Section II
Intrapersonal/Individual Factors Associated With Tobacco-Related Health Disparities

Chapter 3
Genetics, Physiological Processes, and Tobacco-Related Health Disparities
Introduction

This chapter explores the relationships between genetic factors and tobacco use behaviors and tobacco-related cancers. First, the chapter discusses genetic factors associated with nicotine metabolism and smoking initiation, progression to established smoking, smoking prevalence, and smoking cessation. Genetic risk factors typically vary in prevalence across racial/ethnic populations and thus can contribute to the manifestation of tobacco-related health disparities (TRHD) among racial/ethnic groups. Second, the chapter describes genetic factors associated with tobacco-related cancers, specifically lung cancer, and genetic factors that may influence how the body responds to carcinogens in tobacco smoke. It is important to keep in mind that complex interactions between genetic and environmental factors, many of which are correlated, influence interindividual susceptibility to harmful smoking behaviors and to the risk of tobacco-related diseases (Figure 3.1). More data are needed on high-risk segments of the population to pursue important leads about the relative role of genetic factors in TRHD.

The information presented in this chapter is based on a survey of the genetic factors for which the evidence is stronger (i.e., larger, more powered studies and/or replicated associations), not a meta-analysis of each gene investigated with respect to tobacco use and smoking behaviors. Broad search terms pertaining to genetics and smoking were first used to identify specific genetic factors, which were then individually investigated using the relevant gene name or gene region. The absence of a description of a specific genetic factor should not be interpreted as a negative result. Given that the overall purpose of this monograph is to explore and understand TRHD, there is a focus on discussing genetic factors that have been investigated across multiple racial/ethnic groups.

The chapter closes with a discussion of the current state of knowledge about genetic influences on TRHD, including critical knowledge gaps, such as the contribution of genetic factors in the context of complex socioeconomic environments.

Figure 3.1 Contribution of Genetic Factors to TRHD

Note: LGBT = lesbian, gay, bisexual, and transgender.
Genetic factors influence each stage of the tobacco use continuum, from initiation to cessation and to tobacco-related diseases. Tobacco use, smoking behaviors, and tobacco-related diseases are not seen uniformly across populations, and specific racial/ethnic and socioeconomic groups tend to bear a disproportionate burden of tobacco-related health outcomes. For example, African Americans report smoking fewer cigarettes per day compared with Americans of European descent, yet they are less likely to quit smoking and more likely to develop lung cancer. Smoking initiation and progression to daily smoking also differ across racial/ethnic groups even after differences in socioeconomic status (SES) are accounted for. In addition to race/ethnicity and SES, another important marker of tobacco-related health risk is sexual orientation: Smoking prevalence is significantly elevated in lesbian, gay, bisexual, and transgender (LGBT) groups. The importance of these factors for different racial/ethnic groups is highlighted, where known. As of 2017, little research evidence is available on genetic factors and smoking behaviors stratified by SES or LGBT status.

Data Limitations on Genetic Factors Related to Disparate Populations

The evidence for a genetic contribution to smoking behaviors has largely been established in twin studies conducted mostly in European and European American populations. These twin studies have estimated that the heritability of smoking initiation is 36%–75%, and the heritability of cigarette consumption is 51%–86%. Genetic factors also play a strong role in nicotine dependence (59%–75%) and smoking cessation (50%–58%). The broad range for estimated heritability of a given smoking measure reflects the differing relative impacts of genetic and environmental influences and depends on multiple variables, including time (cohort), age, race/ethnicity, and societal and cultural context.

It is noteworthy that the genetic components of each smoking measure are only modestly correlated with each other, indicating that there are unique and common genes contributing to each measure. Specific genes and gene variants associated with smoking measures have been identified in candidate gene studies, which investigate variants chosen for their purported biological role, and genome-wide linkage and association studies have identified genomic regions of interest by testing the association of single nucleotide polymorphisms (tag SNPs) that label intervals across the genome with smoking measures.

Due to the complexity of smoking as a behavior, large-scale genetic association approaches involving data from many thousands of individuals have been favored to facilitate the detection of individual genetic signals. Out of necessity, such approaches tend to simplify or disregard the influence of the socioenvironmental context on the manifestation of genetic factors. Furthermore, investigations have largely been carried out within epidemiology studies primarily or exclusively conducted with people of European descent. Hence, for reasons of statistical power and the avoidance of population stratification, analyses have typically been restricted to the European subgroup, which is an impediment to our understanding of genetic factors in other racial/ethnic groups. In situations where genomic regions were first identified in a population of European descent and then were subsequently investigated in additional racial/ethnic populations, e.g., African Americans, Chinese, the tag SNPs tested in other populations might not be inherited with the same causal genetic variants because patterns of genetic variation are different across racial/ethnic populations. Until the causal variants are identified, it can be difficult to compare the relative impact of variations in a gene across different racial/ethnic groups or to investigate the role of genetic variants between sociodemographic and environmental contexts relevant to TRHD.
Genetic Factors Associated With Nicotine and Smoking

Overview

Genetic factors associated with smoking behaviors include genetic variations in neurotransmitter systems within the brain reward pathways, neuronal plasticity and connectivity, and nicotine metabolism. In particular, and as expected, genetic variations in the nicotinic cholinergic system have been associated with a range of smoking behaviors, including smoking quantity and intensity, the risk of being nicotine dependent, and the level of nicotine dependence. The binding of nicotine to nicotinic acetylcholine receptors in the brain results in the release of neurotransmitters, such as dopamine, which is thought of as a neurotransmitter that signals reward-related events. For this reason, the genes coding for nicotinic receptor subunits have been the subject of intensive research efforts. The α4 and β2 nicotinic receptor subunits, coded by the CHRNA4 and CHRNB2 genes, respectively, were leading candidates; these subunits are the most populous in the brain, form receptors with the highest affinity for nicotine, and alter nicotine self-administration in animals (as investigated using genetic and pharmacological manipulations). However, genetic variations in other nicotinic receptor subunits are the most strongly associated with smoking behaviors.

Genetic variations in the dopaminergic system also modulate smoking behaviors. Cigarette smoking increases dopamine in the brain, and genes in this neurotransmitter system have been investigated as potential modifiers of smoking behaviors. However, genetic associations in the dopaminergic system are less universally reproducible than variations in nicotinic receptor subunits, possibly because dopamine release is a downstream consequence of nicotine binding to nicotinic receptors in the brain. In addition, the dopaminergic system is not specific to smoking; this system is a convergent pathway for many addictive (and other) behaviors, and genetic variations have been associated with multiple substance dependencies.

Associations between genetic variations in other neurotransmitter systems and smoking behaviors are more equivocal. However, the glutamate receptor subunit gene GRIN3A is associated with smoking quantity and nicotine dependence scores (as determined by the Fagerström Test for Nicotine Dependence [FTND], a 6-item questionnaire scored from 0 to 10, which is predictive of relapse and primarily assesses aspects of withdrawal and the urge to smoke). Moreover, the gamma-aminobutyric acid (GABA) receptor subunit genes GABRA2 and GABRA4 are associated with an increased risk of being a dependent (FTND ≥ 4) versus a nondependent (FTND = 0) smoker. Studies focused on genetic variations in the serotonergic system, which are generally centered on a serotonin transporter variant with reduced expression, have largely failed to demonstrate significant associations with smoking initiation, behaviors, or cessation.

