Single nucleotide polymorphisms as susceptibility, prognostic, and therapeutic markers of nonsmall cell lung cancer

Shanbeh Zienolddiny
Vidar Skaug
Section for Toxicology and Biological Work Environment, National Institute of Occupational Health, Oslo, Norway

Abstract: Lung cancer is a major public health problem throughout the world. Among the most frequent cancer types (prostate, breast, colorectal, stomach, lung), lung cancer is the leading cause of cancer-related deaths worldwide. Among the two major subtypes of small cell lung cancer and nonsmall cell lung cancer (NSCLC), 85% of tumors belong to the NSCLC histological types. Small cell lung cancer is associated with the shortest survival time. Although tobacco smoking has been recognized as the major risk factor for lung cancer, there is a great interindividual and interethnic difference in risk of developing lung cancer given exposure to similar environmental and lifestyle factors. This may indicate that in addition to chemical and environmental factors, genetic variations in the genome may contribute to risk modification. A common type of genetic variation in the genome, known as single nucleotide polymorphism, has been found to be associated with susceptibility to lung cancer. Interestingly, many of these polymorphisms are found in the genes that regulate major pathways of carcinogen metabolism (cytochrome P450 genes), detoxification (glutathione S-transferases), adduct removal (DNA repair genes), cell growth/apoptosis (TP53/MDM2), the immune system (cytokines/chemokines), and membrane receptors (nicotinic acetylcholine and dopaminergic receptors). Some of these polymorphisms have been shown to alter the level of mRNA, and protein structure and function. In addition to being susceptibility markers, several of these polymorphisms are emerging to be important for response to chemotherapy/radiotherapy and survival of patients. Therefore, it is hypothesized that single nucleotide polymorphisms will be valuable genetic markers in individual-based prognosis and therapy in future. Here we will review some of the most important single nucleotide polymorphisms in the metabolic pathways that may modulate susceptibility, prognosis, and therapy in NSCLC.

Keywords: lung cancer, susceptibility, polymorphism, single nucleotide polymorphism, non-small cell lung cancer

Introduction
Cancer is a major public health problem throughout the world and is the third most common disease after cardiovascular and communicable diseases. According to 2011 health statistics published by the World Health Organization, 7.6 million deaths (13% of all deaths) were due to cancer in 2008, and it is estimated that this number will rise to 11 million deaths by 2030.1 The latest figures from United States Cancer Statistics indicate that there will be 1.6 million new cancer cases in 2011 in the US alone, of which 0.6 million are estimated to die from the disease.2

Among the most frequent cancer types observed globally (prostate, breast, colorectal, stomach, skin, lung) lung cancer is the one which kills the greatest number of patients. According to recent surveys published by United States Cancer Statistics, lung cancer...
is the leading cause of cancer death in the United States. Establishment of national cancer registries and efforts by the International Agency for Research on Cancer have led to improved cancer statistics nationally and globally. Data from national cancer registries and the International Agency for Research on Cancer have helped to determine accurately the etiological factors for lung cancer and to design national prevention strategies based on knowledge of specific risk factors in each region or country. Furthermore, data on ethnic differences in risk of acquiring lung cancer are becoming available for many countries with multiethnic populations. These data indicate that in addition to external risk factors, genetic factors may also contribute to variations in risk of developing lung cancer. Specifically, normal genetic variations at the DNA level, known as single nucleotide polymorphisms (SNPs) may contribute greatly to the understanding of both interindividual and interpopulation differences, given exposure to similar environmental and lifestyle factors. A large number of genetic polymorphisms have been identified as candidate biomarkers for lung cancer. Many of these polymorphisms have been replicated and validated by the International Lung Cancer Consortium which has a database of lung cancer containing thousands of lung cancer cases and controls from populations in Asia, Europe, and the Americas.

**Single nucleotide polymorphisms**

The genome of living organisms contains multiple types of structural genetic variations (variation in chromosome number, DNA fragment variations, copy number variations) and molecular levels (base modifications, nucleotide substitutions). An SNP is a common type of variation in the DNA sequence which occurs in greater than one percent of the population. As shown in Figure 1, individuals in a population may inherit these base variations from their parents so that they may be homozygous or heterozygous for an SNP at a specific site of the genome. The number of SNPs has been estimated to be 1/100 nucleotide, totaling about $3 \times 10^7$ in a diploid human genome. Many of these SNPs are located within the regulatory regions of the genes which may influence the expression of the gene and some others are located within the exons or ex-intron boundaries which may modify the protein function or the splicing sites, respectively. The biological consequences of SNPs are shown in Figure 2.

**Polymorphisms in major biological pathways for lung carcinogenesis**

Many of the polymorphisms are found in the genetic pathways controlling enzymes involved in the metabolism and detoxification of tobacco smoke carcinogens, repair of harmful DNA adducts, and suppression of mutations induced by the carcinogens. Many of them are also present in the genes regulating the immune system and cellular stress responses. The functional significance of a few of these polymorphisms has been elucidated. However, since it is hypothesized that an SNP may occur in approximately every 100 nucleotides, many more remain undiscovered. In the following sections, we will focus on a selection of polymorphisms for which there is evidence of association with lung cancer risk and finally we will address some prognostic and therapeutic aspects derived from these data.

