cancer. The maximum OR and *P* value per locus as well as variability between studies are illustrated in Fig. 2. Refining susceptibility loci by clinically relevant subgroups is a critical step to reveal functional variants and causative genes (57). Table 3 summarizes the evidence supporting the specificity of lung cancer risk loci by histology, smoking status, gender, ethnicity, and age of onset. For example, convincing evidence supports that the

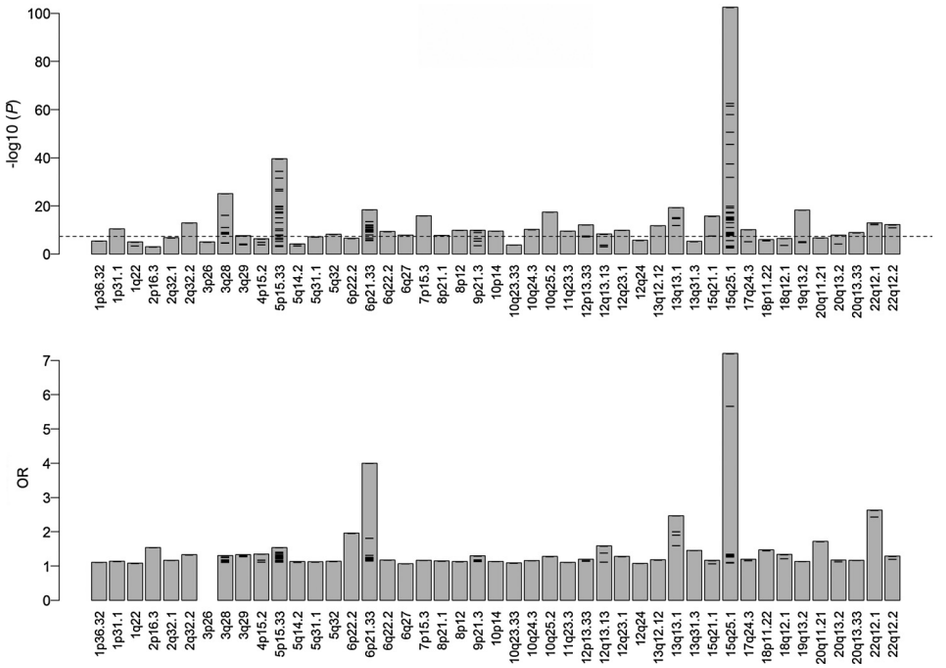

Figure 2.
Magnitude of the genetic associations for the 45 lung cancer susceptibility loci reported in the literature. *P* values and OR for key SNPs indicated in Tables 1 and 2 were collected from reported studies. The top and bottom panels show *P* values and OR, respectively. The *x*-axis shows the 45 lung cancer susceptibility loci ordered on the basis of chromosome position. Bars illustrate the maximum *P* value (smallest *P* values) or the maximum OR reported per locus. Tick symbols inside bars are the results of the individual studies. Top, the *y*-axis represents *P* values in -log10 scale. The dashed horizontal line shows the typical GWAS significance threshold, that is, 5×10^{-8}. Bottom, the *y*-axis represents OR. Note that ORs lower than 1 were converted into their reciprocal (1/OR) for illustration purpose. OR for locus 3p26 was not available.

5p15 locus is specific to lung adenocarcinoma and more strongly associated with never smokers and women (21, 22, 28). Similar genetic association patterns are also emerging for 2p16.3, 5q32, and 6q22. Accordingly, a number of studies have started to delineate the effects of genetic variants in specific subgroups of patients with lung cancer, which is important to reveal the true nature of genetic effects detected in GWAS and narrow the set of genetic variants and genes worthy of functional studies. It is also important to know whether independent variants in the same loci are associated with lung cancer. Convincing evidence supports at least two independent loci on 15q25, 5p15, and 6p21 (Table 3). Two independent loci were also reported at 9p21.3 (37) and 22q12.2 (28), but will require further validation. This knowledge is lacking for other lung cancer loci. Table 3 also provides a glimpse of suspected causal genes at each locus. Further functional and biological analyses will be needed to understand the role of these genes in lung cancer development.

GWAS on Lung Cancer Survival

Interindividual differences in lung cancer survival are observed among lung cancer patients, even among those with the same tumor stage and treatment regimen. The identification of genetic factors associated with lung cancer survival has the potential to guide adjuvant therapy after surgery in early-stage disease, but also to refine prognosis and personalize clinical care in advanced-stages disease. So far, GWAS on lung cancer survival were performed in patients with early-stage NSCLC (58, 59), advanced-stage NSCLC (60–63), and SCLC (64). GWAS were also performed in more specific subgroups of lung cancer patients including never smokers with NSCLC (65) and patients with lung adenocarcinoma (66). Together, these studies have identified 23 loci associated with lung cancer survival (Supplementary Table S1). However, none of these loci was reported in more than one study. The lack of replication may be explained by heterogeneity in treatment regimens. There is also no overlap with GWAS lung cancer susceptibility loci. While the 9p21.3 locus was associated with both susceptibility (22, 37) and survival (62), sentinel SNPs are located more than 1 Mb away from each other, indicating that they are likely not reflecting the same association. GWAS susceptibility loci were specifically evaluated for association with survival in SCLC (67). Briefly, three loci on 20q13.2, 22q12.2, and 5p15 demonstrated some evidence of association with survival. However, none reached genome-wide significance. It should be noted that clinical follow-up of patients are needed to conduct survival analyses and GWAS based on this outcome have thus been performed with much smaller sample sizes compared with studies focused on cancer susceptibility. Larger-scale studies are needed to identify robust lung cancer survival loci.

GWAS on Response to Lung Cancer Therapies

Somatic alterations in the tumor genome are known to modulate the response to anticancer therapy. Less is known about the influence of the host genome on treatment response. The effect of germline variants on sensitivity and toxicity to platinum-based chemotherapy have been examined by GWAS. In patients with SCLC, seven loci demonstrated some evidence of association with treatment response in a discovery set, but were not convincingly replicated in a validation set (68). In NSCLC patients, a locus on 21q22.3 was associated with platinum-induced hepatotoxicity (69) and two loci on 2q24.3 and 17p12 were associated with the risk of platinum-induced myelosuppression (70). More and larger studies are needed to effectively delineate chemosensitive patients that will benefit from treatment and nonresponders that may be spared the adverse side effects associated with chemotherapy.