Aside from genes in neurotransmitter pathways, genes involved with the formation and strengthening of neural connections are also associated with smoking behaviors. NRXN1, which codes for the neurexin 1 cell surface protein, was the strongest signal in a genome-wide association study (GWAS) on the risk of being a dependent (FTND ≥ 4) versus a nondependent (FTND = 0) smoker. In addition, NTRK2, which codes for the brain-derived neurotrophic factor (BDNF) receptor, is associated with smoking quantity and FTND scores. The association between variations in these genes, smoking quantity, and nicotine dependence suggests that genetic differences in learning and memory processes influence smoking behaviors.
Genetic factors that influence nicotine metabolism are also associated with smoking behaviors. Nicotine is the main psychoactive ingredient in cigarettes that establishes and maintains dependence.\textsuperscript{41} The complex biology of nicotine addiction is shown in Figure 3.2. Nicotine from inhaled cigarette smoke is rapidly and extensively metabolized by the liver. On average, less than 10\% of absorbed nicotine is excreted in the urine unchanged.\textsuperscript{42} Smokers adapt their smoking behaviors to maintain a preferred level of nicotine. The manipulation of nicotine clearance, through changes in nicotine metabolism or renal elimination, is associated with compensatory changes in smoking behavior.\textsuperscript{43,44} Thus, genetic factors that affect the amount of nicotine available to bind to receptors in the brain, such as variations in the main nicotine metabolism gene, \textit{CYP2A6}, are associated with cigarette consumption patterns, nicotine dependence, and smoking cessation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{biology_of_nicotine_addiction.png}
\caption{Biology of Nicotine Addiction}
\end{figure}

Source: Benowitz 2010.\textsuperscript{24}
Nicotine Metabolism

Based on the existing evidence, overall rates of nicotine metabolism tend to be faster among populations of European descent; slower in Africans; and slowest in Asians, with Japanese as the slowest characterized population. Early evidence in the Yupik population, a subgroup of Alaska Natives, indicates that nicotine metabolism could be highest in that population. Within a population, rates of nicotine metabolism also show large interindividual variations.

Genetic factors can account for a substantial proportion of the variability observed in the rate of nicotine clearance among individuals and racial/ethnic populations. A twin study in European Americans estimated that additive genetic factors explained 59% of the variability in nicotine clearance.

Cytochrome P450 genes produce enzymes whose action forms (synthesizes) chemicals or breaks them down (metabolizes them) to either non-reactive or reactive metabolites (Figure 3.3). In humans, names of the many different cytochrome P450 genes and their enzymes begin with “CYP.” These CYP enzymes are extensively involved in metabolizing the carcinogens and toxicants such as nicotine found in cigarettes, and other forms of tobacco, drugs, and environmental chemicals typically influence whether the metabolism is fast or slow. These genes can be classified as Phase 1 and Phase 2 genes. Phase 1 genes can activate carcinogens by creating a reactive metabolite that binds to proteins or DNA or metabolize them to metabolites that are excreted. Phase 2 genes generally de-activate these reactive substances, which are also then excreted, as shown in the figure.

Figure 3.3  Phase 1 and Phase 2 Drug-Metabolizing Enzymes

Most nicotine (80%) is metabolically inactivated to cotinine,\(^5^4\) and the \textit{CYP2A6} enzyme mediates approximately 90% of this inactivation pathway.\(^5^5\) In individuals who lack functional \textit{CYP2A6}, there is a dramatic reduction in the rate of nicotine clearance.\(^4^6,5^1,5^6,5^7\) In addition to \textit{CYP2A6}, other CYPs (e.g., \textit{CYP2B6}) make a minor contribution to the inactivation of nicotine to cotinine.\(^5^8\)

Further metabolism of cotinine to trans-3’-hydroxycotinine is exclusively mediated by \textit{CYP2A6}.\(^5^1,5^9\) The ratio of trans-3’-hydroxycotinine to cotinine, often referred to as the nicotine metabolic ratio, is correlated with nicotine clearance and serves as a phenotypic marker for the rate of nicotine metabolism and for \textit{CYP2A6} activity.\(^5^1\) Nicotine is also inactivated to several minor metabolites, including nicotine N-oxide (~4%) and nicotine N-glucuronide (~4%); these pathways of inactivation are catalyzed by flavin-containing monoxygenases (FMOs) such FMO3, and by uridine diphosphate (UDP) glucuronosyltransferases (UGTs) such as UGT2B10, respectively.\(^4^2,6^0-6^3\)

Genetic variants in \textit{CYP2A6} are the most established genetic factors associated with nicotine clearance owing to the substantial contribution of \textit{CYP2A6} to nicotine metabolism and to the characterization of numerous \textit{CYP2A6} alleles (variant forms of the gene) with altered activity. The \textit{CYP2A6} genotype influences many measures of nicotine metabolism and clearance, such as total and nonrenal clearance, clearance to cotinine, nicotine half-life, and the ratio of trans-3’-hydroxycotinine to cotinine.\(^5^0,5^1\) Genetic variants in \textit{CYP2A6} that affect activity include SNPs as well as gene deletions, duplications, and conversions.\(^6^4\)

Examples of \textit{CYP2A6} alleles encoding \textit{CYP2A6} enzymes that are inactive toward nicotine include the \textit{CYP2A6}*2, *4, *7, and *17 alleles.\(^6^4-6^6\) The \textit{CYP2A6}*4 allele is an example of a deletion allele; it has 0% enzymatic activity relative to the nonvariant (wild-type) allele—\textit{CYP2A6}*1. Individuals generally possess two alleles for each gene; thus, individuals possessing two copies of \textit{CYP2A6}*4 have no \textit{CYP2A6} enzymatic activity, resulting in nearly undetectable levels of cotinine and no detectable trans-3’-hydroxycotinine, and the small amount of cotinine formation that takes place is catalyzed by other enzymes.\(^4^6,5^7,6^7\) Possession of one copy each of the 0% activity (i.e., inactive) \textit{CYP2A6}*4 allele and the 100% activity \textit{CYP2A6}*1 allele is associated with a 50% reduction in both \textit{CYP2A6} activity and total nicotine clearance, on average, compared with the possession of two copies of the \textit{CYP2A6}*1 allele.\(^5^0,6^8\)

Some \textit{CYP2A6} alleles have decreased, rather than absent, nicotine metabolic activity; as examples, the \textit{CYP2A6}*9 and *12 alleles have approximately 50% lower activity compared with \textit{CYP2A6}*1.\(^6^4\) On average, individuals possessing one copy of the 50% activity \textit{CYP2A6}*9 or *12 allele and one copy of the 100% activity \textit{CYP2A6}*1 allele have 75% of the \textit{CYP2A6} activity and 80% of the total nicotine clearance of \textit{CYP2A6}*1/*1 individuals.\(^5^0,6^8\)

Due to the large number of low-frequency \textit{CYP2A6} alleles, individuals with reduced-activity alleles are commonly grouped together as \textit{CYP2A6}-reduced (<75% activity) metabolizers for data analyses.\(^6^9\) \textit{CYP2A6}-reduced metabolizers can be further subdivided by genotype into \textit{CYP2A6}-slow (≤50% activity) and \textit{CYP2A6}-intermediate (~75% activity) metabolizers.\(^6^8,7^0,7^1\) A small number of \textit{CYP2A6} alleles with increased activity, such as \textit{CYP2A6}*1B and the duplication allele \textit{CYP2A6}*1X2, have also been characterized.\(^7^2-7^4\)

Thus, genetic variations in \textit{CYP2A6} result in a wide range of enzyme activity and, consequently, are associated with a wide range of nicotine metabolism and clearance rates. Consistent with the prominent role of \textit{CYP2A6} in the metabolic inactivation of nicotine and total nicotine clearance, genetic variations...
in \textit{CYP2A6} are strongly associated with multiple smoking behaviors. The most robust genetic associations pertain to cigarette consumption, followed by nicotine dependence and smoking cessation (discussed in subsections below).

