Genetics of lung-cancer susceptibility

Paul Brennan, Pierre Hainaut, Paolo Boffetta

Lung cancer is the most common form of cancer death worldwide. Although reduction of tobacco consumption remains the most appropriate strategy to reduce lung-cancer burden, identification of genes involved in the cause of disease could contribute to further understanding of the underlying mechanisms, and eventually lead to additional prevention strategies and targeted treatments. Common gene variants involved in lung cancer have been recently identified through large, collaborative, genome-wide association studies. These studies identified three separate loci that are associated with lung cancer (5p15, 6p21, and 15q25) and include genes that regulate acetylcholine nicotinic receptors and telomerase production. However, much about genetic risk remains to be discovered, and rarer gene variants, such as those of the *CHEK2* gene, likely account for most of the remaining risk. There is also a need for studies that investigate how genetic susceptibility is associated with clinical outcome measures, including treatment response and tumour relapse.

Lancet Oncol 2011; 12: 399–408

Published Online
October 14, 2010
DOI:10.1016/S1470-2045(10)70126-1

International Agency for Research on Cancer (IARC), Lyon, France (P Brennan PhD, P Hainaut PhD, P Boffetta MD); **and The Tisch Cancer Institute, Mount Sinai School of Medicine, New York, NY, USA** (P Boffetta)

Correspondence to:
Dr Paul Brennan, Section of Genetics, International Agency for Research on Cancer, 150 cours Albert Thomas, 69372 Lyon Cedex 08, France
gep@iarc.fr

Epidemiology of lung cancer

Lung cancer was a rare disease until the beginning of the 20th century. Since then, occurrence has increased rapidly and this neoplasm has become the most common cancer in men in most countries, and is the main cause of cancer death worldwide. Lung cancer accounts for around 1095 000 new cancer cases and 951 000 deaths each year in men, and 514 000 cases and 427 000 deaths in women,[1] representing about 12·7% of all new cancer cases each year and 18·2% of cancer deaths. Survival from lung cancer has improved only moderately during the past decades and remains poor (around 10% at 5 years).[2] Although early-stage lung cancer can be treated surgically with good survival, most cases are diagnosed at a late stage when surgery is no longer an option. Late-stage lung cancers show poor response to radiotherapy and chemotherapy, although tyrosine kinase inhibitors were recently reported to be effective in reducing tumour burden in non-small-cell lung cancer (NSCLC) with EGFR mutations.[3] Spiral tomography can lead to identification of early pulmonary lesions in high-risk individuals, which can be successfully treated.[4] However, the clinical significance of these lesions is uncertain and it remains to be shown whether this approach leads to a decrease in lung-cancer mortality.[5]

The geographical and temporal patterns of lung-cancer incidence are largely determined by consumption of tobacco. An increase in tobacco consumption leads to an increase in the incidence of lung cancer a few decades later, first evident among young adults. Similarly, a decrease in consumption is followed by a decrease in incidence.

In men, reported lung-cancer incidence is highest in central and eastern Europe (>60 per 100 000) and lowest in Africa and western and southern Asia (<15 per 100 000; figure 1A).[1,6] In women, incidence is high (>20 per 100 000) in the USA, Canada, Denmark, and the UK, and low (around five per 100 000) in countries such as Spain, where the prevalence of smoking in women only recently increased. The lowest incidence (<two per 100 000) is recorded in Africa and India, where detection and registration of lung cancer might be incomplete, particularly in women (figure 1B). Within countries, differences in incidence according to ethnicity are often observed; for example, black men in the USA have high rates of lung cancer. Incidence can differ by geographical location, as in the case of high rates reported among women from several regions of China (35 per 100 000 in Harbin), despite a low prevalence of smoking. In Harbin, high incidence has been associated with emission of coal combustion rich in polycyclic aromatic hydrocarbons, generated by cooking practices in unvented homes.[7]

Lung cancer is declining among men in most high-income countries, and among women in a few countries, such as the UK,[8] mainly because of decreasing patterns of tobacco consumption. Conversely, rates are increasing among men in most low-and middle-income countries, notably China, and among women.[9]

The main histological categories of lung cancer are NSCLC, originating from bronchial epithelial-cell precursors, and small-cell lung carcinoma (SCLC), originating from a neuroendocrine-cell precursor. NSCLC includes three main types—squamous-cell carcinoma, adenocarcinoma, and large-cell carcinoma. The first two types represent about 80% of all lung cancers worldwide.[10] Squamous-cell carcinomas, previously the most common type of NSCLC, have decreased in the past decades and there has been an increase in adenocarcinomas in both sexes.[11,12] Apart from minor differences, the main risk factors for lung cancer affect all histological types.[11] At the molecular level, the most common somatic genetic or epigenetic alterations in lung cancers are *TP53* mutations (which are common in all forms of lung cancers), *KRAS* or *EGFR* mutations (mainly in adenocarcinoma), and alterations in the CDKN2A/ARF/RB1 pathway.[13]

A carcinogenic effect of tobacco smoke on the lung was shown in the 1950s and has been recognised by public health and regulatory authorities since the mid-1960s.[14] The risk of lung cancer is roughly ten-times higher among smokers than never smokers.[14] This overall risk reflects the contribution of different aspects of tobacco smoking: average consumption, duration of smoking, time since quitting, age at start, type of tobacco product,

and inhalation pattern, with duration being the dominant factor.[14] Compared with continuous smokers, the excess risk decreases in ex-smokers after quitting, although a small, lifetime excess risk is likely to persist for long-term quitters.[14] Cumulative risk can be measured by combining relative-risk estimates from case–control studies and national lung-cancer mortality rates. For continuous smokers, the cumulative risk of lung cancer by