Future Directions

Important progress was made to understand host susceptibility to lung cancer using GWAS. This approach is also starting to reveal inherited variants associated with lung cancer survival and response to treatment. During the last decade, progress was driven by enlarging sample sizes, improving methods to genotype and impute SNPs with more comprehensive reference sets (e.g., 1000 Genomes Project), and the creation of large-scale international collaborations and consortia. Additional developments were made by studying patients of different ancestries, never smokers, women, and specific lung cancer histology. Progress was also made by refining results by pathway-based analysis and variant prioritization approaches. The next important step in the field of genomics of lung cancer is to identify the causal genetic variants and genes underpinning GWAS-nominated loci. In addition, the new genomic knowledge must be translated into real benefits for patients. These must be achieved if GWAS are to pay off the huge investment. We foresee different strategies to reach these goals.

eQTL and TWAS

We need to continue to mine GWAS data using more advanced statistical techniques that leverage other sources of data. So far, expression quantitative trait loci (eQTL) mapping studies in a variety of tissues have been used to extent the functional meaning of GWAS in lung cancer (23, 25, 29, 30, 32, 34, 37, 45, 56, 71). More comprehensive methods of colocalization of GWAS and eQTL signals were recently developed and must be performed to reveal genetic associations explained by regulatory effects on gene expression (72). The identification of lung cancer-associated genetic variants associated with the expression of specific genes in a disease relevant tissue is an important step forward to understand the molecular mechanisms underpinning GWAS signals. In addition, the relationships between genetic variants, RNA expression levels, and lung cancer must be further delineated by causality models and Mendelian randomization approaches (73, 74). Large-scale lung cancer GWAS (25) and lung eQTL (75) are also available to perform the first transcriptome-wide association study (TWAS) in lung cancer. In this approach, the *cis* genetic component of expression derived from the eQTL dataset is used to impute expression data for cases and controls used in the GWAS. Imputed genome-wide gene expression levels of sample size orders of magnitude larger than any of the transcriptomic datasets generated so far can then be used to identify genes whose expression is significantly associated with the disease. This approach has the potential to elucidate the most likely molecular drivers of lung cancer in GWAS-nominated loci, but also yield molecular drivers of lung cancer outside GWAS loci. GWAS, eQTL, and TWAS results will also need to be integrated with genes differentially expressed in lung tumor compared with adjacent nontumor lung tissues. For example, we have recently derived a robust list of genes differentially expressed in lung tumor from our own transcriptomic dataset (76) as well as two publicly available datasets (77, 78). These results identified genes consistently deregulated in lung tumor and revealed important insights about the molecular transitions that occur between normal and tumor lung tissues. Accordingly, very promising research is underway exploiting GWAS and gene expression datasets to identify causal genes and molecular drivers of lung cancer.

Deep molecular profiling and biobanking

To make further progress, relevant tissues must be profiled beyond gene expression. GWAS variants of lung cancer may not exert their effects through gene regulation, but other molecular phenotypes such as protein expression, protein state, metabolite levels, and epigenetic marks. Accordingly, deep molecular profiling of human lung tissues will be needed to comprehend the molecular impact of inherited variants on lung cancer. Current biobanking activities to collect high-quality and large numbers of well-annotated lung specimens are the essence of this future development.

Exposome

As depicted in Table 3, the independent contribution of GWAS loci on lung cancer, smoking behavior, and nicotine dependence is still not clearly delineated. Larger-scale gene–exposure interaction studies with established environmental risk factors including tobacco smoke and solid fuel burning are warranted. More comprehensive assessment of environmental factors including radon, asbestos, household and outdoor pollution, and occupational agents will be critically important, but at the same time very challenging to measure accurately in large sample size. A well-orchestrated community effort thus seems necessary (79).

Exome and genome sequencing

New genomic approaches from next-generation of sequencers are also expected to refine GWAS loci and discover new variants unlikely to be found by GWAS. To this effect, whole-exome sequencing in three members of a five-generation family affected by lung cancer has revealed a rare variants in *PARK2* resulting in a loss-of-function of this tumor suppressor gene (80). Although rare, the effect size of this mutation was greater than those reported in GWAS. Similarly, exome-sequencing of sporadic and familial cases of lung cancer identified rare deleterious mutations in GWAS-nominated loci located in the *CDC147* and *DBH* (81) genes. Whole-genome sequencing in a family with very high aggregation of lung adenocarcinoma revealed a functional missense variant in the oncogene *YAP1* (82) associated with the risk of developing the disease. We expect these types of discoveries using exome and genome sequencing to multiply in the near future.

En route for a genetic risk score

Identified lung cancer susceptibility loci provide hope to build tools for targeted screening of high-risk individuals. To date, cumulative effects of loci have shown promising results to improve the discriminatory performance of risk prediction models, but not sufficiently to merit clinical implementation (83). For example, a recent report combining GWAS loci demonstrated only small improvement in lung cancer risk prediction in models including basic clinical factors such as age and smoking (84). Interestingly, the best model may not come from considering only the top GWAS loci. A genetic risk score built from seven telomere-length associated genetic variants was associated with lung cancer risk (85). More recently, it was demonstrated that the cumulative effects of susceptibility variants were better predictors when organized in biological pathways (86). These examples demonstrated the variety of strategies that are currently used to develop new clinical tools to predict lung cancer. Such tools are urgently needed to enable earlier diagnosis. The task is challenging and will require major efforts, but seems more realistically feasible with the outcomes

of GWAS in hands. We hope that this compendium of lung cancer GWAS loci will facilitate further progress in building a clinically useful genetic risk score.

Conclusions

Understanding the genetic factors underlying the development of lung cancer is important to elucidate the etiology of the disease. This genetic knowledge is a prerequisite to develop and improve future clinical strategies for lung cancer management. Discovered loci summarized in this review testify progress made in this field during the last decade. This review also highlights knowledge gaps about causal variants and genes responsible for the underlying genetic associations and proposes some short-term solutions to ensure further progress through eQTL, colocalization, causality models and TWAS. The specificity of many lung cancer loci in terms of histologic subtypes, gender, and ethnicity have been discovered for some loci, but will demand large studies with well-characterized individuals for others. Although smoking and other environmental factors, notably solid fuel burning, are clearly interacting with host factors to cause the disease, the specific variants that come into play are still elusive. Preliminary data provides some clues about inherited variants associated with lung cancer survival and response to treatment, but will require validation in larger-scale studies. On the other hand, robust genetic factors associated with lung cancer derived from GWAS give hope for possible clinical translation. In short term, a genetic risk score to screen high-risk individuals seems realistically achievable and would allow more effective treatments available at earlier stages of the disease. In mid and longer terms, discovering the causal genes underpinning GWAS signals will propel results one step closer to clinical applications by revealing new therapeutic targets and biomarkers to personalize quality of care.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Received October 6, 2016; revised December 6, 2016; accepted April 20, 2017; published first June 14, 2017.