Variations in \textit{CYP2A6} contribute to the racial/ethnic differences observed in nicotine metabolism. Although \textit{CYP2A6} alleles have a similar impact on \textit{CYP2A6} activity and nicotine metabolism among different racial/ethnic populations,\textsuperscript{46,57,71} the frequency of alleles varies significantly across populations, resulting in large differences in the relative rates of nicotine metabolism. The collective frequency of reduced-activity genetic variants parallels the population rates of nicotine metabolism; overall, \textasciitilde{}20\% of whites, \textasciitilde{}40\% of African Americans, \textasciitilde{}55\% of Chinese, and \textasciitilde{}80\% of Japanese populations have \textit{CYP2A6}-reduced metabolism genotypes.\textsuperscript{64} Recent investigations into populations that are less well characterized in terms of \textit{CYP2A6}, such as Alaska Natives, suggest that novel increased-activity \textit{CYP2A6} gene variants could be contributing to the comparatively high rates of nicotine metabolism that persist after accounting for known genetic variants.\textsuperscript{48}

Genetic variants such as the \textit{CYP2A6}*4 and *9 alleles are found in all populations tested to date, albeit at different allele frequencies. For example, the frequency of the \textit{CYP2A6}*4 deletion allele ranges from 0\% to 4\% in white, 0\% to 2\% in African American, 5\% to 15\% in Chinese, and 17\% to 24\% in Japanese populations. Other variants are found predominantly in one population; the \textit{CYP2A6}*7 allele is typically only detected in Asian populations (6\%–13\% frequency),\textsuperscript{46,75–77} and the \textit{CYP2A6}*17 allele is typically only detected in African American or black populations (7\%–11\% frequency).\textsuperscript{46,71} On a population level, the overall percentage of individuals possessing altered-activity \textit{CYP2A6} gene variants (i.e., the portion of the population with reduced activity) matters more to the overall rate of nicotine metabolism than the specific gene variants that are found in that racial/ethnic population, because the impact of characterized \textit{CYP2A6} variants is similar across populations, as assessed using genotype-to-phenotype measurements.

In addition to \textit{CYP2A6}, \textit{CYP2B6} also converts nicotine to cotinine, and there are \textit{CYP2B6} genetic variants with increased and decreased enzymatic activity.\textsuperscript{78} However, \textit{CYP2B6} plays a comparatively minor role in nicotine metabolism, and variations in the \textit{CYP2B6} gene have not been associated with differences in nicotine metabolism when activity at the adjacent \textit{CYP2A6} gene was accounted for.\textsuperscript{79,80} Similarly, variations in the genes coding for other enzymes involved with the metabolic inactivation of nicotine, such as FMOs and UGTs, have a small influence on the total removal of nicotine. Therefore, although not studied extensively, these genes have not been shown to significantly alter smoking behaviors.\textsuperscript{63,64,81} However, although much of the research has focused on clearance of nicotine in major body compartments, some of these non-\textit{CYP2A6} pathways such as FMO may affect the clearance of nicotine from the central nervous system.\textsuperscript{82}

**Smoking Initiation**

Smoking initiation occurs predominantly in adolescence, a time when environmental influences account for a greater proportion of risk of initiation, in contrast with early and middle adulthood, when genetic factors increase in importance (Figure 3.4).\textsuperscript{16,83} Investigating the role of genetic versus environmental factors in smoking initiation is made more challenging by the fact that the relative contributions of these factors change with age. Smoking initiation in adolescence is particularly characteristic of populations of European descent; as outlined in chapter 2, a greater proportion of African Americans and Asian Americans initiate regular smoking after age 18. For a review of twin studies in adolescents, refer to
Specific genetic factors involved with smoking initiation have been less well characterized than genetic influences on later stages of the tobacco use continuum. Studies conducted in adult populations are hampered by recall bias and might fail to account for important environmental influences; thus, prospective longitudinal studies in adolescents are better suited to investigating genetic risk factors for smoking initiation in the context of changing socioenvironmental influences. However, these studies are often difficult to conduct, and few have been undertaken to date. (For recommendations on the genetic modeling of smoking trajectories, see NCI Monograph 20, chapter 6. 84) Nevertheless, one longitudinal study of adolescents assessed tobacco use over a period of years and found that having a higher genetic risk score based on variants in genes involved in tobacco dependence was associated with more rapid progression to tobacco dependence and heavier smoking and more failures in cessation attempts, but was not related to smoking initiation. 85

Smoking initiation has been associated with variations in the dopaminergic system in particular. Variations in the TTC12-ANKK1-DRD2 gene cluster and in the dopamine receptor gene DRD4 are associated with an increased risk of ever smoking. 86–89 Of note, the relationship between smoking initiation and variations in the TTC12-ANKK1-DRD2 gene cluster and in the DRD4 gene is influenced by novelty-seeking and depressive symptoms. 86,88,90 This finding highlights the importance of incorporating endophenotypes (heritable observable characteristics) such as personality traits, through which gene variants could be operating, to influence the risk for smoking initiation. 91 Not all studies in adolescents find that TTC12-ANKK1-DRD2 variations are implicated in smoking initiation; these variations are associated with continued smoking, progression to higher levels of smoking, and daily smoking, but not initiation, in other adolescent populations. 32,86,90

Variations in genes involved with neural connectivity and plasticity are also associated with smoking initiation. Cell adhesion genes such as CDH13 and the BDNF receptor gene NTRK2 are associated with adults’ risk of ever having been a smoker versus never having been a smoker, as are variations in the glutamatergic receptor subunit genes GRIN2B, GRIN2A, GRIK2, and GRM8. 35,92 These studies were mainly conducted in European and European American populations, 93 thus data are currently insufficient to determine whether genetic factors contribute to racial/ethnic differences in the epidemiology of smoking initiation.
Smoking Quantity, Dependence, and Age of Smoking Initiation

Genetic factors influence cigarette consumption and nicotine dependence, which differ by race/ethnicity and SES and are important contributors to tobacco-related health outcomes. This section describes the most robust genetic associations and discusses genetic factors within the nicotinic cholinergic and dopaminergic neurotransmitter systems and variations in the nicotine metabolism gene CYP2A6. These genetic factors currently do not explain differences in average cigarette consumption by race/ethnicity, but they contribute to differences among individual smokers of a given race/ethnicity.