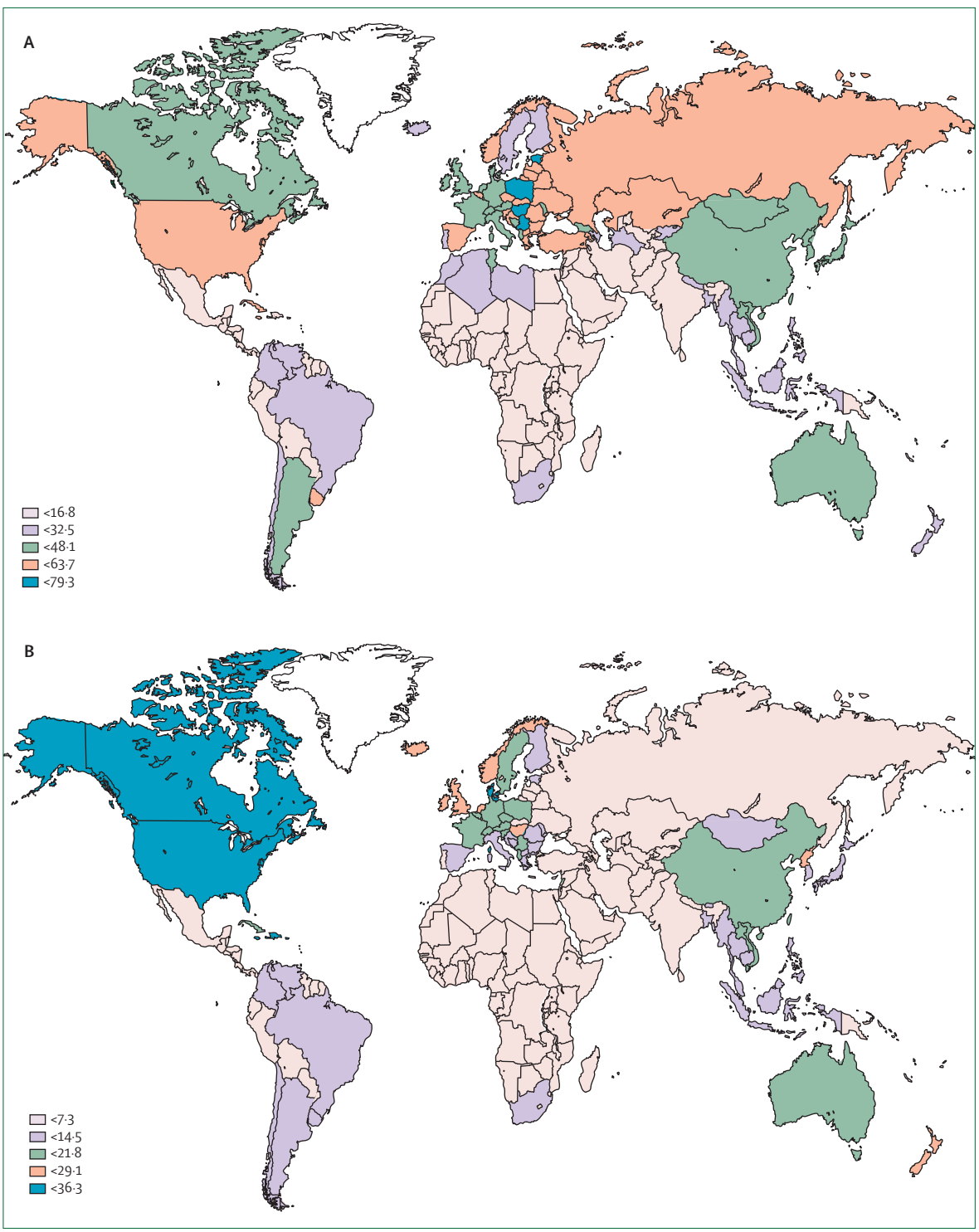

Figure 1: **Lung-cancer incidence (per 100 000) by country, 2008**[1]
Lung-cancer incidence in men (A) and women (B).

age 75 years is about 15%, versus about 1% in never smokers.[15,16] Using lung-cancer case-control data and mortality data in 1999 as an example, cumulative risk was 19% in current smokers, 5% in ex-smokers, and 1% in never smokers (figure 2).[17]

A specific pattern of mutations in *TP53* and *KRAS* (excess of G→T transversions) has been identified in lung cancers of heavy smokers. These mutations often occur at bases that are binding sites for adducts of polycyclic aromatic hydrocarbons.[18] The proportion of lung cancers attributable to tobacco smoking in a population depends on smoking habits in the previous decades and on background rates of lung cancer caused by other risk factors, such as chronic pulmonary infections and occupational exposures. Among men, the proportion of lung cancers caused by tobacco smoking is typically higher than 90% (see table),[19,20] although it is lower in China. Among women, there is greater intercountry variability because of the maturation of the tobacco-smoking epidemic. Large cohort studies show that, for comparable smoking histories, men and women have similar lung-cancer rates, suggesting that they are equally susceptible.[21]

Rates of lung cancer among non-smokers and the geographical and temporal patterns are poorly known, because of the limited number of large-scale studies with detailed information on smoking.[22] Lung cancer in never smokers is more common in women than in men, and the proportion of cases in women varies substantially from region to region. For example, the proportion of women with lung cancer who are never smokers varies from 83% in south Asia to 15% in the USA.[23] It is unclear, however, whether these patterns reflect anything other than the proportion of never smokers among men and women in different populations.

Exposure to involuntary smoking is associated with a risk of lung cancer in non-smokers.[24] The magnitude of the excess risk among non-smokers exposed to involuntary smoking is around 20%.[14] Several other environmental causes of lung cancer have been identified, including ionising radiation (radon decay products and x-rays), asbestos, other agents that mainly occur as occupational exposures, pulmonary diseases including tuberculosis and chronic bronchitis, and outdoor air pollution.[25] However, the effect of these factors on lung cancer is known only in smokers, with the exception of indoor air pollution from heating and cooking, which is an important cause of lung cancer among non-smoking women in China and other Asian countries.[26] A role for dietary factors has been suggested, notably low intake of fresh fruits and vegetables, but no conclusions can be made with the available evidence.[27]

Identifying genes for lung cancer

There are three reasons why the identification of susceptibility variants for lung cancer might be important for prevention and control. First, knowledge of these genes might help to elucidate the underlying process of disease onset or survival, and provide clues for potential lung-cancer therapeutics or chemoprevention drugs. For example, a possible association between polymorphisms in genes involved in the EGFR pathway shows how identification of susceptibility genes could inform therapeutic strategies.[28,29] Second, lung-cancer susceptibility genes might act as unconfounded markers of non-genetic exposures. There is much interest in a protective association for vitamin B6 and lung cancer, although identifying specific causal exposures is difficult because of the strong correlation of nutritional and lifestyle factors with lung-cancer risk.[30] Gene variants that affect plasma concentrations of vitamin B6 have been recently identified.[31] The effects of such variants would not be