References

1. Herbst RS, Heymach JV, Lippman SM. Lung cancer. N Engl J Med 2008;359:1367–80.
2. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014;64:9–29.
3. Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. Br J Cancer 2005;93:825–33.
4. Tokuhata GK, Lilienfeld AM. Familial aggregation of lung cancer in humans. J Natl Cancer Inst 1963;30:289–312.
5. Cote ML, Liu M, Bonassi S, Neri M, Schwartz AG, Christiani DC, et al. Increased risk of lung cancer in individuals with a family history of the disease: a pooled analysis from the International Lung Cancer Consortium. Eur J Cancer 2012;48:1957–68.
6. Schwartz AG, Prysak GM, Bock CH, Cote ML. The molecular epidemiology of lung cancer. Carcinogenesis 2007;28:507–18.
7. Yokota J, Shiraishi K, Kohno T. Genetic basis for susceptibility to lung cancer: recent progress and future directions. Adv Cancer Res 2010;109:51–72.
8. Li W, Li K, Zhao L, Zou H. DNA repair pathway genes and lung cancer susceptibility: a meta-analysis. Gene 2014;538:361–5.
9. Hwang SJ, Cheng LS, Lozano G, Amos CI, Gu X, Strong LC. Lung cancer risk in germline p53 mutation carriers: association between an inherited cancer predisposition, cigarette smoking, and cancer risk. Hum Genet 2003;113:238–43.
10. Kleinerman RA, Tarone RE, Abramson DH, Seddon JM, Li FP, Tucker MA. Hereditary retinoblastoma and risk of lung cancer. J Natl Cancer Inst 2000;92:2037–9.
11. Bell DW, Gore I, Okimoto RA, Godin-Heymann N, Sordella R, Mulloy R, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. Nat Genet 2005; 37:1315–6.
12. You M, Wang D, Liu P, Vikis H, James M, Lu Y, et al. Fine mapping of chromosome 6q23-25 region in familial lung cancer families reveals RGS17 as a likely candidate gene. Clin Cancer Res 2009; 15:2666–74.
13. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42:D1001–6.
14. Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452:633–7.
15. Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452:638–42.
16. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008;40:616–22.
17. Liu P, Vikis HG, Wang D, Lu Y, Wang Y, Schwartz AG, et al. Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer. J Natl Cancer Inst 2008;100:1326–30.
18. Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet 2008;40:1407–9.
19. McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet 2008; 40:1404–6.
20. Broderick P, Wang Y, Vijayakrishnan J, Matakidou A, Spitz MR, Eisen T, et al. Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. Cancer Res 2009;69: 6633–41.
21. Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet 2009;85:679–91.
22. Timofeeva MN, Hung RJ, Rafnar T, Christiani DC, Field JK, Bickeboller H, et al. Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls. Hum Mol Genet 2012;21:4980–95.
23. Wang Y, McKay JD, Rafnar T, Wang Z, Timofeeva MN, Broderick P, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet 2014;46:736–41.
24. Brennan P, McKay J, Moore L, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, et al. Uncommon CHEK2 mis-sense variant and reduced risk of tobacco-related cancers: case control study. Hum Mol Genet 2007;16: 1794–801.
25. McKay JD, Hung RJ, Han Y, Zong X, Carreras-Torres R, Christiani DC, et al. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet 2017;49:1126–32.
26. Yoon KA, Park JH, Han J, Park S, Lee GK, Han JY, et al. A genome-wide association study reveals susceptibility variants for non-small cell lung cancer in the Korean population. Hum Mol Genet 2010;19: 4948–54.
27. Miki D, Kubo M, Takahashi A, Yoon KA, Kim J, Lee GK, et al. Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. Nat Genet 2010;42: 893–6.