Nicotinic Cholinergic System Genetic Factors

Within the nicotinic cholinergic system, genetic variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster are most strongly associated with nicotine dependence and daily cigarette consumption; large-scale meta-analyses of GWASs confirm the association.\(^{20-22}\) Although the genetic association is strong, the effect size of variations in the gene cluster is modest; each minor allele of the most significant genetic marker accounts for only 0.5% of the variance in cigarettes per day, for an increase in daily cigarette consumption of approximately one cigarette.\(^{21}\) However, using cigarette consumption as a surrogate for daily nicotine dose could underestimate the influence of variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster on nicotine intake. Even at a given level of cigarette consumption, individuals differ in how intensively they smoke each cigarette.\(^{94}\) Compared with cigarettes per day, biomarkers of cigarette exposure, such as cotinine and urinary total nicotine equivalents, better reflect smoking level and nicotine intake and can be used to estimate smoking intensity per cigarette.\(^{94-96}\)

There are caveats to comparing cotinine levels among individuals, however since nicotine metabolism varies by age, sex, race, diet, and pregnancy status.\(^{45,97,98}\) Variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster are also associated with differences in cotinine levels among smokers and account for comparatively more of the variation in cotinine than cigarettes per day, which is indicative of more intensive smoking.\(^{99-101}\) For instance, Keskitalo and colleagues discovered that each copy of the genetic variant was associated with an increase in serum cotinine of 77 ng/mL, which would be equivalent to ~6 cigarettes per day, assuming 12–13 ng/mL of cotinine per cigarette, whereas an increase in only 1.2 cigarettes per day per allele was reported, suggesting that the increase in cotinine largely reflected more intensive smoking.\(^{99,102}\) Further evidence for an association of variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster and greater smoking intensity comes from a multiracial/ethnic population study of heavy smoking (greater than 10 cigarettes per day), which assessed total nicotine equivalents adjusted for cigarette consumption.\(^{103}\)

Variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster are also associated with nicotine dependence, as indicated by smoking heaviness and the FTND. Genetic variations in this chromosomal region increase the risk for being a heavy smoker (25 or 30 cigarettes per day) rather than a light smoker (fewer than or equal to 5 cigarettes per day)\(^{104,105}\) (The typical cutoff employed for a light smoker is less than or equal to 10 cigarettes per day.) FTND scores and the risk for being a dependent (FTND \(\geq 4\)) versus a nondependent (FTND = 0) smoker also increase with genetic variations in the cluster.\(^{38,106,107}\)

The GWAS that identified the CHRNA5-CHRNA3-CHRNB4 gene cluster were conducted in European and European American populations.\(^{20-22,108,109}\) Important risk variants in populations of European origin include a specific variation in SNP rs16969968 in exon 5 of CHRNA5, in linkage disequilibrium with rs1051730 in CHRNA3, as well as rs578776 in CHRNA3, a separate SNP.\(^{110}\) Tag SNPs within the region have been subsequently tested in additional racial/ethnic populations. They were found to be a
significant risk factor for cigarette consumption and nicotine dependence in Hispanics, African Americans, and Asians, although the risk SNPs may differ between race/ethnic groups. The cluster was also associated with a biomarker of smoking quantity (total nicotine equivalents) among Alaska Native smokers. Moreover, multiple distinct loci within the CHRNA5-CHRNA3-CHRNB4 gene cluster have been associated with smoking behaviors, but the precise functional variants in the region remain to be confirmed. Once the causal genetic variants have been identified and characterized, it will be feasible to evaluate the role of nicotinic receptor variants in the context of environmental factors that are important to TRHD.

Variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster might also interact with age to influence smoking behaviors. Certain variations in the CHRNA5-CHRNA3-CHRNB4 are also associated with age of onset of smoking. In another study, Weiss and colleagues found that the association between the genomic region and the severity of nicotine dependence was limited to individuals who began daily smoking at or before age 16, whereas Ducci and colleagues found that variations in the region conferred the same degree of risk for smoking at age 14 (regular versus nonsmoker) as at age 31 (heavy versus nonsmoker). A meta-analysis of the CHRNA5-CHRNA3-CHRNB4 SNP rs16969968 compared smoking heaviness among individuals with one risk allele and found that those who started smoking by age 16 were at greater risk for heavy smoking compared with those who started smoking after age 16. Given that the age of smoking initiation varies substantially by race/ethnicity and SES (see chapter 2), age could be an important consideration when interpreting the association of the CHRNA5-CHRNA3-CHRNB4 gene cluster with smoking behaviors.

Additional nicotinic receptor subunit genes associated with smoking include CHRN3 and CHRNA4. Variations in CHRN3 are associated with smoking quantity and the risk of being a dependent versus a nondependent smoker. Variations in CHRNA4 are associated with nicotine dependence, as defined by the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) and the FTND. Variations in CHRN2 might also influence the risk for being a dependent smoker but only through interactions with variations in other genes, such as CHRNA4. Finally, other variants in the cholinergic receptor subunit genes may also be important in African Americans and European-origin populations.

Aside from nicotinic receptor subunits, variations in the choline acetyltransferase gene (ChAT), which codes for a key enzyme in the synthesis of endogenous acetylcholine, are also associated with smoking quantity and FTND scores. Other chromosomal regions and genes might also be important risk factors for cigarette consumption and nicotine dependence in populations of non-European descent. For example, in one study, risk variants for a population of African American but not European origin were found in regions on chromosome 8 and chromosome 14, but not with the CHRNA5-CHRNA3-CHRNA4 cluster. In an Asian population, variants in FRMD4A genes were identified which were associated with nicotine dependence and also age of initiation.

Dopaminergic Neurotransmitter System Genetic Factors

Within the dopaminergic system, genetic variations in the dopamine receptor subunit gene DRD2 have received much attention because of this gene’s central role in the dopamine reward system. The DRD2 gene is part of the TTC12-ANKK1-DRD2 gene cluster, and variations in the cluster are associated with cigarette consumption and the degree of nicotine dependence, as assessed by FTND scores and the
heaviness of smoking index. In addition, variations in *TTC12-ANKK1-DRD2* are also associated with the risk for being nicotine dependent as assessed by the DSM-IV.

The cluster likely influences smoking behaviors through dopamine receptor expression. The commonly investigated Taq1A variant, which resides in the *ANKK1* gene, is associated with reduced *DRD2* mRNA and dopamine D2 receptor density. This genetic region is associated with smoking quantity and nicotine dependence in African Americans and European Americans. However, different locations within the *TTC12-ANKK1-DRD2* gene cluster are associated with dependence risk, possibly owing to differences in the underlying pattern of genetic variations among these racial/ethnic populations.

Few studies have investigated genetic variations in dopamine receptor genes aside from *DRD2*, in part because these other receptors and their genetic variants are not as well characterized. However, functional genetic variants in the *DRD1* and *DRD3* dopamine receptor subunit genes have also been associated with smoking quantity and FTND scores. The relative importance of variations in dopamine receptor genes differs across racial/ethnic populations; family studies suggest that genetic variations in *DRD1* modulate smoking quantity and FTND scores preferentially in African Americans, whereas genetic variations in *DRD3* modulate smoking quantity and FTND scores preferentially in European Americans.

Nicotine Metabolism Gene CYP2A6 Genetic Factors

It is well established that variations in the nicotine metabolism gene *CYP2A6* affect the amount of cigarette consumption. Cigarette smokers who possess *CYP2A6* genotypes that are associated with reduced rates of nicotine metabolism smoke fewer cigarettes per day compared with those who possess normal *CYP2A6* metabolizer genotypes, particularly among racial/ethnic groups characterized by heavier smoking. Among European American smokers, those with *CYP2A6*-reduced metabolizers smoked an average of 20 cigarettes per day, compared with those with *CYP2A6*-normal metabolizers, who smoked 26 cigarettes per day. Among Japanese smokers, cigarette consumption ranged from approximately 15 cigarettes per day in those without functional *CYP2A6* (*CYP2A6* *4/*4) to as many as 30 cigarettes per day in predicted normal metabolizers (*CYP2A6* *1/*1). Chinese smokers with *CYP2A6*-reduced metabolizers also smoked fewer cigarettes per day, but the role of genetic factors in smoking behaviors has not been investigated as extensively in this population. In contrast to heavy-smoking populations, cigarette consumption is not associated with *CYP2A6* genotype groups among African American light smokers or Alaska Native light smokers. Light smoking is also prevalent among Hispanic Americans and in the Asian American aggregate group, but less is known about *CYP2A6* and smoking in these populations.