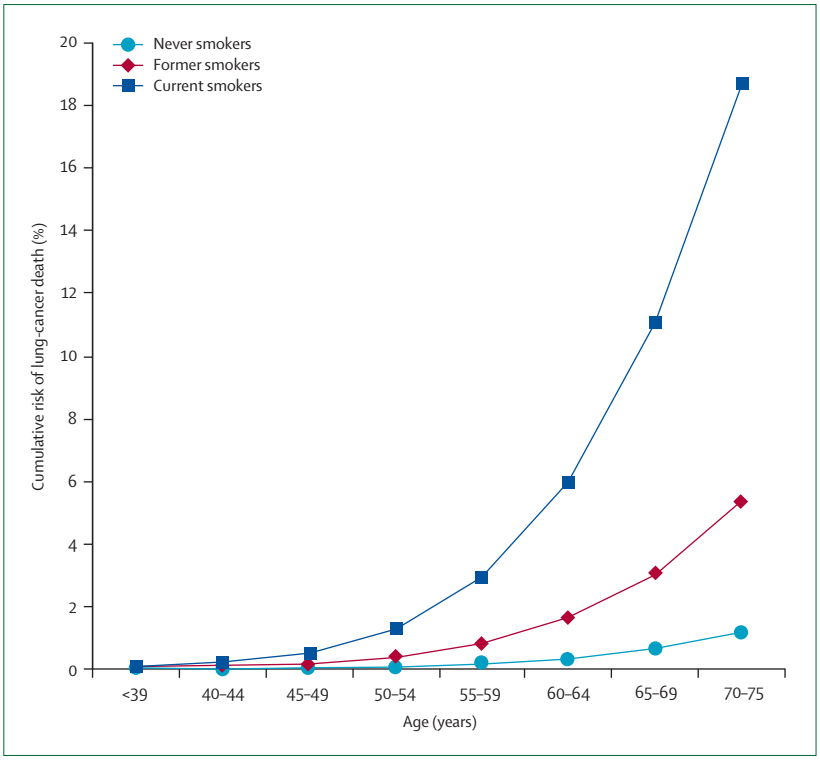

Figure 2: **Cumulative risk of lung-cancer death by age in Polish men, stratified by smoking status**
Risk estimates are derived from a model adjusted by age and centre. Reference mortality rates are from Globocan 2002, Poland. Never smokers are participants who smoked fewer than 100 cigarettes in a lifetime; former smokers are those who stopped smoking more than 1 year before the interview. Mean age at stoppage of smoking was 54·8 years in cases and 45·4 years in controls.

	Men (%)	Women (%)
EU15[19]	91	64
EU10[19]	94	67
USA[19]	91	86
China[20]	75	18

EU15=15 pre-2004 European Union member states. EU10=ten post-2004 European Union members.

Table: **Proportion of lung cancers attributable to tobacco smoking in selected countries**

restricted by confounding of lifestyle factors, and an association with lung cancer would provide strong evidence of a direct protective effect of vitamin B6. Analysis of genes in this manner is known as Mendelian randomisation, and is likened to a randomised controlled trial for the environmental factor.[32] Finally, identification of susceptibility genes could be helpful for personalised prediction, risk estimation, and individualised therapy for lung cancer. In view of the over-riding effect of non-genetic factors for lung cancer, the usefulness of risk estimation by incorporating gene variants is far from clear.

Evidence of a genetic component

Familial clustering is evidence of a genetic component for a particular cancer, where λ is the ratio of observed incidence of a cancer among first-degree relatives to expected frequency.[33] Extensive information on familial clustering for specific cancers is available from linkage of the Utah family database with the Utah cancer registry, which included around 125 000 individuals between 1952 and 1992,[34] and from the Swedish cancer registry, which included more than 1 200 000 cancer cases between 1958 and 1997.[35] The Swedish analysis showed that for almost 89 000 lung-cancer cases, risk of lung cancer among first-degree relatives was increased 1·9 times (95% CI 1·6–2·4). This familial risk was similar to that of other common cancers with a recognised genetic component, including breast (λ=1·5), colon (λ=1·9), and prostate cancer (λ=2·7). The familial risk was higher among siblings (λ=3·1) than among offspring (λ=1·6), which is consistent with an underlying recessive effect.[33,36] The Utah linkage reported a λ of 2·55 for lung cancer among first-degree relatives, which was not modified by age at onset.[34]

A third resource for familial clustering is the Icelandic cancer registry. Although small in size, this registry is of much interest because it allows identification of familial risk among distant relatives. On the basis of 2756 individuals with lung cancer, the familial risk was high for both parents (λ=2·7) and siblings (λ=2·0), and lower among second-degree (λ=1·4), third-degree (λ=1·2), fourth-degree (λ= 1·1), and fifth-degree (λ=1·0) relatives.[37,38]

Increased familial risk of lung cancer is a necessary indication of a genetic contribution, although not a sufficient one, since it could be influenced by familial clustering of non-genetic risk factors, particularly cigarette smoking. Bermejo and Hemminki[39] estimated that the familial risk of lung cancer among offspring would be expected to increase by about 20% with increased familial tendency to smoke. However, this increase does not fully account for the observed increase in familial risk, suggesting that other factors, including genetic susceptibility, are the main contributors to this increased risk.[39]

Genes that affect lung-cancer risk

Genes that might affect lung-cancer risk fall into three categories: rare high-risk variants (risk of 10 or higher and prevalence of 1% or less), moderate-risk variants (risk of around 2–5 and prevalence of not more than 5%), and common low-risk variants (risk of between 1·1–1·5 and prevalence of more than 5%). Other types of gene variants are unlikely to exist because of evolutionary pressures (eg, common high-risk variants), or are undetectable by current study designs (eg, rare low-risk variants). For each category, it is possible for several individual variants to exist within a locus. High-risk gene mutations that show some association with lung cancer include *TP53* mutations in families with Li-Fraumeni syndrome. Although the frequency of lung cancer in carriers of *TP53* mutation is only slightly higher than that of the general population, the age of cancer occurrence is typically lower.[40,41]

Linkage studies to detect high-risk susceptibility genes must include families with several members who have lung cancer. The Genetic Epidemiology of Lung Cancer Consortium (GELCC) reported on 52 families with three or more cases of lung or larynx cancer, and identified a potential lung-cancer susceptibility locus in chromosome region 6q23–25.[42] Fine-mapping of this region led to identification of *RGS17* as a possible susceptibility gene;[43,44] however, the precise functional effect of the gene in lung cancer is unknown, and confirmation studies in families with several cases of lung cancer are needed.

Most genetic risk is likely to involve several genes of moderate and low risk. Such risk variants have, until recently, been tested on a candidate gene basis. Although many preliminary studies have reported an association between specific gene variants and lung cancer, most associations have not been robustly replicated in large studies or combined meta-analyses. Two possible exceptions exist. The glutathione S transferase M1 (*GSTM1*) genotype is one of the most widely analysed genetic variants for lung cancer, with a hypothesised increased risk among null carriers because of their decreased ability to detoxify environmental carcinogens.[45,46] A meta-analysis of 43 studies comprising more than 7000 cases and 10 000 controls found an increased risk of 1·17 (95% CI 1·07–1·27), with no evidence of publication bias.[47] *GSTM1* is associated with bladder cancer, another cancer strongly associated with tobacco carcinogens,[48] adding to its plausibility as a lung-cancer susceptibility variant, although confirmation in additional large studies would be helpful.