28. Hu Z, Wu C, Shi Y, Guo H, Zhao X, Yin Z, et al. A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. Nat Genet 2011;43:792–6.
29. Dong J, Hu Z, Wu C, Guo H, Zhou B, Lv J, et al. Association analyses identify multiple new lung cancer susceptibility loci and their interactions with smoking in the Chinese population. Nat Genet 2012;44:895–9.
30. Dong J, Jin G, Wu C, Guo H, Zhou B, Lv J, et al. Genome-wide association study identifies a novel susceptibility locus at 12q23.1 for lung squamous cell carcinoma in han chinese. PLoS Genet 2013;9:e1003190.
31. Jin G, Zhu M, Yin R, Shen W, Liu J, Sun J, et al. Low-frequency coding variants at 6p21.33 and 20q11.21 are associated with lung cancer risk in Chinese populations. Am J Hum Genet 2015;96:832–40.
32. Shiraishi K, Kunitoh H, Daigo Y, Takahashi A, Goto K, Sakamoto H, et al. A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. Nat Genet 2012;44: 900–3.
33. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers–a different disease. Nat Rev Cancer 2007;7:778–90.
34. Li Y, Sheu CC, Ye Y, de Andrade M, Wang L, Chang SC, et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. Lancet Oncol 2010;11:321–30.
35. Hsiung CA, Lan Q, Hong YC, Chen CJ, Hosgood HD, Chang IS, et al. The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. PLoS Genet 2010;6:e1001051.
36. Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z, et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. Nat Genet 2012;44:1330–5.
37. Wang Z, Seow WJ, Shiraishi K, Hsiung CA, Matsuo K, Liu J, et al. Meta-analysis of genome-wide association studies identifies multiple lung cancer susceptibility loci in never-smoking Asian women. Hum Mol Genet 2016;25:620–9.
38. Ahn MJ, Won HH, Lee J, Lee ST, Sun JM, Park YH, et al. The 18p11.22 locus is associated with never smoker non-small cell lung cancer susceptibility in Korean populations. Hum Genet 2012;131:365–72.
39. Kim JH, Park K, Yim SH, Choi JE, Sung JS, Park JY, et al. Genome-wide association study of lung cancer in Korean non-smoking women. J Korean Med Sci 2013;28:840–7.
40. Shi J, Chatterjee N, Rotunno M, Wang Y, Pesatori AC, Consonni D, et al. Inherited variation at chromosome 12p13.33, including RAD52, influences the risk of squamous cell lung carcinoma. Cancer Discov 2012;2:131–9.
41. Spitz MR, Gorlov IP, Amos CI, Dong Q, Chen W, Etzel CJ, et al. Variants in inflammation genes are implicated in risk of lung cancer in never smokers exposed to second-hand smoke. Cancer Discov 2011;1:420–9.
42. Wang M, Liu H, Liu Z, Yi X, Bickeboller H, Hung RJ, et al. Genetic variant in DNA repair gene GTF2H4 is associated with lung cancer risk: a large-scale analysis of six published GWAS datasets in the TRICL consortium. Carcinogenesis 2016;37:888–96.
43. Lee D, Lee GK, Yoon KA, Lee JS. Pathway-based analysis using genome-wide association data from a Korean non-small cell lung cancer study. PLoS One 2013;8:e65396.
44. Li Y, Huang J, Amos CI. Genetic association analysis of complex diseases incorporating intermediate phenotype information. PLoS One 2012;7: e46612.
45. Poirier JG, Brennan P, McKay JD, Spitz MR, Bickeboller H, Risch A, et al. Informed genome-wide association analysis with family history as a secondary phenotype identifies novel loci of lung cancer. Genet Epidemiol 2015;39:197–206.
46. Brenner DR, Amos CI, Brhane Y, Timofeeva MN, Caporaso N, Wang Y, et al. Identification of lung cancer histology-specific variants applying Bayesian framework variant prioritization approaches within the TRICL and ILCCO consortia. Carcinogenesis 2015;36:1314–26.
47. Zhang R, Chu M, Zhao Y, Wu C, Guo H, Shi Y, et al. A genome-wide gene-environment interaction analysis for tobacco smoke and lung cancer susceptibility. Carcinogenesis 2014;35:1528–35.
48. Wei S, Wang LE, McHugh MK, Han Y, Xiong M, Amos CI, et al. Genome-wide gene-environment interaction analysis for asbestos exposure in lung cancer susceptibility. Carcinogenesis 2012;33:1531–7.
49. Liu CY, Stucker I, Chen C, Goodman G, McHugh MK, D'Amelio AM Jr, et al. Genome-wide gene-asbestos exposure interaction association study identifies a common susceptibility variant on 22q13.31 associated with lung cancer risk. Cancer Epidemiol Biomarkers Prev 2015;24:1564–73.
50. Hosgood HD III, Song M, Hsiung CA, Yin Z, Shu XO, Wang Z, et al. Interactions between household air pollution and GWAS-identified lung cancer susceptibility markers in the Female Lung Cancer Consortium in Asia (FLCCA). Hum Genet 2015;134:333–41.
51. Malhotra J, Sartori S, Brennan P, Zaridze D, Szeszenia-Dabrowska N, Swiatkowska B, et al. Effect of occupational exposures on lung cancer susceptibility: a study of gene-environment interaction analysis. Cancer Epidemiol Biomarkers Prev 2015;24:570–9.
52. Chu M, Zhang R, Zhao Y, Wu C, Guo H, Zhou B, et al. A genome-wide gene-gene interaction analysis identifies an epistatic gene pair for lung cancer susceptibility in Han Chinese. Carcinogenesis 2014; 35:572–7.
53. Rafnar T, Sulem P, Stacey SN, Geller F, Gudmundsson J, Sigurdsson A, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet 2009;41:221–7.
54. Jin G, Ma H, Wu C, Dai J, Zhang R, Shi Y, et al. Genetic variants at 6p21.1 and 7p15.3 are associated with risk of multiple cancers in Han Chinese. Am J Hum Genet 2012;91:928–34.
55. Hung RJ, Ulrich CM, Goode EL, Brhane Y, Muir K, Chan AT, et al. Cross cancer genomic investigation of inflammation pathway for five common cancers: lung, ovary, prostate, breast, and colorectal cancer. J Natl Cancer Inst 2015;107:pii: djv246.
56. Fehringer G, Kraft P, Pharoah PD, Eeles RA, Chatterjee N, Schumacher FR, et al. Cross-cancer genome-wide analysis of lung, ovary, breast, prostate, and colorectal cancer reveals novel pleiotropic associations. Cancer Res 2016;76:5103–14.
57. Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. Nat Rev Genet 2009;10: 318–29.
58. Huang YT, Heist RS, Chirieac LR, Lin X, Skaug V, Zienolddiny S, et al. Genome-wide analysis of survival in early-stage non-small-cell lung cancer. J Clin Oncol 2009;27:2660–7.
59. Tang S, Pan Y, Wang Y, Hu L, Cao S, Chu M, et al. Genome-wide association study of survival in early-stage non-small cell lung cancer. Ann Surg Oncol 2015;22:630–5.
60. Sato Y, Yamamoto N, Kunitoh H, Ohe Y, Minami H, Laird NM, et al. Genome-wide association study on overall survival of advanced non-small cell lung cancer patients treated with carboplatin and paclitaxel. J Thorac Oncol 2011;6:132–8.
61. Wu X, Ye Y, Rosell R, Amos CI, Stewart DJ, Hildebrandt MA, et al. Genome-wide association study of survival in non-small cell lung cancer patients receiving platinum-based chemotherapy. J Natl Cancer Inst 2011;103: 817–25.
62. Hu L, Wu C, Zhao X, Heist R, Su L, Zhao Y, et al. Genome-wide association study of prognosis in advanced non-small cell lung cancer patients receiving platinum-based chemotherapy. Clin Cancer Res 2012;18: 5507–14.
63. Lee Y, Yoon KA, Joo J, Lee D, Bae K, Han JY, et al. Prognostic implications of genetic variants in advanced non-small cell lung cancer: a genome-wide association study. Carcinogenesis 2013;34:307–13.
64. Wu C, Xu B, Yuan P, Miao X, Liu Y, Guan Y, et al. Genome-wide interrogation identifies YAP1 variants associated with survival of small-cell lung cancer patients. Cancer Res 2010;70:9721–9.
65. Wu X, Wang L, Ye Y, Aakre JA, Pu X, Chang GC, et al. Genome-wide association study of genetic predictors of overall survival for non-small cell lung cancer in never smokers. Cancer Res 2013;73: 4028–38.
66. Galvan A, Colombo F, Frullanti E, Dassano A, Noci S, Wang Y, et al. Germline polymorphisms and survival of lung adenocarcinoma patients: a genome-wide study in two European patient series. Int J Cancer 2015; 136:E262–71.
67. Li D, Wei L, Xu B, Yu D, Chang J, Yuan P, et al. Association of GWAS-identified lung cancer susceptibility loci with survival length in patients with small-cell lung cancer treated with platinum-based chemotherapy. PLoS One 2014;9:e113574.
68. Wu C, Xu B, Yuan P, Ott J, Guan Y, Liu Y, et al. Genome-wide examination of genetic variants associated with response to platinum-based chemotherapy in patients with small-cell lung cancer. Pharmacogenet Genomics 2010;20:389–95.