Among light smokers, biomarkers that operate as reliable indicators of smoking levels in European heavy-smoking populations, such as plasma cotinine, and exhaled carbon monoxide, have limited utility. The relationship between self-reported cigarette consumption and either exhaled carbon monoxide or plasma cotinine is substantially weaker in African American light smokers, making it difficult to investigate the role of *CYP2A6* genotype groups. As of 2015, it was unclear whether differences in the utility of biomarkers could be ascribed to light smokers generally or to African American light smokers specifically. Further complicating the utility of cotinine, the most commonly used marker of nicotine intake, it has been demonstrated that the relationship between cotinine and nicotine dose is affected by *CYP2A6* genetics and sex. Future studies employing more reliable biomarkers of nicotine intake, such as urinary total nicotine equivalents, are necessary to
determine whether variations in \textit{CYP2A6} are an important determinant of smoking behaviors in light-smoking populations. Another biochemical indicator of consumption, carbon monoxide, appears to yield strong associations with \textit{CHRNA5-CHRNA3-CHRNB4} as well as \textit{CYP2A6}.\textsuperscript{139}

As with the \textit{CHRNA5-CHRNA3-CHRNB4} gene cluster, the use of cigarette consumption as a surrogate for daily nicotine dose can underestimate the influence of variations in \textit{CYP2A6} on nicotine intake. Smokers might titrate (i.e., adjust) nicotine levels through their cigarette smoke inhalation patterns as well as change the number of cigarettes smoked. The \textit{CYP2A6} genotype is associated with smoking intensity among European American smokers, with \textit{CYP2A6}-reduced metabolizers taking smaller volume puffs compared with \textit{CYP2A6}-normal metabolizers.\textsuperscript{140} Nicotine titration was also evident in an open-label clinical trial of nicotine replacement therapy (NRT), where \textit{CYP2A6}-reduced metabolizers achieved similar nicotine plasma levels when compared with normal metabolizers by using fewer daily sprays of nicotine nasal spray.\textsuperscript{68}

Variations in \textit{CYP2A6} also influence smokers’ progression toward nicotine dependence and final level of dependence. In European American adolescents, \textit{CYP2A6}-slow metabolizers progress in nicotine dependence at a slower rate and reach a stable level of dependence more quickly compared with normal metabolizers.\textsuperscript{141–143} Slow \textit{CYP2A6} metabolism, however, is a risk factor for acquiring nicotine dependence in adolescence, and the existence of one to two copies of the inactive alleles, \textit{CYP2A6}*2 or *4 increases the risk of conversion to nicotine dependence.\textsuperscript{141} Once dependent, slow metabolizers consume fewer cigarettes compared with normal metabolizers.\textsuperscript{141,142} Thus, longitudinal cohort studies of adolescents have suggested that \textit{CYP2A6}-slow metabolizers acquire nicotine dependence sooner (after initial exposure to nicotine), reach a plateau in their degree of dependence earlier, and have lower levels of cigarette consumption and nicotine dependence compared with normal metabolizers.\textsuperscript{143} Given the importance of environmental influences in adolescence, the association of \textit{CYP2A6} and dependence trajectories must be assessed by race/ethnicity and SES.

Lower levels of nicotine dependence in \textit{CYP2A6}-reduced versus \textit{CYP2A6}-normal metabolizers are also seen in adult European American cigarette smokers, with \textit{CYP2A6}-reduced (<75\% activity, as predicted by genotype) versus \textit{CYP2A6}-normal metabolizers having significantly lower FTND scores.\textsuperscript{135} A component of the FTND is the time to first cigarette in the morning; \textit{CYP2A6}-slow metabolizers trend toward a reduced likelihood of smoking within the first 5 minutes of waking compared with normal metabolizers.\textsuperscript{68} Japanese smokers with \textit{CYP2A6}-reduced metabolism genotypes also have lower FTND scores and are less likely to smoke their first cigarette within 5 minutes of waking compared with \textit{CYP2A6}-normal metabolizers.\textsuperscript{144}

\textbf{Smoking Cessation}

Genetic factors have been demonstrated to affect smoking cessation. Disparities in the ability to quit smoking and the ability to quit smoking with and without nicotine pharmacotherapy are both important risk factors for tobacco-related adverse health outcomes. As outlined in chapter 2, the frequency and success of quit attempts differ by race/ethnicity and SES. For instance, African Americans are less likely to achieve smoking cessation than European Americans. In addition, light smokers appear to have lower abstinence rates compared with moderate-to-heavy smokers on either placebo or pharmacotherapy; however, this finding has not been tested directly.\textsuperscript{145,146}
As with smoking behaviors, multiple neurotransmitter systems are implicated in the ability to quit smoking, with the most replicated associations found among genetic variants in the dopaminergic system. In this system, two variants in the TTC12-ANKK1-DRD2 gene cluster, Taq1A (ANKK1 rs1800497 C>T) and DRD2-141 Ins/Del C, are associated with smoking cessation in clinical trials and in a general care setting, although an earlier meta-analysis found no association between variations in the TTC12-ANKK1-DRD2 gene cluster and smoking cessation. Individuals with the Taq1A T/T (also known as A1/A1) genotype are more likely to be abstinent, regardless of the type of treatment, compared with those with the Taq1A C/C (A2/A2) genotype, but individuals with the Taq1A C/C genotype benefit more from bupropion versus placebo. Regarding the DRD2-141 Ins/Del C variant, individuals with a Del C allele have higher quit rates with transdermal NRT compared with those with the InsC/InsC genotype in an open-label NRT study. Individuals with the InsC/InsC genotype benefit more from bupropion versus placebo. In addition to the TTC12-ANKK1-DRD2 gene cluster, variations in DRD4 are associated with a reduced likelihood of abstinence, regardless of therapy. In contrast, genetic variations in DRD2 were not associated with spontaneous cessation in a large population-based sample of smokers. It is noteworthy that sex might modify the influence of genetic factors in the dopaminergic system; the strength of the association between variations in DRD2 and smoking cessation is related to the proportion of men in the study population.

The catechol-O-methyltransferase (COMT) enzyme, which metabolizes catecholamines, including dopamine, is also associated with transdermal NRT quit rates. Specifically, individuals homozygous for the decreased-activity rs4680 A variant (rs4680 G>A, also known as Val108,158 Met) have a greater probability of quitting with NRT than placebo compared with individuals with the G/G or G/A genotype. Variations in the COMT gene have also been associated with responses to bupropion in which only individuals with the rs4680 A allele benefit from bupropion rather than placebo. Moreover, variations in the dopamine transporter gene, SLC6A3, were associated with cessation in a meta-analysis but not in a more recent population-based study. In addition to clinical studies of drug therapies, an association between the CHRNA5-CHRNA3-CHRNB4 gene cluster and smoking cessation in large human population studies has been observed. In the cholinergic system, variations in the nicotinic acetylcholine receptor subunit CHRNA2 are associated with the ability to quit smoking on either bupropion or placebo even 6 months following treatment. Variations in CHRNA2 were also associated with the ability to remain abstinent for a longer period on transdermal NRT versus placebo in a crossover study. Variations in the choline acetyltransferase gene ChAT are associated with cessation outcomes on transdermal NRT in open-label studies. Initially, variations in the CHRNA5-CHRNA3-CHRNB4 gene cluster appeared to have a negligible role in smoking cessation. However, subsequent investigations revealed an association between SNPs in this cluster and a reduced ability to quit unaided, which was mitigated by pharmacological treatment (NRT in particular). One study also suggested that variants in CHRNA2, CHRNA4, and CHRNA7 and in the CHRNA5-CHRNA3-CHRNB4 gene cluster influence abstinence while an individual is taking varenicline, a pharmacological treatment that acts at nicotinic receptors by partially mimicking the effects of nicotine. Variation in the CHRNA5-CHRNA3-CHRNB4 gene cluster may also influence response to the investigational cessation aid selegiline. Importantly, variation in this gene cluster is also associated with smoking abstinence on active pharmacotherapy among African American smokers.