A second variant with convincing evidence of an association with lung cancer is the I157T variant of *CHEK2*,[49,50] a key cell-cycle control gene that activates cell-cycle checkpoints in response to DNA damage. Thus, *CHEK2* has a main role in maintenance of genetic integrity. The I157T variant increases risk of several cancers, notably colon, breast, and prostate,[51,52] and is present at a relatively high frequency in populations from northern and central Europe (5–7%). In a multicentre study from central Europe, carriers of the rare variant had a significantly lower incidence of lung cancer than individuals with the common homozygous genotype

(OR 0·44, p<0·00001).[49] This unexpected result was replicated in an independent study from Poland (OR 0·3, p<0·0000001).[50] In the latter study, other rare *CHEK2* mutations were also found to reduce the risk of lung cancer, implying that impaired function of *CHEK2* might, in some circumstances, increase risk of breast and other cancers yet be protective for lung cancer. The functional basis of this effect is unclear. One hypothesis is that dysfunctional *CHEK2* might hinder the capacity of cells with damaged DNA to undergo cell-cycle arrest and DNA repair at defined checkpoints, resulting in increased mitotic failure and elimination of cells containing heavily damaged DNA. Understanding of this association would be important for development of CHEK2-based therapeutics.[53]

Genome-wide association studies

An alternative to candidate gene studies is the recent development of genome-wide association (GWA) studies, which aim to cover most genetic variation by genotyping up to 1000000 genetic variants (or single-nucleotide polymorphisms [SNPs]), and do not require prior knowledge of the functional significance of the variants studied. GWA studies do require very large sample sizes and provide only partial coverage of common variants (around 70–90% for European populations based on recent chip panels), and limited coverage of rarer variants.[54] The power of this approach for identifying new susceptibility genes for several chronic diseases has become apparent over the past 3 years. By the end of 2009, GWA studies identified more than 80 common variants independently associated with different cancer sites.[55]

In April, 2008, three separate genome-wide studies of lung cancer were published.[56–58] These included a study by the International Agency for Research on Cancer (IARC) of genome-wide data on 1989 lung-cancer cases and 2625 controls, a study by MD Anderson (Houston, TX, USA) including 1154 cases and 1137 controls, and a study from DeCode, Iceland, including 665 cases and more than 10000 controls. All three studies provided strong evidence of a susceptibility region in 15q25.1, with a very consistent measure of effect between the studies: OR 1·30, $p=10^{-20}$ for the IARC study; OR 1·31, $p=10^{-8}$ for the DeCode study; OR 1·32, $p=10^{-17}$ for the MD Anderson study. This locus was identified by all three studies because the effect size is large and the minor allele is common. A SNP in chromosome 6p21 affecting lung-cancer risk was also reported in the IARC study.[56] Evidence for an association with this variant, situated in the *HLA* gene region, has been supported in additional studies,[59] although the association is not entirely reproducible.[60] A third gene region for lung cancer has been subsequently confirmed by bringing together larger numbers of cases and controls.[59,61] The IARC study was expanded to 3259 cases of lung cancer and 4159 controls with genome-wide data, and reported an additional susceptibility locus in chromosome 5p, in a region that includes *TERT* (human telomerase reverse transcriptase gene), and *CLPTM1L* (cleft lip and palate transmembrane-1-like gene).[61] This region was also identified in a meta-analysis combining overall odds ratios from the MD Anderson and IARC studies and an additional UK study for over 5000 cases and 5000 controls.[59]

GWA studies of lung cancer and tobacco smoking

Two further large and partly overlapping GWA studies of lung cancer have been recently reported.[60,62] An analysis of more than 21000 variants based on genome-wide results of more than 7500 cases and 8200 controls did not find any additional susceptibility loci for lung cancer.[62] Results from a recent GWA study of almost 6000 case–control pairs were pooled with another ten studies, resulting in an analysis of 13300 primary lung-cancer cases and 19666 controls.[60] No additional susceptibility loci were observed, suggesting that in populations of European descent, no further common risk variants that increase risk by at least 20% or more remain to be detected. However, the investigators did detect strong heterogeneity by histology of the 5p15 locus, with the effect largely restricted to adenocarcinomas.

A large genome-wide study of lung cancer among never-smokers provided preliminary evidence for a susceptibility locus in region 13q31.3, with additional gene-expression data suggesting possible involvement of the *GPC5* gene.[63] The association was not statistically significant genome wide, and additional data to confirm this effect are needed. With the high rate of lung cancer among never smokers in some parts of the world, further elucidation of genetic susceptibility among this group will be of much interest.

Association of the 15q25 locus with lung cancer

The 15q25 susceptibility region contains six identified coding regions, including three cholinergic nicotine-receptor genes (*CHRNA3*, *CHRNA5*, and *CHRNB4*), encoding nicotinic acetylcholine receptors (nAChRs) in neuronal and other tissues. Two of the GWA studies identified variants directly via their association with lung-cancer risk,[56,57] whereas the third study identified an association with the same genetic region and smoking quantity, and concluded that the variant increases lung-cancer risk indirectly through smoking.[58] Since nAChRs mediate sensitivity to nicotine, it has been proposed that variant receptors might increase addiction to tobacco and, therefore, exposure to tobacco carcinogens. In addition to lung cancer, the variant allele is associated with peripheral arterial disease and chronic obstructive pulmonary disease,[64] and with head and neck cancers.[65] However, although an association between 15q25 variants and smoking intensity has been repeatedly noted, this alone does not account for the strength of the association with lung cancer (around an 80% increase in lung-cancer risk for individuals who inherit two risk alleles). Evidence

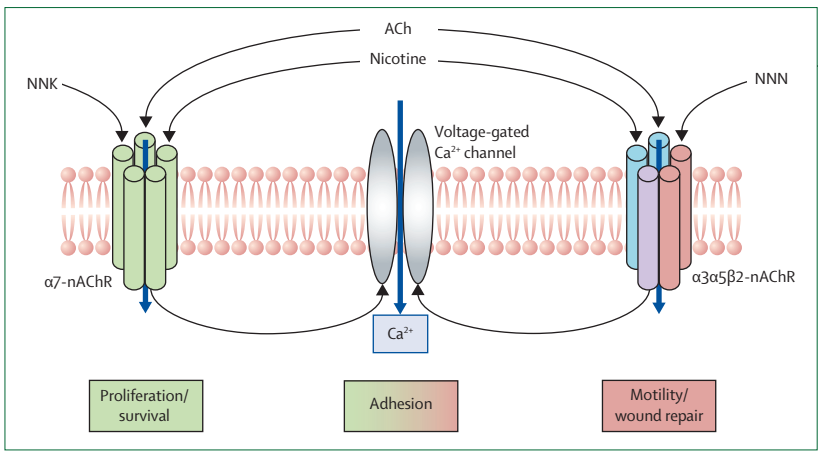

Figure 3: **Nicotine cholinergic signalling system in bronchial mucosal cells**
The two main categories of nAChRs are shown in green (α7-nAChR) or pink (α3α5β3-nAChR). The same colours show the participation of these receptors in downstream effects. nAChR=nicotinic acetylcholine receptor. NNK=4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. ACh=acetylcholine. NNN=N-nitrosonornicotine. Ca=calcium.