69. Cao S, Wang C, Ma H, Yin R, Zhu M, Shen W, et al. Genome-wide association study on platinum-induced hepatotoxicity in non-small cell lung cancer patients. Sci Rep 2015;5:11556.
70. Cao S, Wang S, Ma H, Tang S, Sun C, Dai J, et al. Genome-wide association study of myelosuppression in non-small-cell lung cancer patients with platinum-based chemotherapy. Pharmacogenomics J 2016;16:41–6.
71. Nguyen JD, Lamontagne M, Couture C, Conti M, Pare PD, Sin DD, et al. Susceptibility loci for lung cancer are associated with mRNA levels of nearby genes in the lung. Carcinogenesis 2014;35:2653–9.
72. Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet 2014;10: e1004383.
73. Lamontagne M, Timens W, Hao K, Bossé Y, Laviolette M, Steiling K, et al. Genetic regulation of gene expression in the lung identifies CST3 and CD22 as potential causal genes for airflow obstruction. Thorax 2014; 69:997–1004.
74. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet 2016;48:481–7.
75. Hao K, Bossé Y, Nickle DC, Pare PD, Postma DS, Laviolette M, et al. Lung eQTLs to help reveal the molecular underpinnings of asthma. PLoS Genet 2012;8:e1003029.
76. Bossé Y, Sazonova O, Gaudreault N, Bastien N, Conti M, Pagé S, et al. Transcriptomic microenvironment of lung adenocarcinoma. Cancer Epidemiol Biomarkers Prev 2017;26:389–96.
77. Selamat SA, Chung BS, Girard L, Zhang W, Zhang Y, Campan M, et al. Genome-scale analysis of DNA methylation in lung adenocarcinoma and integration with mRNA expression. Genome Res 2012;22:1197–211.
78. Landi MT, Dracheva T, Rotunno M, Figueroa JD, Liu H, Dasgupta A, et al. Gene expression signature of cigarette smoking and its role in lung adenocarcinoma development and survival. PLoS One 2008;3:e1651.
79. Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, et al. Mutational signatures associated with tobacco smoking in human cancer. Science 2016;354:618–22.
80. Xiong D, Wang Y, Kupert E, Simpson C, Pinney SM, Gaba CR, et al. A recurrent mutation in PARK2 is associated with familial lung cancer. Am J Hum Genet 2015;96:301–8.
81. Liu Y, Kheradmand F, Davis CF, Scheurer ME, Wheeler D, Tsavachidis S, et al. Focused analysis of exome sequencing data for rare germline mutations in familial and sporadic lung cancer. J Thorac Oncol 2016;11:52–61.
82. Chen HY, Yu SL, Ho BC, Su KY, Hsu YC, Chang CS, et al. R331W missense mutation of oncogene YAP1 is a germline risk allele for lung adenocarcinoma with medical actionability. J Clin Oncol 2015;33: 2303–10.
83. Wang X, Oldani MJ, Zhao X, Huang X, Qian D. A review of cancer risk prediction models with genetic variants. Cancer Inform 2014;13:19–28.
84. Weissfeld JL, Lin Y, Lin HM, Kurland BF, Wilson DO, Fuhrman CR, et al. Lung cancer risk prediction using common SNPs located in GWAS-identified susceptibility regions. J Thorac Oncol 2015;10:1538–45.
85. Machiela MJ, Hsiung CA, Shu XO, Seow WJ, Wang Z, Matsuo K, et al. 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 2015;137:311–9.
86. Qian DC, Han Y, Byun J, Shin HR, Hung RJ, McLaughlin JR, et al. A novel pathway-based approach improves lung cancer risk prediction using germline genetic variations. Cancer Epidemiol Biomarkers Prev 2016;25:1208–15.
87. McDonald JM, Dunlap S, Cogdell D, Dunmire V, Wei Q, Starzinski-Powitz A, et al. The SHREW1 gene, frequently deleted in oligodendrogliomas, functions to inhibit cell adhesion and migration. Cancer Biol Ther 2006;5:300–4.
88. Habbig S, Bartram MP, Muller RU, Schwarz R, Andriopoulos N, Chen S, et al. NPHP4, a cilia-associated protein, negatively regulates the Hippo pathway. J Cell Biol 2011;193:633–42.
89. Nath S, Mukherjee P. MUC1: a multifaceted oncoprotein with a key role in cancer progression. Trends Mol Med 2014;20:332–42.
90. Kharbanda A, Rajabi H, Jin C, Tchaicha J, Kikuchi E, Wong KK, et al. Targeting the oncogenic MUC1-C protein inhibits mutant EGFR-mediated signaling and survival in non-small cell lung cancer cells. Clin Cancer Res 2014;20:5423–34.
91. Kuefer R, Day KC, Kleer CG, Sabel MS, Hofer MD, Varambally S, et al. ADAM15 disintegrin is associated with aggressive prostate and breast cancer disease. Neoplasia 2006;8:319–29.
92. Nussbaum J, Xu Q, Payne TJ, Ma JZ, Huang W, Gelernter J, et al. Significant association of the neurexin-1 gene (NRXN1) with nicotine dependence in European- and African-American smokers. Hum Mol Genet 2008;17: 1569–77.
93. Hawryluk-Gara LA, Platani M, Santarella R, Wozniak RW, Mattaj IW. Nup53 is required for nuclear envelope and nuclear pore complex assembly. Mol Biol Cell 2008;19:1753–62.
94. Li SR, Gyselman VG, Lalude O, Dorudi S, Bustin SA. Transcription of the inositol polyphosphate 1-phosphatase gene (INPP1) is upregulated in human colorectal cancer. Mol Carcinog 2000;27: 322–9.
95. Dasgupta S, Chakraborty SB, Roy A, Roychowdhury S, Panda CK. Differential deletions of chromosome 3p are associated with the development of uterine cervical carcinoma in Indian patients. Mol Pathol 2003;56:263–9.
96. Braga E, Senchenko V, Bazov I, Loginov W, Liu J, Ermilova V, et al. Critical tumor-suppressor gene regions on chromosome 3P in major human epithelial malignancies: allelotyping and quantitative real-time PCR. Int J Cancer 2002;100:534–41.
97. Katoh I, Aisaki KI, Kurata SI, Ikawa S, Ikawa Y. p51A (TAp63gamma), a p53 homolog, accumulates in response to DNA damage for cell regulation. Oncogene 2000;19:3126–30.
98. Flores ER. The roles of p63 in cancer. Cell Cycle 2007;6:300–4.
99. Bonne A, Vreede L, Kuiper RP, Bodmer D, Jansen C, Eleveld M, et al. Mapping of constitutional translocation breakpoints in renal cell cancer patients: identification of KCNIP4 as a candidate gene. Cancer Genet Cytogenet 2007;179:11–8.
100. Janknecht R. On the road to immortality: hTERT upregulation in cancer cells. FEBS Lett 2004;564:9–13.
101. Lantuejoul S, Salon C, Soria JC, Brambilla E. Telomerase expression in lung preneoplasia and neoplasia. Int J Cancer 2007;120:1835–41.
102. Mavrogiannou E, Strati A, Stathopoulou A, Tsaroucha EG, Kaklamanis L, Lianidou ES. Real-time RT-PCR quantification of human telomerase reverse transcriptase splice variants in tumor cell lines and non-small cell lung cancer. Clin Chem 2007;53:53–61.
103. Wu TC, Lin P, Hsu CP, Huang YJ, Chen CY, Chung WC, et al. Loss of telomerase activity may be a potential favorable prognostic marker in lung carcinomas. Lung Cancer 2003;41:163–9.
104. Strazisar M, Mlakar V, Glavac D. The expression of COX-2, hTERT, MDM2, LATS2 and S100A2 in different types of non-small cell lung cancer (NSCLC). Cell Mol Biol Lett 2009;14:442–56.
105. Xie M, Chen Q, He S, Li B, Hu C. Silencing of the human TERT gene by RNAi inhibits A549 lung adenocarcinoma cell growth in vitro and in vivo. Oncol Rep 2011;26:1019–27.
106. Yamamoto K, Okamoto A, Isonishi S, Ochiai K, Ohtake Y. A novel gene, CRR9, which was up-regulated in CDDP-resistant ovarian tumor cell line, was associated with apoptosis. Biochem Biophys Res Commun 2001;280:1148–54.
107. Zienolddiny S, Skaug V, Landvik NE, Ryberg D, Phillips DH, Houlston R, et al. The TERT-CLPTM1L lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. Carcinogenesis 2009; 30:1368–71.
108. James MA, Wen W, Wang Y, Byers LA, Heymach JV, Coombes KR, et al. Functional characterization of CLPTM1L as a lung cancer risk candidate gene in the 5p15.33 locus. PLoS One 2012;7: e36116.
109. Ni Z, Tao K, Chen G, Chen Q, Tang J, Luo X, et al. CLPTM1L is over-expressed in lung cancer and associated with apoptosis. PLoS One 2012;7: e52598.
110. James MA, Vikis HG, Tate E, Rymaszewski AL, You M. CRR9/CLPTM1L regulates cell survival signaling and is required for Ras transformation and lung tumorigenesis. Cancer Res 2014;74:1116–27.
111. Yu H, Zhao H, Wang LE, Han Y, Chen WV, Amos CI, et al. An analysis of single nucleotide polymorphisms of 125 DNA repair genes in the Texas