Genetic variations in the μ-opioid receptor have also been implicated in smoking cessation. Nicotine stimulates the release of endorphin, which binds to μ-opioid receptors, and these receptors mediate
feelings of withdrawal. \(^{170}\) Hence, candidate gene studies have investigated the association between variations in the \(\mu\)-opioid receptor gene \(OPRM1\) and smoking cessation—in particular, the rs1799971 A>G variant, because the G allele is associated with reduced receptor expression. \(^{57}\) In an open-label study of NRT, individuals with the rs1799971 G/G or G/A genotype had higher abstinence rates on transdermal NRT, \(^{171}\) whereas in a placebo-controlled trial of NRT, individuals with the rs1799971 A/A genotype had higher abstinence rates on active treatment. \(^{172}\) These observations, coupled with the \(DRD2\) and \(CHRNA5-CHRNA3-CHRNB4\) findings, highlight the importance of treatment condition in genetic association studies of smoking cessation; the effect of a genetic variant that reduces general quit ability might be mitigated by pharmacotherapy.

Genetic variations in \(CYP2A6\) are associated with smoking cessation, both unassisted and assisted by pharmacotherapy. \(CYP2A6\)-reduced metabolizers appear to have higher levels of smoking cessation, as European, European American, and Japanese individuals possessing the inactive \(CYP2A6^{*4}\) allele have a lower likelihood of being a current smoker and have a greater likelihood of quitting compared with \(CYP2A6\)-normal metabolizers. \(^{70,173,174}\) Consistent with these findings, the proportion of \(CYP2A6\)-slow metabolizers (<50% activity, as predicted by genotype) among current smokers decreases as smoking duration increases. \(^{70}\) The influence of \(CYP2A6\) is even apparent in adolescent ever-smokers, as a greater proportion of slow versus normal metabolizers had quit smoking for at least 1 year. \(^{175}\) Additional evidence for an increased ability to quit smoking among \(CYP2A6\)-slow metabolizers comes from the placebo arm of clinical trials. The slowest quartile of nicotine metabolizers, as assessed by the nicotine metabolic ratio (a biomarker of \(CYP2A6\) activity and genotype), have higher quit rates on placebo compared with the fastest three quartiles of nicotine metabolizers in European Americans and African Americans. \(^{71,176}\)

Rates of nicotine metabolism, which influence how nicotine levels fluctuate after a cigarette, can affect the development of conditioned responses among smokers; functional brain imaging has demonstrated greater neural responses to smoking cues in faster versus slower \(CYP2A6\) genotypes. \(^{177}\) Slow versus fast nicotine metabolism is also predictive of higher abstinence rates on transdermal NRT and of less intense cigarette cravings in the week following the target quit date in populations of predominantly European descent. \(^{178,179}\) Furthermore, in a clinical trial comparing normal and extended NRT, only those with \(CYP2A6\)-reduced metabolism genotypes were found to benefit from extended therapy. \(^{69}\) \(CYP2A6\) is related to cessation success in those getting NRT. There is evidence that cessation success is related to NRT and not to bupropion pharmacotherapy, and that the contribution of \(CYP2A6\) is independent of that of \(CHRNA5-CHRNA3-CHRNB4\). \(^{180}\) Together, these studies suggest that variations in \(CYP2A6\) modulate smoking cessation outcomes; \(CYP2A6\)-reduced metabolizers benefit more from transdermal NRT compared with normal metabolizers and, more importantly, are able to quit more easily even without the use of pharmacotherapy. A clinical trial found that normal nicotine metabolizers were more successful in smoking cessation on varenicline than on nicotine patches, but this was not the case among the slow nicotine metabolizers who also experienced more side effects on varenicline than did the normal metabolizers. This suggests the potential for nicotine metabolism testing to help in identifying who may benefit most from specific cessation therapies. \(^{181}\)

Genetic variations in \(CYP2B6\) also appear to influence smoking cessation outcomes. \(CYP2B6\) genetic variants could be a risk factor for a reduced likelihood of quitting smoking, as reduced-expression variants are associated with increased cigarette craving and reduced abstinence rates in the placebo arm of clinical trials. \(^{80,182,183}\) Importantly, \(CYP2B6\) is the main enzyme that metabolizes the smoking cessation drug bupropion, \(^{184}\) and variations in \(CYP2B6\) are associated with bupropion treatment response.
and with altered levels of hydroxybupropion (an active metabolite) in pharmacogenetic studies of smoking cessation.\textsuperscript{167,185} Thus, although variations in \textit{CYP2B6} might not be a significant genetic factor in nicotine metabolism and cigarette consumption, they might predict the likelihood of quitting smoking unassisted or assisted by bupropion, but possibly not assisted by nicotine replacement therapy.\textsuperscript{183}

\textbf{Summary for Tobacco Smoking Initiation, Nicotine Dependence, and Cessation}

Smoking is a complex behavior, with genetic and environmental influences operating at each stage along the tobacco use continuum. The relative contribution of genetic and environmental factors to a given smoking behavior differs by cohort, age, race/ethnicity, and sociocultural context, and genetic factors characterized in one context cannot be readily extrapolated to another. Most of the evidence concerning genetic factors comes from studies in populations of European descent.

This section described genetic factors in neurotransmitter systems, neuronal connectivity and plasticity, and nicotine metabolism and their effects on smoking behaviors. Common and unique genetic risk factors in neurotransmitter systems and neural connectivity are associated with smoking behaviors along the tobacco use continuum, from initiation to cessation.\textsuperscript{186} Genetic factors related to smoking initiation include variations in dopamine receptor genes and neuronal connectivity genes. Nicotine dependence and smoking levels are strongly associated with variations in nicotinic receptor genes, in particular the \textit{CHRNA5-CHRNA3-CHRNB4} gene cluster. Variations in dopaminergic and neuronal connectivity genes are also associated with nicotine dependence, and variations in dopaminergic genes are associated with smoking cessation. Although racial/ethnic populations appear to share certain genetic factors, such as variations in the \textit{CHRNA5-CHRNA3-CHRNB4} and \textit{TTC12-ANKK1-DRD2} gene clusters, the specific risk SNPs are sometimes different in the different population groups. Studies encompassing each racial/ethnic group are necessary to determine the precise genetic variants operating within each population and the importance of genetic factors relative to sociocultural factors.