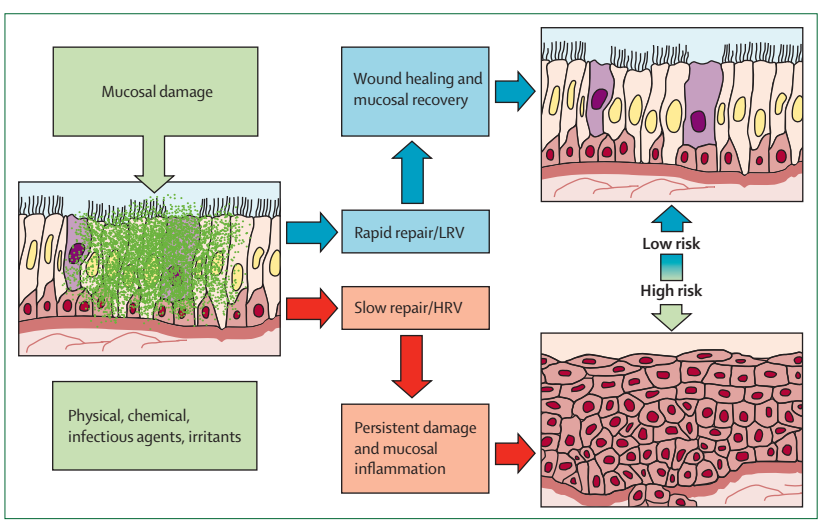

Figure 4: **Role of *CHRNA* variants as determinants of rapid or slow repair of the bronchial mucosa after damage**
LRV=low-risk variants. HRV=high-risk variants.

that this risk region is relevant for lung cancer independent of smoking intensity comes from studies of never smokers, although results from European and US studies are not consistent, with some studies reporting an excess risk for never smokers and others detecting no increase in risk.[56,57] These differences might be due, at least in part, to misclassification of smoking status.[65] Studies in Asian populations might be more relevant, since tobacco surveys have consistently reported that very few older women have ever smoked tobacco, and misclassification of smoking status among women is unlikely.[24] Wu and colleagues[66] recently reported that four common 15q25 variants were associated with lung cancer, in a large series of 3500 lung-cancer cases and 3300 controls in two Chinese case–control studies.[66] The effect of all four SNPs was similar among never and ever smokers, and among men and women. Given the low linkage disequilibrium between these variants and those reported in European and US populations, these results raise the possibility that the region might contain several distinct lung-cancer susceptibility loci. A Japanese study of the same 15q25 variants that increase risk in European populations found that these variants were associated with lung-cancer risk to a similar extent among never and ever smokers, although results were not reported separately for women.[67]

Biological basis of susceptibility with CHRNA genes

nAChR are ubiquitous cell-surface receptors composed of five subunits, each a transmembrane protein with four membrane-spanning domains (figure 3). Nicotine mimics the effect of acetylcholine by binding to a subunit and inducing a conformational change on the internal side of the membrane that allows the flow of cations (Ca^{2+}, Na^{+}, K^{+}) into the cytoplasm. This Ca^{2+} influx opens the gates of voltage-activated calcium channels, with many effects on the activation of calcium-dependent signalling pathways.[68,69] In bronchial epithelial cells, the main forms of nAChR are α7 homopentamers and α3α5β2 heteropentamers. These two receptor types show different distributions within the epithelium. In wounded epithelium, α7-nAChR is primarily expressed at lateral edges of differentiated, non-migrating cells, whereas α3α5β2-nAChr is highly expressed in migrating cells at the wound edge. Moreover, low levels of the latter receptor are also detected at the surface of basal cells, the compartment providing progenitors for normal differentiation and for flat cells that migrate during wound repair. These results suggest that α3α5β2-nAChR have a role in the repair of wounded bronchial epithelium.[70]

Alterations in nAChR signalling might contribute to lung cancer.[69] Gene-expression studies show that NSCLC in non-smokers exhibits higher levels of α3-nAChR and α6-nAChR than in smokers, and have identified a 65-gene expression signature associated with the α3/α6-nAChR expression pattern.[71] In a recent study, we found that *CHRNA3*, but not *CHRNA5*, is systematically hypermethylated and down-regulated in lung cancers. Forced expression of *CHRNA3* in lung-cancer cells resulted in enhanced apoptosis, suggesting that hypermethylation provides cancer cells with a survival advantage.[72,73]

Moreover, the tobacco-specific nitrosamines 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and N-nitrosonornicotine (NNN) bind to distinct nAChRs with affinities higher than for nicotine itself, suggesting that these receptors might enhance the targeting of bronchial cells by tobacco carcinogens.[69]

Figure 4 summarises current hypotheses on how *CHRNA* polymorphisms predispose to lung cancer independent of addiction properties. First, these variants might modulate cell migration and wound repair in bronchial mucosa injured by inhaled toxic substances. Polymorphisms associated with low cell motility might lead to more persistent mucosal damage and

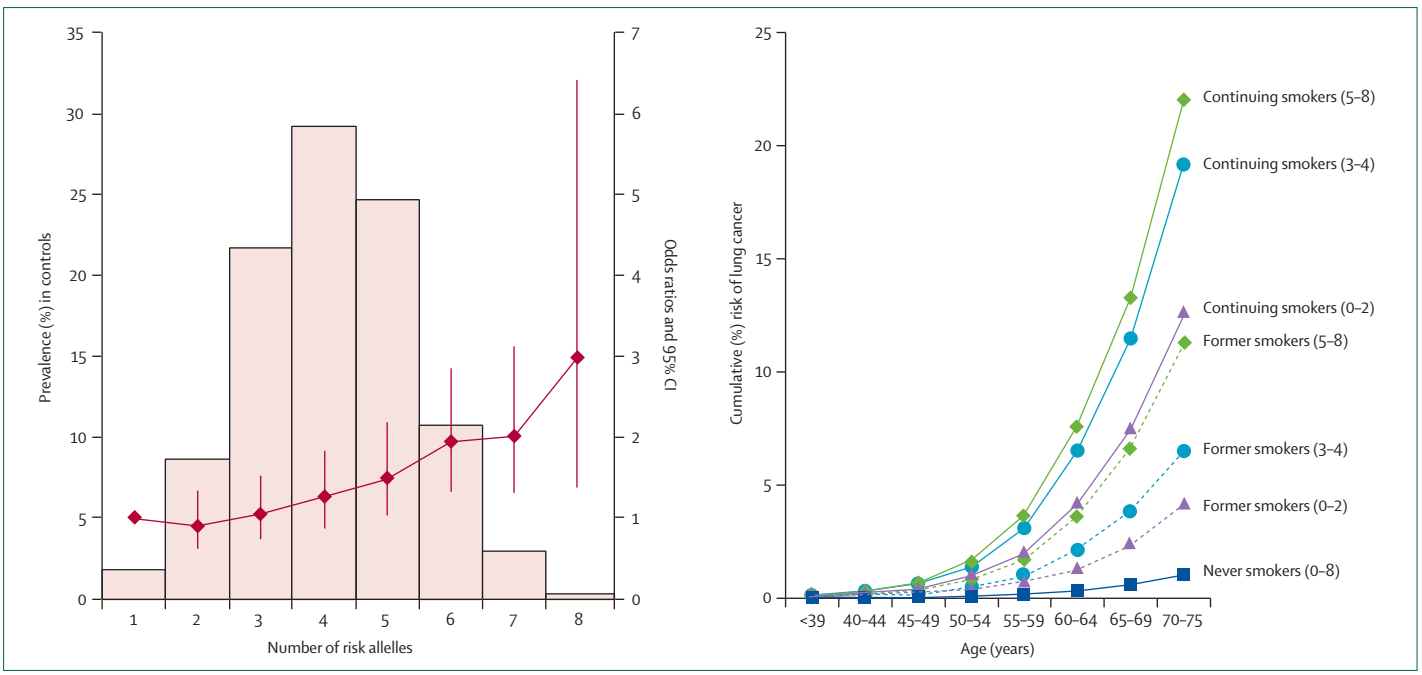

Figure 5: **Lung-cancer risk and cumulative risk of lung-cancer death stratified by number of risk alleles in Polish men**
(A) Left axis and histogram show prevalence of lung cancer in controls. Right axis, diamonds, and vertical bars show risk of lung cancer. (B) Cumulative risk of lung-cancer death by age, stratified by smoking status and number of risk alleles. Reference mortality rates are from Globocan 2002. Risk estimates are derived from a model adjusted by age and centre. Mean age at stoppage of smoking for the (0–2), (3–4), and (5–8) allele categories are 54·3, 54·7, and 56·3 years, respectively, in cases and 45·5, 43·2, and 43·4 years in controls.