genome-wide association study of lung cancer with a replication for the XRCC4 SNPs. DNA Repair 2011;10:398–407.
112. Teodoro JG, Parker AE, Zhu X, Green MR. p53-mediated inhibition of angiogenesis through up-regulation of a collagen prolyl hydroxylase. Science 2006;313:968–71.
113. Dougan M, Li D, Neuberg D, Mihm M, Googe P, Wong KK, et al. A dual role for the immune response in a mouse model of inflammation-associated lung cancer. J Clin Invest 2011;121:2436–46.
114. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, et al. Genetics and beyond–the transcriptome of human monocytes and disease susceptibility. PLoS One 2010;5:e10693.
115. Chen J, Brunzell DH, Jackson K, van der Vaart A, Ma JZ, Payne TJ, et al. ACSL6 is associated with the number of cigarettes smoked and its expression is altered by chronic nicotine exposure. PLoS One 2011;6:e28790.
116. Vazquez A, Kulkarni D, Grochola LF, Bond GL, Barnard N, Toppmeyer D, et al. A genetic variant in a PP2A regulatory subunit encoded by the PPP2R2B gene associates with altered breast cancer risk and recurrence. Int J Cancer 2011;128:2335–43.
117. Tan J, Lee PL, Li Z, Jiang X, Lim YC, Hooi SC, et al. B55beta-associated PP2A complex controls PDK1-directed myc signaling and modulates rapamycin sensitivity in colorectal cancer. Cancer Cell 2010;18:459–71.
118. Okita K, Motohashi S, Shinnakasu R, Nagato K, Yamasaki K, Sato Y, et al. A set of genes associated with the interferon-gamma response of lung cancer patients undergoing alpha-galactosylceramide-pulsed dendritic cell therapy. Cancer Sci 2010;101:2333–40.
119. Sasaki T, Gan EC, Wakeham A, Kornbluth S, Mak TW, Okada H. HLA-B-associated transcript 3 (Bat3)/Scythe is essential for p300-mediated acetylation of p53. Genes Dev 2007;21:848–61.
120. Grover A, Izzo AA. BAT3 regulates Mycobacterium tuberculosis protein ESAT-6-mediated apoptosis of macrophages. PLoS One 2012;7:e40836.
121. Mu QF, Luo GH, Chen LJ, Wei J, Zheng L, Zhang XY, et al. [Apolipoprotein M expression in human colorectal cancer tissues and its clinicopathological relevance]. Zhonghua Wei Chang Wai Ke Za Zhi 2012;15:855–8.
122. Luo G, Zhang X, Mu Q, Chen L, Zheng L, Wei J, et al. Expression and localization of apolipoprotein M in human colorectal tissues. Lipids Health Dis 2010;9:102.
123. Levy P, Ripoche H, Laurendeau I, Lazar V, Ortonne N, Parfait B, et al. Microarray-based identification of tenascin C and tenascin XB, genes possibly involved in tumorigenesis associated with neurofibromatosis type 1. Clin Cancer Res 2007;13:398–407.
124. Tompkins JD, Wu X, Her C. MutS homologue hMSH5: role in cisplatin-induced DNA damage response. Mol Cancer 2012;11:10.
125. Robson T, James IF. The therapeutic and diagnostic potential of FKBPL; a novel anticancer protein. Drug Discov Today 2012;17:544–8.
126. McKeen HD, Brennan DJ, Hegarty S, Lanigan F, Jirstrom K, Byrne C, et al. The emerging role of FK506-binding proteins as cancer biomarkers: a focus on FKBPL. Biochem Soc Trans 2011;39:663–8.
127. Guo H, Deng Q, Wu C, Hu L, Wei S, Xu P, et al. Variations in HSPA1B at 6p21.3 are associated with lung cancer risk and prognosis in Chinese populations. Cancer Res 2011;71:7576–86.
128. Wang Y, Wei Y, Gaborieau V, Shi J, Han Y, Timofeeva MN, et al. Deciphering associations for lung cancer risk through imputation and analysis of 12,316 cases and 16,831 controls. Eur J Hum Genet 2015;23:1723–8.
129. Koshikawa K, Osada H, Kozaki K, Konishi H, Masuda A, Tatematsu Y, et al. Significant up-regulation of a novel gene, CLCP1, in a highly metastatic lung cancer subline as well as in lung cancers in vivo. Oncogene 2002;21:2822–8.
130. Nagai H, Sugito N, Matsubara H, Tatematsu Y, Hida T, Sekido Y, et al. CLCP1 interacts with semaphorin 4B and regulates motility of lung cancer cells. Oncogene 2007;26:4025–31.
131. Acquaviva J, Wong R, Charest A. The multifaceted roles of the receptor tyrosine kinase ROS in development and cancer. Biochim Biophys Acta 2009;1795:37–52.
132. Takeuchi K, Soda M, Togashi Y, Suzuki R, Sakata S, Hatano S, et al. RET, ROS1 and ALK fusions in lung cancer. Nat Med 2012;18:378–81.