**Polymorphisms in genes involved in biotransformation of carcinogens**

It has been known for a long time that there are individual differences in the metabolism, detoxification, and removal of carcinogens from the body. The cytochrome P450 (CYP450) system is the biological pathway to deal with metabolism of chemical carcinogens and synthetic drugs that enter the body. In the human genome, CYP450 is comprised of 115 genes, of which 57 code for bioactive enzymes whereas the remainder are considered to be pseudogenes.

Tobacco smoke is thought to contain more than 4000 chemicals, of which more than 60 are carcinogens, with polycyclic aromatic hydrocarbons (PAHs) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNKs) as the major groups. PAHs are biotransformed by CYP450 enzymes.

![Figure 1](image)
enzymes to epoxides which are then converted by epoxide hydrolase to diol-epoxides that may bind to DNA, forming PAH-DNA adducts. Benzo(a)pyrene is a typical PAH, which is mainly bioactivated to B(a)P-7,8 dihydro-epoxide by CYP4501A1 (CYP1A1) and CYP450 1B1 (CYP1B1). Subsequent hydrolyzation of B(a)P-7,8 dihydro-epoxide by microsomal epoxide hydrolase (EPHX1) and another round of metabolism by CYP1B1, the ultimate carcinogen benzo[a]pyrene-7,8-dihydrodiol-9,10-epoxide, is formed, which is highly mutagenic.14 CYP1A1 and CYP1B1 are found overexpressed in lung cancer.15 Both enzymes are induced by tobacco smoke and their induction may affect the level of PAH-DNA adducts in cells.16 An almost four-fold higher expression of the CYP1A1 gene was found in current female smokers17 and there was a positive correlation between PAH-DNA adducts with CYP1A1 expression in the lung, which may partly explain gender differences in susceptibility to NSCLC.18,19 Higher expression of CYP3A5, another CYP enzyme involved in biotransformation of PAHs, has also been reported to correlate with higher adduct levels in bronchoalveolar macrophages in smokers.20 The CYP1A2 gene is also an important CYP gene which is closely linked to CYP1A1 and may affect carcinogen metabolism, but its expression has not been investigated.21 The NNKs found in tobacco are also linked to the risk of lung cancer. NNK is biotransformed by CYP2A6, and the enzyme is also involved in the metabolism of nicotine, which may affect smoking behavior and addiction.22

The extent of induction of CYP450 genes in the liver and lung varies between smokers in the population, and may be due to the presence of functional polymorphisms in these genes.23,24 Several SNPs in the CYP genes have been associated with the risk of nonsmall cell lung cancer (NSCLC), but only a few of these have been confirmed in multiple studies to be of functional importance. MspI and Ile462Val polymorphisms in the CYP1A1 gene have been studied in relation to lung cancer in many populations. A recent meta-analysis of 71 studies including all ethnic groups has concluded that both polymorphisms may be associated with an increased risk of lung cancer, with odds ratios of 1.19 (95% confidence interval [CI] 1.11–1.28) for MspI and 1.20 (95% CI 1.08–1.33) for Ile462Val.25 However, in a subgroup analysis by ethnicity, significantly increased risks were found for the MspI and Ile462Val polymorphisms among East Asians. No associations were found with Ile462Val, and only marginal significant associations were detected for the MspI polymorphism in Caucasians.

Several polymorphisms in the CYP1B1 gene have been found to be associated with risk of NSCLC.26-28 A meta-analysis including results from 10 studies involving more than 16,000 subjects concluded that CYP1B1 SNPs Leu432Val (rs10566836), Ala119Ser (rs10566827), Arg48Gly (rs10012), and Asn453Ser (rs1800440) were associated with risk for lung cancer.27 The data suggested that homozygote carriers of the variant 432Val/Val polymorphism had 40% higher risk of lung cancer, whereas the Val allele alone was associated with a 26% increased risk. Individuals with 119Ser/Ser genotype
Compared with 119 Ala/Ala had a two-fold increased risk of lung cancer, whereas for the individuals with the 48Gly/Gly genotype, the risk was four-fold compared with carriers of the wild-type 48 Arg/Arg genotype. The Asn453Ser polymorphism was only associated with risk in women.

Although functional data on risk-associated polymorphisms are mostly lacking, some in vivo and in vitro results support higher gene expression and enzymatic activity for the variant alleles of CYP1B1. Moreover, some of the genetic variants in the CYP1B1 gene have been found to affect survival and anticancer drug therapy in NSCLC patients.