Variations in nicotine metabolism genes are also associated with smoking behaviors. \textit{CYP2A6} is the main enzyme that metabolically inactivates nicotine, and variations in the \textit{CYP2A6} gene are an established genetic factor affecting nicotine metabolism and, consequently, nicotine clearance. In line with evidence that cigarette smokers titrate their nicotine intake from cigarettes to maintain a preferred level of nicotine, individuals with \textit{CYP2A6}-reduced metabolism genotypes smoke fewer cigarettes per day and take smaller puff volumes compared with \textit{CYP2A6}-normal metabolizers. \textit{CYP2A6}-reduced metabolizers progress more slowly in their level of nicotine dependence as youths and have lower nicotine dependence scores as adults. Variations in \textit{CYP2A6} are also associated with increased smoking cessation, both the ability to quit without pharmacotherapy and the ability to quit on transdermal NRT. Although the frequency of \textit{CYP2A6}-reduced metabolism genotypes varies according to race/ethnicity, the functional impact of \textit{CYP2A6} genotypes on nicotine metabolism appears to be consistent across racial/ethnic groups. In heavier smoking populations, principally men of Japanese, Chinese, and European descent, a consistent association between variations in \textit{CYP2A6} and smoking behaviors such as cigarette consumption is emerging. Less is known about the effects of \textit{CYP2A6} genetic variations in lighter smoking populations, such as African Americans. Genetic factors that influence smoking behaviors among racial/ethnic populations characterized by lighter smoking are not well understood in part due to the inadequacy of common cigarette smoke exposure biomarkers in light smokers.
Genetic Factors Associated With the Risk for Lung Cancers

Multiple factors contribute to the risk for developing tobacco-related cancers and to population disparities in risk such as differences in smoking behaviors (e.g., the prevalence and/or amount of smoking) and socioeconomic and environmental influences. One such factor, genetic factors, is the focus of this section. Cigarette smoking and secondhand smoke exposure are associated with numerous cancers, including those of the oral cavity, oropharynx, nasopharynx, and hypopharynx, stomach, colorectum, liver, pancreas, nasal cavity and paranasal sinuses, larynx, lung, uterine cervix, ovary (mucinous), urinary bladder, kidney (body and pelvis), ureter, and bone marrow (myeloid leukemia).

Cigarette smoking is the leading risk factor for lung cancer: 80% of all lung cancer deaths in the United States are attributable to tobacco smoke exposure. The risk for lung cancer increases with the level of daily cigarette consumption and the duration of cigarette smoking. Despite a strong association between increasing cigarette exposure and increasing lung cancer risk, susceptibility to lung cancer varies greatly among smokers, particularly among racial/ethnic, low SES, and LGBT groups, and genetic factors have been shown to modulate the risk of developing lung cancer. Although the biological impact of individual genetic variants is predicted to be similar, the frequency of risk factors varies across populations. Furthermore, although heavier cigarette smoking is itself such a strong predictor of lung cancer risk, genetic risk factors are often more pronounced and detectable at lower levels of cigarette smoke exposure.

This section introduces the general mechanisms by which genetic factors influence cancer risk among smokers, with a specific focus on lung cancer risk, and provides examples, where possible, of genetic factors and their importance by race/ethnicity. As of 2015, there was limited, if any, evaluation of genetic factors and lung cancer risk stratified by SES or LGBT status. The intersection of those environments affects health and health services use and could obscure the detection and understanding of genetic factors.

Genetic factors modulate cancer risk among smokers via three general mechanisms (Figure 3.5). First, genetic factors influence cigarette smoke exposure levels by modifying smoking behaviors and smoking cessation outcomes, as discussed in the previous section. Second, genetic factors influence the body’s response to carcinogenic compounds in cigarette smoke, such as the processing of carcinogens for removal, and the ability of carcinogens to interfere with endogenous growth-signaling pathways. The genetic influence on both of these factors, which act proximate to carcinogen exposure, will be the focus of this section. Lastly, genetic factors influence the body’s underlying susceptibility to the damage caused by carcinogens. Gene variations that regulate DNA repair, the cell cycle, and apoptosis are associated with the risk of developing lung cancer—for example, the cell cycle genes p53 and p27; the chromosome 6p21.33 region, which contains a DNA mismatch repair gene, MSH5; and an apoptosis and DNA damage response gene BAT3. Such genetic factors are beyond the scope of this review, but have been discussed by others. Genetic factors might also modulate cancer risk through more than one of the mechanisms described above. For example, genetic variations in CYP2A6 and in CHRNA5-CHRNA3-CHRNB4, which will be discussed, modulate lung cancer risk both indirectly, by influencing smoking behaviors, and directly, by influencing the carcinogenicity of tobacco smoke compounds.
Cigarette smoke delivers at least 69 carcinogens, of which nitrosamines, polycyclic aromatic hydrocarbons (PAHs), aromatic amines, volatile organic chemicals, and heavy metals are among the most potent. Many of the chemicals in cigarette smoke are procarcinogens—chemicals that exert their full carcinogenicity following metabolic activation. In general, chemicals that enter the body, including cigarette smoke carcinogens entering through the lungs, undergo sequential steps of biotransformation (metabolic processing). Initially, such enzymes as CYPs typically make chemical substrates more hydrophilic and more reactive by adding polar chemical groups. These more polar metabolites then become the substrates of classes of transferase enzymes, such as UDP-glucuronosyl transferases (UGTs) and glutathione S-transferases (GSTs), which facilitate elimination by conjugating large, bulky hydrophilic groups onto the metabolites. Alternatively, these transferases could also directly interact with procarcinogens. Thus, in general, CYPs metabolically activate procarcinogens into reactive species capable of damaging DNA, and genetic variations that result in reduced CYP activity are anticipated to reduce cancer risk. In contrast, transferase enzymes metabolically inactivate (i.e., detoxify) carcinogens, and genetic variations that result in reduced transferase activity are anticipated to increase cancer risk by allowing carcinogens to reside in the body for longer periods. In addition to DNA damage, nitrosamines might also foster carcinogenesis by interfering with endogenous nicotinic receptor signaling, which is vital to managing proper cell growth, and genetic variations in nicotinic receptor subunits are associated with lung cancer risk among smokers.
Genetic factors associated with lung cancer risk could offer insight into disparities in the smoking-related risk for lung cancer among racial/ethnic populations. Differences in cigarette consumption cannot readily explain racial/ethnic susceptibility to lung cancer, as African American smokers, a typically light-smoking population, appear to be at a higher risk of lung cancer than European American and Japanese smokers, both of whom are generally heavier smoking populations. The excess risk observed among African Americans may result from the interplay of genetic and environmental factors that need more detailed study.

Lung cancer is a heterogeneous disease defined by histological subtypes and increasingly by molecular signatures. Lung cancers are broadly classified as small cell carcinomas and non-small-cell carcinomas; the latter includes two of the most prevalent histological subtypes of lung cancer—adenocarcinoma and squamous cell carcinoma. Adenocarcinoma is the most prevalent histological subtype overall and within smokers (40% of all lung cancers in the United States), followed by squamous cell carcinoma (20% of lung cancers). Smoking is a risk factor for all histological subtypes of lung cancer. The heterogeneity of lung cancer can confound genetic investigations, as genetic risk factors are unlikely to be uniformly associated with all histological subtypes of lung cancer and could predominantly influence a single subtype. There is a trend in this direction. For example, in a GWAS of lung cancer involving almost 30,000 cases and 56,000 controls, certain susceptibility loci were specific to lung adenocarcinoma.

In addition to smoking behavior, cigarette brand could affect lung cancer histology. The shift from a predominance of squamous cell carcinoma to adenocarcinoma over the past 50 years has coincided with changes in cigarette composition and design, resulting in greater relative exposure to tobacco-specific nitrosamines than PAHs and fostering deeper inhalation. Thus, product preference and trends in use might further complicate the relationship between genetic factors and lung cancer and could possibly contribute to TRHD.

PAHs and nitrosamines are the procarcinogens most strongly associated with lung cancer. Nitrosamines can also foster carcinogenesis by binding to nicotinic receptors, and evidence for the association of variations in nicotinic receptor signaling with lung cancer follows. The next section discusses (1) genetic variations in enzymes that metabolize PAHs and are associated with lung cancer risk, then (2) variations in enzymes that metabolize nitrosamines. Aromatic amines are also strong procarcinogens, but they are more strongly associated with bladder cancer than with lung cancer, therefore genetic variations in metabolizing enzymes that predominantly interact with aromatic amines (such as N-acetyltransferases) are not discussed here.