133. Bergethon K, Shaw AT, Ou SH, Katayama R, Lovly CM, McDonald NT, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol 2012;30:863–70.
134. Fernandez-Cuesta L, Plenker D, Osada H, Sun R, Menon R, Leenders F, et al. CD74-NRG1 fusions in lung adenocarcinoma. Cancer Discov 2014;4:415–22.
135. Gil J, Peters G. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. Nat Rev Mol Cell Biol 2006;7:667–77.
136. Kotake Y, Nakagawa T, Kitagawa K, Suzuki S, Liu N, Kitagawa M, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. Oncogene 2011;30:1956–62.
137. Yu W, Gius D, Onyango P, Muldoon-Jacobs K, Karp J, Feinberg AP, et al. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. Nature 2008;451:202–6.
138. Enciso-Mora V, Broderick P, Ma Y, Jarrett RF, Hjalgrim H, Hemminki K, et al. A genome-wide association study of Hodgkin's lymphoma identifies new susceptibility loci at 2p16.1 (REL), 8q24.21 and 10p14 (GATA3). Nat Genet 2010;42:1126–30.
139. Maurice D, Hooper J, Lang G, Weston K. c-Myb regulates lineage choice in developing thymocytes via its target gene Gata3. EMBO J 2007;26:3629–40.
140. Tkocz D, Crawford NT, Buckley NE, Berry FB, Kennedy RD, Gorski JJ, et al. BRCA1 and GATA3 corepress FOXC1 to inhibit the pathogenesis of basal-like breast cancers. Oncogene 2012;31:3667–78.
141. Asselin-Labat ML, Sutherland KD, Vaillant F, Gyorki DE, Wu D, Holroyd S, et al. Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. Mol Cell Biol 2011;31:4609–22.
142. Wu Q, Wang H, Zhao X, Shi Y, Jin M, Wan B, et al. Identification of G-protein-coupled receptor 120 as a tumor-promoting receptor that induces angiogenesis and migration in human colorectal carcinoma. Oncogene 2013;32:5541–50.
143. Bose A, Guilherme A, Huang S, Hubbard AC, Lane CR, Soriano NA, et al. The v-SNARE Vti1a regulates insulin-stimulated glucose transport and Acrp30 secretion in 3T3-L1 adipocytes. J Biol Chem 2005;280:36946–51.
144. Petridou ET, Mitsiades N, Gialamas S, Angelopoulos M, Skalkidou A, Dessypris N, et al. Circulating adiponectin levels and expression of adiponectin receptors in relation to lung cancer: two case-control studies. Oncology 2007;73:261–9.
145. Bass AJ, Lawrence MS, Brace LE, Ramos AH, Drier Y, Cibulskis K, et al. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. Nat Genet 2011;43:964–8.
146. Mortensen UH, Lisby M, Rothstein R. Rad52. Curr Biol 2009;19:R676–7.
147. Danoy P, Michiels S, Dessen P, Pignat C, Boulet T, Monet M, et al. Variants in DNA double-strand break repair and DNA damage-response genes and susceptibility to lung and head and neck cancers. Int J Cancer 2008;123:457–63.
148. Benson FE, Baumann P, West SC. Synergistic actions of Rad51 and Rad52 in recombination and DNA repair. Nature 1998;391:401–4.
149. Feng Z, Scott SP, Bussen W, Sharma GG, Guo G, Pandita TK, et al. Rad52 inactivation is synthetically lethal with BRCA2 deficiency. Proc Natl Acad Sci USA 2011;108:686–91.
150. Treuner K, Helton R, Barlow C. Loss of Rad52 partially rescues tumorigenesis and T-cell maturation in Atm-deficient mice. Oncogene 2004;23:4655–61.
151. Lieberman R, Xiong D, James M, Han Y, Amos CI, Wang L, et al. Functional characterization of RAD52 as a lung cancer susceptibility gene in the 12p13.33 locus. Mol Carcinog 2016;55:953–63.
152. Su GH, Bansal R, Murphy KM, Montgomery E, Yeo CJ, Hruban RH, et al. ACVR1B (ALK4, activin receptor type 1B) gene mutations in pancreatic carcinoma. Proc Natl Acad Sci USA 2001;98:3254–7.
153. Burdette JE, Jeruss JS, Kurley SJ, Lee EJ, Woodruff TK. Activin A mediates growth inhibition and cell cycle arrest through Smads in human breast cancer cells. Cancer Res 2005;65:7968–75.
154. Modica S, Murzilli S, Salvatore L, Schmidt DR, Moschetta A. Nuclear bile acid receptor FXR protects against intestinal tumorigenesis. Cancer Res 2008;68:9589–94.
155. Park JY, Helm JF, Zheng W, Ly QP, Hodul PJ, Centeno BA, et al. Silencing of the candidate tumor suppressor gene solute carrier family 5 member 8 (SLC5A8) in human pancreatic cancer. Pancreas 2008;36:e32–9.
156. Rafnar T, Vermeulen SH, Sulem P, Thorleifsson G, Aben KK, Witjes JA, et al. European genome-wide association study identifies SLC14A1 as a