inflammation, which could render mucosal cells more sensitive to mutagens and to the formation of precursor lesions, such as squamous metaplasia or glandular hyperplasia. Second, *CHRNA* polymorphisms might modulate cancer-cell invasion and metastasis. How *CHRNA* variants affect these processes remains unclear and the mechanistic basis of these effects at the receptor level remain to be elucidated. Of the SNPs strongly associated with susceptibility, rs16969968 is non-silent and specifies an amino-acid change from aspartic acid to asparagine at codon 398, in a loop of the α5 subunit which is directly involved in the closure of Ca^{2+} channels. However, it is still not known whether the two polymorphic variant proteins differ in their functional properties.

Lung cancer and 5p15 locus

The susceptibility locus at 5p15.33 contains two biologically relevant genes for lung cancer, *TERT* and *CLPTM1L*. Two variants (rs402710 and rs2736100) that are not strongly associated with each other were both reported to be associated with lung-cancer risk.[59,61] This association was further clarified by an international coordinated analysis comprising genotype data from an additional 11 645 lung-cancer cases and 14 954 controls, of whom 85% were white and 15% were Asian.[74] This study confirmed a significant association with lung-cancer risk in white people for both rs2736100 (OR 1·15, $p=1\times10^{-10}$) and rs402710 (OR 1·14, $p=5\times10^{-8}$), and identified a similar effect in Asians (rs2736100: OR 1·23, $p=2\times10^{-5}$; rs402710: OR 1·15, p=0·007). The study also reported a histology-specific role of rs2736100 and rs402710 in adenocarcinoma in both ethnic groups, similar to the recent large meta-analysis of genome-wide studies for lung cancer.

Current knowledge of the functions of *TERT* and *CLPTM1L* implicate *TERT* as the more plausible lung-cancer gene candidate. *TERT* is the reverse transcriptase component of telomerase that is essential for telomerase enzymatic activity and maintenance of telomeres.[75] Telomerase is responsible for telomere regeneration and up to 90% of human tumour samples (including lung cancer) show telomerase activity, suggesting that regeneration of telomeres is a vital step for most forms of carcinogenesis.[76] *TERT* expression is high in stem cells and progenitors but very low in most types of normal cells. Telomerase inhibitors are of much interest for potential chemoprevention and treatment of cancer.[77] DNA sequencing has revealed that there is little common genetic variation in the *TERT* coding region, which along with high conservation between species, implies that the gene itself is under strong evolutionary restraint. The presence of two independent associations for lung cancer in this region suggests that additional susceptibility variants might be present.

The functions of *CLPTM1L* are poorly understood and a possible role in cancer is a matter of speculation. The gene derives its name from its homology with *CLPTM1*, which is located on chromosome 19p and segregated with a phenotype of cleft lip and palate in one family.[78]

CLPTM1L has been cloned as a cisplatin-resistance factor in ovarian cancer-cell lines, under the name *CRR9* (cisplatin-resistance related protein 9).[79] Finally, there is also evidence that the 5p15.33 region might be a susceptibility locus for several cancer types, including pancreatic, urinary bladder, and cervical cancer.[80,81]

Identification of susceptible subgroups

Taking the three susceptibility loci for lung cancer identified through GWA studies (5p15, 6p21, and 15q25), and including both 5p15 variants, an individual can potentially inherit up to eight risk variants. This is a conservative estimate since, as discussed, each region might contain more than one susceptibility gene. The risk profile, based on the population from the IARC study of lung-cancer in central Europe shows that, as expected, most of the population will have between one and five risk variants, with risk increasing by about two-times between these groups (figure 5A). If the risk profile is stratified according to the bottom third of the population (zero to two risk variants), middle half (three to four risk variants), and the remainder (more than four risk variants), and calculated by smoking status (never, former, or current), there is clear evidence that smoking status is the over-riding predictor of lung-cancer risk (figure 5B).

Future GWA studies of lung cancer and tobacco smoking

Less than 10% of the familial risk of lung cancer is explained by the three known susceptibility loci at 15q25, 5p14, and 6p21. Assuming familial risk is largely driven by shared genetic factors, this suggests that the genetic architecture of lung cancer might include several rare variants, perhaps clustered within susceptibility loci, such as that for 5p15. Another possibility is that inherited structural variation, including copy-number variation, small deletions, and insertions, account for much of the missing variability. It is likely that these questions will be answered by larger studies, including more comprehensive panels of SNPs that allow rare-variant detection and assessment of structural variation, as well as parallel studies in several ethnic groups.

Implications for treatment and survival

Survival after a diagnosis of lung cancer is poor, largely because of the high metastatic potential of the tumour and limited response to treatment. The potential for genes to affect disease outcome is apparent from family studies of parent-child or sibling-sibling pairs who both have lung cancer.[82] In both situations, having a parent or sibling with a good survival from lung cancer was a significant positive indicator of outcome. Having a spouse with good survival from lung cancer was not associated with survival. Although specific tumour characteristics are known to be associated with treatment response and survival, most notably specific *EGFR* mutations, comprehensive evidence for an association between germline variation and outcome is not yet available. The need for large, comprehensive genetic studies within the context of therapeutic clinical trials is clearly a priority. Moreover, elucidation of the functional effect of *CHRNA* variants is also an area for further development, since these receptors are potential targets for specific agonist or antagonist drugs.

Search strategy and selection criteria

References for this review were identified through searches of Medline and the reference lists of relevant articles up to March, 2010. Main search terms used were "lung cancer" and "genetic susceptibility". Papers selected included the largest-scale and highest-quality evidence for each topic discussed. Only references published in English were included.

Contributors

All authors contributed to the initial draft and redrafting of the report, and accept equal responsibility for its content.

Conflicts of interest

The authors declared no conflicts of interest.

References

1 Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. GLOBOCAN 2008: cancer incidence and mortality worldwide: IARC Cancer Base, no 10. Lyon: International Agency for Research on Cancer, 2010.

2 Sant M, Allemani C, Santaquilani M, et al. EUROCARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. *Eur J Cancer* 2009; **45:** 931–91.