The glutathione S-transferases (GSTs) and N-acetyltransferases (NATs) are involved in conjugation and detoxification of reactive metabolites produced by CYP enzymes. The GSTM1, GSTT1, GSTP1, and glutathione peroxidases are the most important members of the phase II enzymes, all of which are polymorphic. The ary1 and heterocyclic amines are metabolized by NAT1 and NAT2 enzymes that act via N- and O-acetylation reactions. Polymorphisms in the NAT genes result in variants of fast, slow, or intermediate acetylator phenotypes in humans. We and others have found significant associations between polymorphisms in several of these genes and risk of NSCLC. For example, the GSTT Met139Ile and GSTP1 Ala114Val substitutions were significantly associated with NSCLC risk. There are multiple studies showing evidence of risk of lung cancer associated with the NAT1 fast acetylator phenotype. For some genes, such as Ala222-Val in methylenetetrahydrofolate reductase (MTHFR), the associations with NSCLC risk are not consistent, although functional analysis of this polymorphism has shown that the Val allele may result in lower enzyme activity, lower folate levels, and lower methylation levels.

In addition to PAHs and NNKs, reactive oxygen/nitrogen species in tobacco smoke may contribute to lung carcinogenesis. Reactive oxygen/nitrogen species are detoxified by a number of antioxidant enzymes, including microsomal EPHX1, myeloperoxidase, catecho1-O-methyltransferase (COMT), and manganese superoxide dismutase. Several functional polymorphisms in these genes have been identified, which may affect gene expression or enzyme activities. His139 Arg polymorphism in the EPHX1 gene, the Val16 Ala in the SOD2 gene, and Val158Met in the COMT gene have been found to modify lung cancer risk.

**Polymorphisms in DNA repair genes**

Cells are exposed to thousands of DNA lesions resulting from exposure to endogenous and exogenous carcinogenic agents. If un repaired, this large number of lesions could lead to global genomic instability in the form of chromosome aberrations, mutations, and DNA rearrangements, which are commonly found in most cancer cells. To avoid genomic instability, cells have evolved damage-specific DNA repair pathways that are tightly regulated and orchestrated in normal cells. A reduced or diminished capacity to repair DNA lesions predisposes individuals to an increased susceptibility to cancer. Individuals with lung cancer have been shown to have a lower DNA repair capacity. However, reduced DNA repair capacity in patients may be beneficial for therapy with alkylating agents which act by inducing damage to the DNA of tumor cells.

There are at least five major DNA repair pathways, each dealing with specific DNA damage, with some proteins having overlapping or backup functions. Bulky DNA damage (ie, PAH adducts) are mainly repaired via nucleotide excision repair pathway, which consists of more than 30 proteins involved in various steps of damage recognition, DNA incision/opening, DNA synthesis, and ligation. The nucleotide excision repair genes mostly studied in relation to cancer are xeroderma pigmentosum (XP) also known as “excision repair cross complementing” (ERCC) where XPA, XPC, ERCC1, ERCC2/XPD, ERCC4/XPF, and ERCC5/XPG are key enzymes. Another DNA repair mechanism that deals with large strand breaks such as double-strand breaks is DNA double-strand break repair, which involves homologous and nonhomologous recombinational end-joining repair pathways, including several protein families such as the RAD51, ataxia telangiectasia-mutated, and x-ray repair cross complementing (XRCC) proteins.

Presence of reactive oxygen/nitrogen species may lead to formation of 7,8-dihydro-8-oxoguanine (8-oxo-G), apurinic/apyrimidinic (AP) sites and single-strand breaks in DNA. These lesions are repaired by the base-excision repair pathway. 8-oxo-G is a major form of DNA damage that mispairs with thymine, leading to mutations (G to T transversion). 8-oxo-G is excised from DNA by the 8-oxoguanine glycosylase-1 (OGG1) protein. The enzymatic activity of OGG1 creates an AP site which is then repaired by the AP endonuclease-1 (APEX/APE1) enzyme. Complete repair of diverse DNA damage requires activity of many other proteins, including the x-ray repair cross complementation-1 (XRCC1), methylpurine glycosylase, polymerase B, ligase-3, exonuclease-1, and proliferating cell nuclear antigen. The base-excision repair pathway has been considered to be important in sensitzation to radiotherapy of lung tumors.
The alkylating agents in tobacco smoke lead to addition of various methyl and alkyl groups to the DNA bases. Among others, the O6-methylguanine (O6-meG) is a major adduct that may mispair with thymine during replication, resulting in G to A mutations. O6-meG is removed from DNA by the suicidal O6-meG-DNA methyltransferase protein in human cells. This protein has been shown to have an important role in the response of patients treated with DNA-alkylating drugs.56,57

A recent meta-analysis including results from 241 associations has confirmed that polymorphisms in the DNA repair genes, APEX/APE, XRCC1, ERCC1, ERCC2, ERCC5, NBN, XPA, POLI, MGMT, and OGG1 could affect the risk of lung cancer.58 A follow-up study of this meta-analysis has found strong cumulative evidence for an association between polymorphisms in the ERCC2 gene and lung cancer risk.59 It is interesting to note that polymorphisms in the ERCC1 and ERCC2 genes have also been found to modulate survival and response to drugs used in therapy for NSCLC.60 A newly published prospective study from a Texas population has also confirmed the association between SNPs in the XRCC4 gene and lung cancer risk.61