new urinary bladder cancer susceptibility gene. Hum Mol Genet 2011; 20:4268–81.
157. Eby MT, Jasmin A, Kumar A, Sharma K, Chaudhary PM. TAJ, a novel member of the tumor necrosis factor receptor family, activates the c-Jun N-terminal kinase pathway and mediates caspase-independent cell death. J Biol Chem 2000;275:15336–42.
158. McAllister KA, Bennett LM, Houle CD, Ward T, Malphurs J, Collins NK, et al. Cancer susceptibility of mice with a homozygous deletion in the COOH-terminal domain of the Brca2 gene. Cancer Res 2002;62:990–4.
159. De Cat B, David G. Developmental roles of the glypicans. Semin Cell Dev Biol 2001;12:117–25.
160. Filmus J. Glypicans in growth control and cancer. Glycobiology 2001;11: 19R–23R.
161. Yu CT, Hsia JY, Hseih YC, Su LJ, Lee TC, Ku CF, et al. The novel protein suppressed in lung cancer down-regulated in lung cancer tissues retards cell proliferation and inhibits the oncokinase Aurora-A. J Thorac Oncol 2011;6:988–97.
162. Liu Y, Liu P, Wen W, James MA, Wang Y, Bailey-Wilson JE, et al. Haplotype and cell proliferation analyses of candidate lung cancer susceptibility genes on chromosome 15q24-25.1. Cancer Res 2009;69:7844–50.
163. Ruthenburg AJ, Li H, Milne TA, Dewell S, McGinty RK, Yuen M, et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. Cell 2011;145:692–706.
164. Medina PP, Sanchez-Cespedes M. Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. Epigenetics 2008;3:64–8.
165. Coste B, Mathur J, Schmidt M, Earley TJ, Ranade S, Petrus MJ, et al. Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels. Science 2010;330:55–60.
166. Shimomura Y, Agalliu D, Vonica A, Luria V, Wajid M, Baumer A, et al. APCDD1 is a novel Wnt inhibitor mutated in hereditary hypotrichosis simplex. Nature 2010;464:1043–7.
167. Takahashi M, Fujita M, Furukawa Y, Hamamoto R, Shimokawa T, Miwa N, et al. Isolation of a novel human gene, APCDD1, as a direct target of the beta-Catenin/T-cell factor 4 complex with probable involvement in colorectal carcinogenesis. Cancer Res 2002;62:5651–6.
168. Tashiro K, Tsunematsu T, Okubo H, Ohta T, Sano E, Yamauchi E, et al. GAREM, a novel adaptor protein for growth factor receptor-bound protein 2, contributes to cellular transformation through the activation of extracellular signal-regulated kinase signaling. J Biol Chem 2009; 284:20206–14.
169. Park KH, Lo Han SG, Whang YM, Lee HJ, Yoo YD, Lee JW, et al. Single nucleotide polymorphisms of the TGFB1 gene and lung cancer risk in a Korean population. Cancer Genet Cytogenet 2006;169:39–44.
170. Liao Q, Zeng Z, Guo X, Li X, Wei F, Zhang W, et al. LPLUNC1 suppresses IL-6-induced nasopharyngeal carcinoma cell proliferation via inhibiting the Stat3 activation. Oncogene 2014;33:2098–109.
171. Lin R, White JH. The pleiotropic actions of vitamin D. BioEssays 2004; 26:21–8.
172. Peehl DM, Skowronski RJ, Leung GK, Wong ST, Stamey TA, Feldman D. Antiproliferative effects of 1,25-dihydroxyvitamin D3 on primary cultures of human prostatic cells. Cancer Res 1994;54:805–10.
173. Anderson MG, Nakane M, Ruan X, Kroeger PE, Wu-Wong JR. Expression of VDR and CYP24A1 mRNA in human tumors. Cancer Chemother Pharmacol 2006;57:234–40.
174. Chen G, Kim SH, King AN, Zhao L, Simpson RU, Christensen PJ, et al. CYP24A1 is an independent prognostic marker of survival in patients with lung adenocarcinoma. Clin Cancer Res 2011;17:817–26.
175. Wrensch M, Jenkins RB, Chang JS, Yeh RF, Xiao Y, Decker PA, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat Genet 2009;41:905–8.
176. Cybulski C, Masojc B, Oszutowska D, Jaworowska E, Grodzki T, Waloszczyk P, et al. Constitutional CHEK2 mutations are associated with a decreased risk of lung and laryngeal cancers. Carcinogenesis 2008;29: 762–5.
177. Chen Y, Deng J, Fujimoto J, Kadara H, Men T, Lotan D, et al. Gprc5a deletion enhances the transformed phenotype in normal and malignant lung epithelial cells by eliciting persistent Stat3 signaling induced by autocrine leukemia inhibitory factor. Cancer Res 2010; 70:8917–26.
178. Niwa H, Ogawa K, Shimosato D, Adachi K. A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature 2009;460:118–22.