3 Linardou H, Dahabreh IJ, Bafaloukos D, Kosmidis P, Murray S. Somatic *EGFR* mutations and efficacy of tyrosine kinase inhibitors in NSCLC. *Nat Rev Clin Oncol* 2009; **6:** 352–66.

4 Henschke CI, Yankelevitz DF. CT screening for lung cancer: update 2007. *Oncologist* 2008; **13:** 65–78.

5 Bach PB, Jett JR, Pastorino U, Tockman MS, Swensen SJ, Begg CB. Computed tomography screening and lung cancer outcomes. *JAMA* 2007; **297:** 953–61.

6 Curado MP, Edwards B, Shin HR, et al, eds. Cancer incidence in five continents, vol IX. IARC scientific publications, no 160. Lyon: International Agency for Research on Cancer, 2007.

7 De Marini DM, Landi S, Tian D, et al. Lung Tumor KRAS and *TP53* mutations in nonsmokers reflect exposure to PAH-rich coal combustion emissions. *Cancer Res* 2001; **61:** 6679–81.

8 Bray F, Tyczynski JE, Parkin DM. Going up or coming down? The changing phases of the lung cancer epidemic from 1967 to 1999 in the 15 European Union countries. *Eur J Cancer* 2004; **40:** 96–125.

9 Parkin MD, Bray F, Ferlay J, Pisani P. Global Cancer Statistics, 2002. *CA Cancer J Clin* 2005; **55:** 74–108.

10 Devesa SS, Bray F, Vizcaino P, Parkin DM. International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. *Int J Cancer* 2005; **117:** 294–99.

11 Travis WD, Lubin J, Ries L, Devesa S. United States lung carcinoma incidence trends: declining for most histologic types among males, increasing among females. *Cancer* 1996; **77:** 2464–70.

12 Tyczynski JE, Bray F, Parkin DM. Lung cancer in Europe in 2000: epidemiology, prevention, and early detection. *Lancet Oncol* 2003; **4:** 45–55.

13 Gazdar A, Franklin WA, Brambilla E, Hainaut P, Yokota J, Harris CC. Genetic and molecular alterations. In: Travis WD, Brambilla E, Müller-Hermelink HK, Harris CC, eds. World Health Organization classification of tumours. Pathology and genetics of tumours of the lung, pleura, thymus and heart. Lyon: International Agency for Research on Cancer, 2004: 21–23.

14 IARC monographs on the evaluation of carcinogenic risks to humans, vol 83. Tobacco smoke and involuntary smoking. Lyon, France: International Agency for Research on Cancer, 2004.

15 Peto R, Darby S, Deo H, Silcocks P, Whitley E, Doll R. Smoking, smoking cessation, and lung cancer in the UK since 1950: combination of national statistics with two case-control studies. *BMJ* 2000; **321:** 323–29.

16 Crispo A, Brennan P, Jöckel KH, et al. The cumulative risk of lung cancer among current, ex- and never-smokers in European men. *Br J Cancer* 2004; **91:** 1280–86.

17 Brennan P, Crispo A, Zaridze D, et al. High cumulative risk of lung cancer death among smokers and nonsmokers in central and eastern Europe. *Am J Epidemiol* 2006; **164:** 1233–41.

18 Pfeiffer GP, Denissenko MF, Olivier M, Tretyakova N, Hecht SS, Hainaut P. Tobacco smoke carcinogens, DNA damage and *p53* mutations in smoking-associated cancers. *Oncogene* 2002; **21:** 7435–51.

19 Peto R, Lopez AD, Boreham J, Thun M. Mortality from smoking in developed countries 1950-2000, 2nd edn. Oxford: Oxford University Press, 2006.

20 Wang JB, Jiang Y, Wei WQ, Yang GH, Qiao YL, Boffetta P. Estimation of cancer incidence and mortality attributable to smoking in China. *Cancer Causes Control* 2010; **21:** 959–65.

21 Thun MJ, Henley SJ, Burns D, Jemal A, Shanks TG, Calle EE. Lung cancer death rates in lifelong nonsmokers. *J Natl Cancer Inst* 2006; **98**: 691–99.

22 Thun MJ, Hannan LM, Adams-Campbell LL, et al. Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies. *PLoS Med* 2008; **5:** 185.

23 Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers—a different disease. *Nat Rev Cancer* 2007; **7:** 778–90.

24 Boffetta P, Agudo A, Ahrens W, et al. Multicenter case-control study of exposure to environmental tobacco smoke and lung cancer in Europe. *J Natl Cancer Inst* 1998; **90:** 1440–50.

25 Boffetta P, Trichopoulos D. Cancer of the lung, larynx and pleura. In: Adami HO, Hunter D, Trichopoulos D, eds. Textbook of cancer epidemiology, 2nd edn. Oxford: Oxford University Press, 2008.

26 Straif K, Baan R, Grosse Y, Secretan B, El Ghissassi F, Cogliano V; WHO International Agency for Research on Cancer Monograph Working Group. Carcinogenicity of household solid fuel combustion and of high-temperature frying. *Lancet Oncol* 2006; **7:** 977–78.

27 World Cancer Research Fund and American Institute for Cancer Research. Second expert report. Food, nutrition, physical activity and the prevention of cancer: a global perspective. Lyon: International Agency for Research on Cancer, 2007.

28 Bell DW, Gore I, Okimoto RA, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in *EGFR*. *Nat Genet* 2005; **37:** 1315–16.

29 Heist RS, Christiani D. EGFR-targeted therapies in lung cancer: predictors of response and toxicity. *Pharmacogenomics* 2009; **10:** 59–68.

30 Johansson M, Relton C, Ueland PM, et al. Serum B vitamin levels and risk of lung cancer. *JAMA* 2010; **303:** 2377–85.

31 Hazra A, Kraft P, Lazarus R, et al. Genome-wide significant predictors of metabolites in the one-carbon metabolism pathway. *Hum Mol Genet* 2009; **18:** 4677–87.

32 Davey Smith G, Ebrahim S, Lewis S, Hansell AL, Palmer LJ, Burton PR. Genetic epidemiology and public health: hope, hype, and future prospects. *Lancet* 2005; **366:** 1484–98.

33 Risch N. The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. *Cancer Epidemiol Biomarkers Prev* 2001; **10:** 733–41.

34 Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic population-based assessment of cancer risk in first-degree relatives of cancer probands. *J Natl Cancer Inst* 1994; **86:** 1600–8.

35 Lindelöf B, Eklund G. Analysis of hereditary component of cancer by use of a familial index by site. *Lancet* 2001; **358:** 1696–98.

36 Dong C, Hemminki K. Modification of cancer risks in offspring by sibling and parental cancers from 2,112,616 nuclear families. *Int J Cancer* 2001; **92:** 144–50.

37 Amundadottir LT, Thorvaldsson S, Gudbjartsson DF, et al. Cancer as a complex phenotype: pattern of cancer distribution within and beyond the nuclear family. *PLoS Med* 2004; **1:** 65.

38 Jonsson S, Thorsteinsdottir U, Gudbjartsson DF, et al. Familial risk of lung carcinoma in the Icelandic population. *JAMA* 2004; **292:** 2977–83.