We and others have also found polymorphisms in double-strand break repair genes to modulate lung cancer susceptibility.62–64 An almost three-fold increased risk of lung cancer was found for the XRCC-2 Arg188His and homozygote group compared with subjects with the Arg188 Arg genotype.64 Mutation or deletion of the amino acid at this site of XRCC-2 protein reduces cell survival following DNA damage. In a case-control study, we found that the Thr297Ile polymorphism in exon 7 of the XRCC9/FANCG gene was associated with a reduced risk of NSCLC where only 0.3% of lung cancer cases carried variant genotypes compared with 2.5% of controls.65 It was also found that the ATR Thr211Met SNP was associated with a decreased risk of NSCLC. ATR (ataxia telangiectasia and Rad3-related) protein plays a central role in phosphorylation of DNA damage-sensing proteins, such as the ataxia telangiectasia-mutated protein and TP53. The target proteins are phosphorylated by ATR protein in response to double-strand breaks, reactive oxygen/nitrogen species, hypoxia, and inflammation. It has been shown that drugs such as gefitinib may sensitize NSCLC tumors to radiotherapy via inhibition of the ataxia telangiectasia-mutated pathway.66

**Polymorphisms in the TP53 pathway**

TP53 is a tumor suppressor protein and is activated to prevent carcinogenic effects of DNA damage. Activation of TP53 induces cell cycle arrest, allowing DNA repair machinery to repair damage.66,67 Cells with unreparable DNA damage are removed by apoptosis, which is also induced by TP53. The importance of this protein is supported by the fact that the TP53 gene is mutated in more than half of NSCLC tumors, and that mutation of the gene is an early event in lung carcinogenesis.68 The murine double minute 2 (MDM2, also called HDM2), plays a central role in regulation of TP53 activities through its ubiquitin ligase activity leading to rapid degradation of TP53 protein. Furthermore, TP53 induces transcription of the MDM2 gene, generating an autoregulatory feedback loop.69 Overexpression of MDM2 protein has also been observed in lung tumors, which may indicate a complementary mechanism to TP53 inactivation.70

Polymorphisms in the TP53 and MDM2 genes have been reported to affect cancer risk.71 The most functional SNPs in this pathway are those located in exon 4 of the TP53 gene (Arg72Pro) and SNP T309G in promoter of the MDM2 gene.67,72 Our laboratory has shown that lung cancer patients carrying the Pro72 allele may have a lower frequency of TP53 mutations in tumors.71 The 72Pro allele has been suggested to lead to higher transcription of DNA repair genes than the Arg72 allele.74 In contrast, the Arg72 variant has been associated with greater apoptotic potential. The T309G SNP located at position 309 in the first intron of the MDM2 gene has been thoroughly investigated, both in association studies and in functional experiments. This T to G substitution was shown to lead to higher transcription of the gene.72 Our study of SNP309 indicated that the G/G genotype was more frequently found among NSCLC cases than in healthy controls.75 Interestingly, the strongest effect of SNP309 polymorphism was seen among women, suggesting that the effect of overexpression of MDM2 caused by SNP309 may be enhanced by sex hormones, as proposed by Bond et al.67 Some studies have shown the SNP309 to be associated with lower age at diagnosis but our results indicated different roles of SNP309 depending on the TP53 status in the tumors. In lung cancer patients with tumors having mutations in the TP53 gene, the G allele was associated with older age at diagnosis. Conversely, no such effect of age was seen in cases without mutations in TP53. Although few studies have not found SNP309 to be associated with a risk of NSCLC,76,77 arguing that ethnic differences may exist, most of the recently published meta-analyses have found the GG genotype of SNP309 to be associated with a 12%–27% increased risk of lung cancer.78,79 Furthermore, an interaction between Arg72Pro polymorphism of TP53 and SNP309 of MDM2 has been proposed.80 It has also been shown that both of these polymorphisms may affect clinical outcome in NSCLC.81,82
The cysteine-dependent aspartate-specific proteases (CASPs) are also important in apoptotic cell signaling. The CASP8 is a key enzyme in this pathway and a six nucleotide insertion-deletion polymorphism (−652 6N ins/del, rs3834129) in the CASP8 gene promoter was found to be associated with susceptibility to lung cancer although not all studies are consistent.93 In our study, we found that only men with homozygote del/del genotypes had an increased risk of NSCLC, and that the risk was only elevated in patients with a nonmutated TP53 gene.84

**Polymorphisms in inflammatory genes**

Inflammation has been recognized as a factor contributing to the pathogenesis of many cancers where chronic inflammation, arising as a result of continuous exposure to tobacco components, may result in an excess of oxidative stress and contribute to tumor promotion and progression in the lung.85,86 Epidemiological studies have shown an increased risk of lung cancer in patients with inflammatory airway phenotypes, such as chronic obstructive pulmonary disease, emphysema, bronchitis, and asthma.87 The components of cigarette smoke stimulate epithelial and immune cells to overproduce proinflammatory proteins, including interleukins 1-beta (IL-1β), IL-6, IL-8, tumor necrosis factor alfa (TNF-α), and cyclo-oxygenase-2 (COX-2, also known as prostaglandin-endoperoxide synthase 2). Increased levels of these proteins have been observed in NSCLC tumors.88 Moreover, upregulation of the COX-2 protein has been associated with poorer survival in NSCLC in some studies, whereas others have not found such a correlation.91,92 Due to overexpression of COX-2 in tumors, several clinical trials have been designed to investigate the chemopreventive effect of nonsteroidal anti-inflammatory drugs, of which COX-2 is a major target, particularly in smokers.93–95 Furthermore, it has been observed that the main mechanism of action of drugs such as carboxamidotriazole may be through inhibition of proinflammatory cytokines.96 In several studies, we and others have shown that polymorphisms in some of the inflammation genes may predispose individuals to NSCLC.97,100

IL-1β is a central proinflammatory cytokine important in initiating and balancing the level of inflammatory response. In several epidemiological and molecular biological studies we have investigated SNPs in the IL-1B gene in relation to NSCLC and expression of this gene in the lung. We have found statistically significant effects on NSCLC in association with the T-31C, T-3893 G, and −1464 G regulatory SNPs.100,101 Haplotype analyses of the risk SNPs showed that G-3893 A, G-1464C, C-511T, and T-31C SNPs formed a specific haplotype in NSCLC cases but not in controls.101 It was also found that this haplotype was associated with higher IL-1B expression in the lung. In further molecular studies, we focused on the T-31C SNP since the C allele of the SNP disrupted the TATA-box and was thus believed to affect transcription. Indeed, the experimental data from our laboratory showed that the IL-1B promoter with T in this position had higher transcriptional activity compared with the C-SNP.102 On the other hand, the C-SNP created a novel binding site for transcription factor YY1, which can regulate transcription either as an activator or repressor.103 In addition, it was observed that the polymorphisms could affect mutations in the TP53 gene in tumors.104

The IL-6 G-174C SNP was associated with a risk of developing squamous cell carcinoma of the lung.97 This SNP, which is located in the promoter of the IL-6 gene, has been shown to be associated with increased IL-6 and C-reactive protein levels and a decline in lung function in patients with chronic obstructive pulmonary disease.104,105 Indeed, IL-6 has been found to be the predominant cytokine expressed by tumor-associated lymphocytes and macrophages in lung tumors.98 Our data suggested a decreased risk of NSCLC associated with the T-251A SNP of the IL-8 gene in female subjects only, which may partly support gender differences in lung cancer risk. A recent meta-analysis has confirmed the role of IL8 T-251A polymorphism as a susceptibility marker for lung cancer.106 This IL-8 gene promoter SNP has also been reported to affect IL-8 gene expression.107

The cytokine TNF-α is known to influence the pathogenesis of cancer, and several polymorphisms of this gene, such as G-308 A and G488 A, have been suggested to be associated with susceptibility to lung cancer.108 Other polymorphisms in the TNF-α pathway have also been associated with the risk of NSCLC.109 In our case-control study, a moderate association between risk of NSCLC and the TNF-α G488A polymorphism was found.110 Furthermore, it was found that a combination of three out of four polymorphisms in the genes Casp8, IL10, MMP1, and Seps1 was associated with a four-fold increase in the risk of NSCLC.110

A polymorphism in the 3-UTR region of the COX-2 gene (COX-2.8473) may be associated with an increased risk of NSCLC whereas other polymorphisms have not shown associations.97,111 The COX-2 8473 SNP has been suggested to alter stability of the COX-2 mRNA levels, although there are no specific functional data.
Polymorphisms in nicotinic acetylcholine and dopaminergic receptors

Genome-wide association studies have recently reported that polymorphisms at chromosome 5p15.33 (TERT-CLPTM1L), 6p21.33, and 15q25.1 (CHRNA3-CHRNA5) loci influence the risk of developing lung cancer.112,113 We also found an increased and dose-dependent association with risk of NSCLC associated with the studied variants, consistent with previously published data.26 Additionally, our results provided evidence that individuals carrying these SNPs accumulated higher levels of smoking-related DNA adducts in the lung. There is also some evidence that polymorphisms at 15q25.1 that harbor the nicotinic acetylcholine receptor genes, CHRNA3 and CHRNA5, may be related to smoking behavior in smokers.114–116 The SNP at 5p15.13 (rs402710) maps within intron 4 of the cleft lip and palate transmembrane 1-like (CLPTM1L) gene which is upregulated in cisplatin-resistant (CDDP) ovarian tumor cell lines and its overexpression has been shown to induce apoptosis in CDDP-sensitive cells.117

Dopamine receptors have been identified both in human lung cancer cell lines and in the normal lung. It is known that there is great individual variation in activity of the dopaminergic system.118,119 We studied 11 SNPs in the DRD2, DRD4, and DAT1/SLC6A3 genes encoding dopamine receptor or transporter molecules in relation to lung cancer susceptibility. We found that carriers of the DRD2–141Cdel, G3208T, TaqIB_G > A, DRD4 C-521T, and SLC6A3/DAT1 C-1476A had an increased risk of NSCLC.120 It was also found that the G allele of DRD2–1412A > G had a protective effect.