39 Lorenzo Bermejo J, Hemminki K. Familial lung cancer and aggregation of smoking habits: a simulation of the effect of shared environmental factors on the familial risk of cancer. *Cancer Epidemiol Biomarkers Prev* 2005; **14:** 1738–40.

40 Birch JM, Alston RD, McNally RJ, et al. Relative frequency and morphology of cancers in carriers of germline *TP53* mutations. *Oncogene* 2001; **20:** 4621–28.

41 Olivier M, Goldgar DE, Sodha N, et al. Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and *TP53* genotype. *Cancer Res* 2003; **63:** 6643–50.

42 Bailey-Wilson JE, Amos CI, Pinney SM, et al. A major lung cancer susceptibility locus maps to chromosome 6q23-25. *Am J Hum Genet* 2004; **75:** 460–74.

43 You M, Wang D, Liu P, 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.

44 Sierra DA, Gilbert DJ, Householder D, et al. Evolution of the regulators of G-protein signaling multigene family in mouse and human. *Genomics* 2002; **79:** 177–85.

45 Shi X, Zhou S, Wang Z, Zhou Z, Wang Z. *CYP1A1* and *GSTM1* polymorphisms and lung cancer risk in Chinese populations: a meta-analysis. *Lung Cancer* 2008; **59:** 155–63.

46 Vineis P, Anttila S, Benhamou S, et al. Evidence of gene gene interactions in lung cancinogenesis in a large pooled analysis. *Carcinogenesis* 2007; **28:** 1902–5.

47 Benhamou S, Lee WJ, Alexandrie AK, et al. Meta- and pooled analyses of the effects of glutathione S-transferase M1 polymorphisms and smoking on lung cancer risk. *Carcinogenesis* 2002; **23:** 1343–50.

48 García-Closas M, Malats N, Silverman D, et al. *NAT2* slow acetylation, *GSTM1* null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 2005; **366:** 649–59.

49 Brennan P, McKay J, Moore L, et al. Uncommon *CHEK2* mis-sense variant and reduced risk of tobacco-related cancers: case control study. *Hum Mol Genet* 2007; **16:** 1794–801.

50 Cybulski C, Masojc B, Oszutowska D, et al. Constitutional *CHEK2* mutations are associated with a decreased risk of lung and laryngeal cancers. *Carcinogenesis* 2008; **29:** 762–65.

51 Kilpivaara O, Vahteristo P, Falck J, et al. *CHEK2* variant I157T may be associated with increased breast cancer risk. *Int J Cancer* 2004; **111:** 543–47.

52 Cybulski C, Górski B, Huzarski T, et al. *CHEK2* is a multiorgan cancer susceptibility gene. *Am J Hum Genet* 2004; **75:** 1131–35.

53 Antoni L, Sodha N, Collins I, Garrett MD. CHK2 kinase: cancer susceptibility and cancer therapy—two sides of the same coin? *Nat Rev Cancer* 2007; **7:** 925–36.

54 Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. *PLoS Genet* 2009; **5:** 1000477.

55 Orr N, Chanock S. Common genetic variation and human disease. *Adv Genet* 2008; **62:** 1–32.

56 Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* 2008; **452:** 633–37.

57 Amos CI, Wu X, Broderick P, 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.

58 Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 2008; **452:** 638–42.

59 Wang Y, Broderick P, Webb E, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet* 2008; **40:** 1407–9.

60 Landi MT, Chatterjee N, Yu K, 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.

61 McKay JD, Hung RJ, Gaborieau V, et al. Lung cancer susceptibility locus at 5p15.33. *Nat Genet* 2008; **40:** 1404–6.

62 Broderick P, Wang Y, Vijayakrishnan J, et al. Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res* 2009; **69:** 6633–41.

63 Li Y, Sheu CC, Ye Y, et al. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. *Lancet Oncol* 2010; **11:** 321–30.

64 Pillai SG, Ge D, Zhu G, et al. A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet* 2009; **5:** 1000421.

65 Lips EH, Gaborieau V, McKay JD, et al. Association between a 15q25 gene variant, smoking quantity and tobacco-related cancers among 17 000 individuals. *Int J Epidemiol* 2010; **39:** 563–77.

66 Wu C, Hu Z, Yu D, et al. Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Res* 2009; **69:** 5065–72.

67 Shiraishi K, Kohno T, Kunitoh H, et al. Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. *Carcinogenesis* 2009; **30:** 65–70.

68 Grando SA. Cholinergic control of epidermal cohesion. *Exp Dermatol* 2006; **15:** 265–82.

69 Schuller HM. Is cancer triggered by altered signalling of nicotinic acetylcholine receptors? *Nat Rev Cancer* 2009; **9:** 195–205.

70 Tournier JM, Maouche K, Coraux C, et al. Alpha3alpha5beta2-Nicotinic acetylcholine receptor contributes to the wound repair of the respiratory epithelium by modulating intracellular calcium in migrating cells. *Am J Pathol* 2006; **168:** 55–68.

71 Al-Wadei HA, Schuller HM. Nicotinic receptor-associated modulation of stimulatory and inhibitory neurotransmitters in NNK-induced adenocarcinoma of the lungs and pancreas. *J Pathol* 2009; **218:** 437–45.

72 Paliwal A, Vaissière T, Krais A, et al. Aberrant DNA methylation links cancer susceptibility locus 15q25.1 to apoptotic regulation and lung cancer. *Cancer Res* 2010; **70:** 2779–88.

73 Lam DC, Girard L, Ramirez R, et al. Expression of nicotinic acetylcholine receptor subunit genes in non-small-cell lung cancer reveals differences between smokers and nonsmokers. *Cancer Res* 2007; **67:** 4638–47.

74 Truong T, Hung RJ, Amos CI, et al. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. *J Nat Cancer Inst* 2010; published online June 14. DOI:10.1093/jnci/djq178.

75 Lansdorp PM. Telomeres and disease. *EMBO J* 2009; **28:** 2532–40.

76 Fernandez-Garcia I, Ortiz-de-Solorzano C, Montuenga LM. Telomeres and telomerase in lung cancer. *J Thorac Oncol* 2008; **3:** 1085–88.

77 Chen H, Li Y, Tollefsbol TO. Strategies targeting telomerase inhibition. *Mol Biotechnol* 2009; **41:** 194–99.

78 Yoshiura K, Machida J, Daack-Hirsch S, et al. Characterization of a novel gene disrupted by a balanced chromosomal translocation t(2;19)(q11.2;q13.3) in a family with cleft lip and palate. *Genomics* 1998; **54:** 231–40.

79 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.

80 Rafnar T, Sulem P, Stacey SN, et al. Sequence variants at the *TERT-CLPTM1L* locus associate with many cancer types. *Nat Genet* 2009; **41:** 221–27.

81 Petersen GM, Amundadottir L, Fuchs CS, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet* 2010; **42:** 224–28.

82 Lindström LS, Hall P, Hartman M, Wiklund F, Czene K. Is genetic background important in lung cancer survival? *PLoS One* 2009; **4:** 5588.