Individual predisposition to lung cancer prognosis and consequences for therapy

Epidemiological studies reveal poor overall survival for patients with lung cancer, with only a slight improvement in recent years.121,122 Mean age-adjusted relative 5-year survival was found to be 12% in the Eurocare database, but there are wide variations between countries.123 Such differences may be due in part to sociodemographic factors, delays in diagnosis, and different therapeutic regimes. Recent data on patient cancer survival from the Nordic cancer registries showed no differences in survival rates in the Nordic countries 3–5 years after diagnosis.124 However, when analyzing survival, irrespective of the follow-up period, shorter life expectancy was registered in Denmark, possibly related to unfavorable stage distribution. The staging of lung cancer at the time of diagnosis reflects the size and grade of the tumor and also the extent of spread from the primary tumor, thus staging is an essential predictor for survival. According to a Surveillance, Epidemiology, and End Results (SEER) report, the 5-year relative survival by stage for lung cancer patients treated in the United States during 2001–2007 was as follows (stage distribution in parentheses): localized tumors (confined to primary site) (15%): 52.0%, regional (spread to local lymph nodes) (22%): 24.2% distant (with metastasis: (56%)/3.6%, and unknown (untagged) (7%): 8.1.

Other predictors of prognosis have shown inconsistent results across studies and within stages at the time of diagnosis. Young age at diagnosis indicates greater survival compared with elderly patients,125–129 although not in all studies.130 Elderly patients present more comorbidity and are generally treated less with primary surgery than younger lung cancer patients.131 Females have better survival132–134 but this is not consistent across all studies.135 Also histological subtypes may influence outcome to some extent, but the results are inconsistent. The largest progress on survival in recent years, related to therapy has been shown to be for adenocarcinoma, as shown in the data from the SEER Cancer Statistics Review, 1975–2008.131 In these SEER data, the one-year survival rate in patients with adenocarcinoma histology increased from 15% to 23%, mainly in the 2002–2005 period in which erlotinib, gefitinib, and pemetrexed were approved.136

Individual susceptibility accounts for additional variation in prognosis within the various prognostic groups that have been under study. Such differences have been shown for an increasing number of polymorphisms in cancer-related genes, although staging is the most important predictive factor for prognosis.126 Even with the introduction of targeted agents and the establishment of multiple lines of therapy, median survival for patients with advanced NSCLC does not considerably extend beyond 1 year. Genetic susceptibility factors are nevertheless crucial for continuous new research and development towards targeted therapy for individual lung cancer patients.

In a systematic review of 90 studies on genetic polymorphisms and lung cancer outcomes published between 1990 and 2009, 170 genetic variants were reported in relation to overall survival and toxicity. The most frequently studied were polymorphisms in the carcinogen metabolism/biotransformation, DNA repair, TP53/MDM2, and inflammation pathways, some of which are reviewed here. Heterogeneity in study populations, incomplete reporting of important population or study characteristics, inadequate power, and inconsistencies in methodology were common.
Large-scale confirmatory studies on individual polymorphisms are generally not found.\textsuperscript{137}

High susceptibility to smoking-related lung cancer has been associated with polymorphisms of the \textit{CYP1A1} gene. The recent meta-analysis of 71 studies demonstrated that the MspI and Ile462Val polymorphism of the \textit{CYP1A1} gene is a risk factor for lung cancer susceptibility, and that these associations vary in different ethnic populations.\textsuperscript{25} Lung cancer patients from Turkey showed no polymorphisms in \textit{CYP450} genes that were important for prognosis.\textsuperscript{138} Individuals with a susceptible \textit{CYP1A1} genotype have been found to be at higher risk when the genotype is combined with a deficient \textit{GSTM1} genotype\textsuperscript{139} and other risk factors such as \textit{p16} promoter hypermethylation, DAPK, and RAR beta in smokers with null \textit{GSTM1},\textsuperscript{140} as well as \textit{EGFR} polymorphism and its treatment with \textit{EGFR} tyrosine kinase inhibitors.\textsuperscript{141}

The polymorphisms \textit{CYP2D6*4} (C188T), \textit{CYP3A4*3} (G-44 A), and \textit{CYP3A5*3} were found to be associated with response to vinorelbine in a small (\textit{n} = 59) Chinese lung cancer patient group.\textsuperscript{142} Functional polymorphisms in DNA repair genes may interact with chemotherapy and reduce the treatment effect. Cisplatin is a commonly used adjuvant chemotherapeutic agent which acts by cross-linking DNA strands, thus inhibiting DNA replication and disfavors apoptosis of tumor cells. DNA repair enzymes that counteract these effects may diminish the chemotherapeutic effect. Enhanced activity in the nucleotide excision repair pathway may thus unintentionally confer platinum resistance, through enhanced recognition, removal, and resynthesis of DNA. Lung cancer patients with loss of expression of the \textit{ERCC1} protein (by immuno-histochemistry) showed improved survival during cisplatin treatment, compared with patients expressing \textit{ERCC1} in their tumors. In the case of nonadjuvant therapy, subjects expressing \textit{ERCC1} had increased survival compared with patients with abolished expression in their tumors.\textsuperscript{60} This indicates a beneficial effect when it comes to adjuvant cisplatin therapy. In a recent systematic review of \textit{ERCC1} gene polymorphisms in patients with advanced stages of NSCLC, the sensitivity of platinum-based chemotherapy was significantly associated with polymorphism of \textit{ERCC1} C1181T, but not with \textit{ERCC2/XPD} Asp312 Asn and Lys751Gln polymorphisms.\textsuperscript{143} Improved survival for patients harboring the \textit{ERCC1} C1181T has been shown by others\textsuperscript{144,145} in a recent study which also included the C8092A polymorphism with similar improved prognosis when treated with platinum-based chemotherapy.\textsuperscript{146} The \textit{ERCC1} C8092A polymorphism has been shown to be effective for adjuvant platinum therapy in a Japanese population independently of \textit{ERCC1} protein expression\textsuperscript{147} and the variant allele has also been proven as a prognostic marker when adjuvant chemotherapy was given to lung cancer patients in Greece and Spain.\textsuperscript{148,149} In a Chinese population, among two functional variants in the \textit{ERCC1} 5′-flanking region, \textit{−433T > C} and \textit{−262G > T}, which cooperatively influence transcriptional regulation of \textit{ERCC1}, the 262G allele was associated with better drug response and longer survival time compared with the 262T allele.\textsuperscript{150} Patients with the \textit{XRCC7} 399 Arg/Arg and \textit{ERCC1} 8092CA genotypes are susceptible to treatment with gefitinib.\textsuperscript{151} Other variant alleles in the DNA repair pathways are also sensitive when lung cancer patients are treated with chemotherapeutic agents other than platinum.\textsuperscript{147,152,153}

The \textit{TP53} pathway, involving \textit{TP53}, \textit{MDM2}, \textit{MDM4}, and \textit{P73} among others, mediates cellular stress responses inducing DNA repair, cell-cycle arrest, senescence, and apoptosis. These effects enhance the ability to suppress tumor formation and to respond to targeted cancer therapy. Mutations in \textit{TP53} gene in tumors are in most studies related to shortened survival of patients.\textsuperscript{154} Most clinical studies suggest that NSCLC patients with \textit{TP53} alterations have a worse prognosis and may be relatively more resistant to chemotherapy and radiation.\textsuperscript{68} The Arg72Pro polymorphism in \textit{TP53} has been shown to be associated with high-grade NSCLC and significantly shorter survival.\textsuperscript{155} Additionally, there are data that support prognostic effects of intron 3 polymorphism of \textit{TP53} gene that may result in a significantly poorer prognosis in NSCLC.\textsuperscript{155}

The \textit{MDM2} SNP309 was associated with \textit{MDM2} transcripts, mRNA levels, and survival in stage I NSCLC patients with wild-type \textit{TP53} gene in the tumors,\textsuperscript{156} whereas this SNP was not shown to be a predictive marker for survival in other studies.\textsuperscript{157} In stage III–IV NSCLC patients on platinum-based chemotherapy, the combined effect of multiple biologically relevant functional polymorphisms in the \textit{TP53} pathway genes \textit{TP53} Arg72Pro, \textit{P73} G4C14-to-A4T14, and \textit{MDM2} T309G had a higher impact on shortening survival when combined, in a gene-dose-dependent manner, compared with each single polymorphism.\textsuperscript{158} Other polymorphisms in the \textit{TP53} pathway may also be related to prognosis, such as the Arg allele of the p21 \textit{WAF1/CIP1} codon 31, which is a downstream target of \textit{TP53}.\textsuperscript{159} The combined effect of \textit{TP53} with another tumor suppressor gene, \textit{PTEN}, has been shown to decrease survival rates in NSCLC, thus identifying a subgroup of patients with particularly aggressive disease.\textsuperscript{68,160}
Somatic mutations in the *EGFR* tyrosine kinase domain have been demonstrated as being essential for the treatment of NSCLC.\(^{161}\) Polymorphisms in the *EGFR* (rs373506, rs759165, and rs6958497) were shown to be independent candidate biomarkers to predict NSCLC survival in a Chinese population.\(^{162}\) However, there are great interethnic differences\(^{163}\) and the findings are not consistent, due to differences in study design and analysis.\(^{164}\) Functional *EGFR* germ line polymorphisms may confer a risk for *EGFR* somatic mutations in NSCLC with a predominant effect on exon 19 micro deletions.\(^{165}\)

Polymorphisms in the genes involved in the pulmonary inflammatory response may be tightly linked to the carcinogenic process. NSCLC patients with the *CRP* 1846T/T genotype in the C-reactive protein (*CRP*) gene showed a poorer prognosis.\(^{166}\) Furthermore, polymorphisms in the *COX-2* gene associated with reduced overall survival of NSCLC patients has been shown.\(^{167}\) A functional polymorphism of the innate immunity gene *MBL2* has also been associated with lung cancer survival among white patients.\(^{168}\) Polymorphisms in inflammatory genes have been shown to increase the risk of inflammatory side effects in NSCLC patients treated with radiation or chemoradiation. Among 37 inflammation-related genes in 173 patients with stage III disease, 59 SNPs were profiled. Increased risk for esophagitis was observed for nine SNPs of *COX-2* variants, ie, rs20417 (hazards ratio [HR] 1.93, 95% CI 1.10–3.39), rs5275 (HR 1.58, 95% CI 1.09–2.27), and rs689470 (HR 3.38, 95% CI 1.58, 95% CI 1.09–2.27). Polymorphisms in the proinflammatory genes *IL-1A* (rs1800587), *IL-8* (rs4073), *TNF-α* (rs1799724), *TNFRSF1B* (rs1061622), and *MIF* (rs7555622), were related to increased risk for pneumonitis, whereas a polymorphism in the *NOS-3* gene (rs1799983) displayed a protective effect with a 45% reduction in pneumonitis risk (HR 0.55, 95% CI 0.31–0.96). The risk of pneumonitis was also modulated by polymorphisms in anti-inflammatory genes, including genetic variation in *IL13* (rs20541 and rs180925) each associated with increased risk (HR 2.95, 95% CI 1.14–7.63 and HR 3.23, 95% CI 1.03–10.18, respectively).\(^{169}\)

MicroRNAs (miRNAs) are novel regulators of gene expression that are linked to the main oncogene networks, including the *TP53* pathway.\(^{170}\) *TP53* regulates the maturation process of miRNAs including miR-16 and miR-143. The presence of sequence variants in miRNA genes may influence their processing, expression, and binding to target miRNAs. A single miRNA may have many mRNAs as targets, and therefore variations in miRNAs may affect their expression which, in turn, could affect the respective target mRNAs.\(^{171}\) There is increasing evidence that altered microRNA expression is associated with tumor progression and survival in cancer patients. Several studies have analyzed the role of miRNAs as prognostic markers in NSCLC patients.\(^{172}\) In one study, NSCLC patients were divided into three groups according to expression levels of miR-16, and it was shown that patients with normal levels had the best outcome while those with high levels had the worst outcome. Disease-free survival was 22 months for patients with high levels, 72 months for those with normal levels, and 56 months for those with low levels. The overall survival was 24 months for patients with high levels, 98 months for those with normal levels, and 64 months for those with low levels. The authors concluded that high miR-16 levels may be an independent prognostic factor for poor disease-free survival for NSCLC. Another investigation by Saito et al found that increased levels of miR-21, miR-17, and miR-155 were associated with worse cancer death in cohorts of stage I adenocarcinoma patients from Maryland in the United States, and in Norway and Japan.\(^{173}\) The authors highlight the importance of miR-21 as a target for therapy or a biomarker for early-stage prognosis in adenocarcinoma of the lung. In yet another study, it was found that expression of miR-146b, miR-221, let-7a, miR-155, miR-17–5p, miR-27a, and miR-106a was significantly reduced in the serum of NSCLC cases while miR-29c was significantly increased. The expression profiles of miRNAs showed considerable differences between African-American and European Americans.\(^{174}\) Furthermore, it was found that reduced plasma expression of let-7b was modestly associated with worse cancer-specific mortality in all patients and reduced serum expression of miR-223 was modestly associated with cancer-specific mortality in stage...
I/A/B patients. The authors indicate that expression of let-7b may be associated with prognosis in NSCLC. This study did not find any relationship between miRNA expression and response to adjuvant cisplatin-based chemotherapy.

**Conclusion and perspectives**

Lung cancer has been considered a disease related to lifestyle and environmental factors. However, recent molecular epidemiological studies have identified normal genetic variations in population that may modify the carcinogenic effect of environmental and lifestyle carcinogens. It is also evident from available data that these genetic variations may affect prognosis and possibly treatment outcomes in lung cancer patients (Figure 3). Studies are still needed to replicate, validate, and test functionally the biological role of polymorphisms associated with risk,\(^{175}\) prognosis, and therapy.\(^{175}\) Additionally, two other types of variations known as copy number variations and epigenetic variation affect the expression of the genes and we have not reviewed these variations here.\(^ {177,178}\) These variations are currently being investigated in relation to susceptibility, prognosis, and therapy in lung cancer and there are no conclusive results yet. Studies are also warranted to test the combined effects of both genetic and epigenetic variations. Future large validation studies are needed on the impact of functional SNPs on survival and in therapy. It is also important to address genetic biomarkers coming out of genome-wide association studies which may be beneficial in the future.\(^ {179,180}\) Historical molecular epidemiological studies on genetic susceptibility will also have to address and adapt to the changing staging criteria over time.\(^ {181}\)

In conclusion, detection of functional SNPs in metabolic pathways which are associated with increased risk for lung cancer are important, being candidate markers for early diagnostics, hence early treatment and increased likelihood of longer survival. SNPs are also candidate markers for poor prognosis allowing identification of individuals in need for more aggressive therapy. Finally SNPs may help in choosing the most sensitive treatment and for reducing the risk of toxic side effects from chemotherapy and radiotherapy.

**Acknowledgment**

The financial support of the Norwegian Cancer Society is acknowledged.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