Genetic Factors Associated With Carcinogen Metabolism

Variations in genes involved with the metabolism of PAHs are associated with differences in lung cancer risk. PAHs such as benzo[a]pyrene are metabolically activated into epoxides and further into diol epoxides capable of reacting with DNA and myeloperoxidase (MPO) are among the enzymes capable of metabolically activating benzo[a]pyrene intermediates and have been studied widely as potential genetic risk factors in lung cancer. Genetic variations in CYP1A1 have been associated with lung cancer in several populations, including African Americans, Asians, Europeans, European Americans, and Indians. They have also been associated with other respiratory cancers, such as oral and pharyngeal cancers. Large-scale meta-analyses supported an association between the CYP1A1 Msp1 variant and increased lung cancer risk in Asian and non-Asian populations, whereas the
**CYP1A1** Ile<sup>462</sup>Val variant is predominantly associated with increased lung cancer risk in Asian populations, probably because of the lower prevalence of this variant in non-Asian populations. Variations in **CYP1A1** appear to be predominantly associated with squamous cell carcinoma. MPO genetic variants with decreased transcription are associated with a reduction in lung cancer risk in populations of European descent, but not all studies have supported this association. Variations in **CYP1B1** have also been investigated, albeit to a lesser extent, and study results have been mixed.

Detoxification enzymes for PAHs include GSTs. Genes coding for GSTs, such as **GSTM1**, **GSTT1**, and **GSTP1**, have been widely studied as genetic factors in cancer risk, and many studies have found them to be associated with lung cancer and head and neck cancer risk in smokers, but not all studies report an association. Some investigations have confirmed the association of the **GSTM1** deletion allele with increased lung cancer risk in Asian populations. A meta-analysis of the **GSTT1** deletion allele also reported a significant association between **GSTT1** and lung cancer risk in Asians but not in Europeans, which was likely due to the lower frequency of the deletion allele in people of European descent, thereby reducing the power to detect the association. The results were more equivocal for a reduced-activity **GSTP1** allele in a combined meta-analysis and pooled analysis in Asian and European populations.

Variations in genes involved with the metabolism of nitrosamines are also associated with lung cancer risk. The drug-metabolizing enzymes **CYP2A6** and **CYP2A13** activate tobacco-specific nitrosamines such as 4-(methylnitrosamino)-1-(3-pyridyl)-1-butane (NNK) and N'-nitrosonornicotine (NNN). Cigarette smokers with **CYP2A6**-reduced metabolism genotypes are associated with a reduction in lung cancer risk compared with **CYP2A6**-normal metabolizers; this finding has been observed in Japanese, European, and European American populations. In contrast, one study found that the loss-of-activity **CYP2A6**<sup>4</sup> allele was associated with increased lung cancer in a Chinese population; however, a subsequent larger Chinese study found no association between **CYP2A6**<sup>4</sup> and lung cancer, and a small study in a Thai population also failed to find an association between **CYP2A6**<sup>4</sup> and nasopharyngeal cancer. The relevance of these findings to smokers is unclear; the Chinese studies included a significant proportion of never-smokers, whereas the Thai study provided no details on the smoking status of participants.

Earlier studies among European populations also did not find an association between **CYP2A6** and the risk for developing lung cancer, but these studies assessed only one or two variants of low frequency in Europeans (~1%). In terms of histology, Japanese studies have reported an association between variations in **CYP2A6** and squamous and small cell carcinomas, whereas in a European American population a stronger association with adenocarcinoma was noted. Some researchers offer the caveat that genetic association studies in cancer are typically not powered to assess genetic risk within each lung cancer subtype, and they can be further confounded by differences in smoking level, which contribute unequally to the risk of each histological subtype of cancer.

**CYP2A6**-reduced metabolism genotypes are also associated with lower cigarette use and lower nitrosamine exposure. However, even after controlling for cigarette exposure, the association between **CYP2A6** variations and lung cancer risk remains significant, suggesting that genetic differences in the **CYP2A6**-mediated activation of nitrosamines contribute to differing lung cancer risk in addition to **CYP2A6**-mediated influences on smoking behavior.
The involvement of CYP2A6 metabolic activation in carcinogenesis is corroborated by human data showing that the inhibition of CYP2A6 in cigarette smokers is associated with increased routing of NNK to the metabolite NNAL (i.e., evidence of the reduced activation of NNK), and by mouse data demonstrating that the inhibition of the mouse version of CYP2A6 reduces the occurrence of NNK-induced adenomas. Variations in CYP2A6 are also associated with other respiratory tract cancers—oral cancer in North Indians, head and neck cancers in Sri Lankans, and upper aerodigestive cancers in Europeans.

Genetic variations in CYP2A13 are less well characterized, but two different functional variants have been associated with altered lung cancer risk in Chinese and European populations. Variations in another nitrosamine-activating gene, CYP2E1, have been extensively investigated as potential genetic risk factors for lung cancer, but results have largely been equivocal.

Although CYP2A6 and CYP2A13 metabolically activate nitrosamines, UGTs typically detoxify the metabolites of nitrosamines and PAHs into noncarcinogenic glucuronide conjugates, and there is growing evidence that genetic variations in UGTs can influence the risk of lung cancer and other tobacco-related cancers. A UGT2B17 deletion variant with a reduced ability to detoxify nitrosamines was associated with increased lung cancer risk in European American female smokers. Reduced-activity variants in the UGT1A7 gene were associated with an increased risk for lung cancer among Japanese, increased orolaryngeal cancer among European Americans and African Americans, and increased proximal digestive cancers (e.g., esophageal, orolaryngeal, and gastric cancers) among Europeans, presumably because of the reduced ability of these gene variants to detoxify PAHs. However, increased activity UGT1A7 variants were associated with an increased risk, as opposed to a decreased risk, of head and neck cancers, underscoring the need to better characterize the functional and biological consequences of variations in UGT genes.

Variations in UGT1A10, another UGT involved with the detoxification of PAHs, are also associated with the risk for orolaryngeal cancer in African Americans but are unlikely to be detected as a risk factor for cancer in Europeans or Asians due to the lower prevalence of the variants in these other populations (less than 1%). Genetic variations in UGTs could also be an important consideration in biomarker studies. The ratio of NNAL-glucuronide to NNAL (the main metabolites of NNK) is a proposed biomarker for nitrosamine detoxification and for cancer risk; thus, variations in UGTs could confound the interpretation of this biomarker.

Hereditary factors, in addition to diet and cigarette design, are hypothesized to contribute to lower lung cancer incidence in Japanese men compared with American men, despite a higher prevalence of smoking and heavier smoking. Genetic variations that result in reduced metabolic activation of carcinogens could contribute to the lower lung cancer susceptibility. For example, the CYP2A6 deletion allele is significantly more prevalent in Japanese compared with Europeans, European Americans, and African Americans, and cigarette smokers homozygous for the deletion allele have a substantially lower risk of lung cancer compared with CYP2A6-normal metabolizers, with an odds ratio of 0.29 (95% confidence interval 0.15–0.56). Therefore, the high prevalence of the CYP2A6 deletion allele could contribute to the lower average lung cancer risk observed in Japanese smokers. However, genetic variants in CYP1A1 and GSTs, which are associated with increased lung cancer risk, are also more prevalent in Asian versus non-Asian populations. Thus, variations in metabolic genes associated with increased and decreased cancer risk are prevalent in Japanese populations, underscoring the need to assess the concurrent impact of variations in multiple genes.
In African Americans, genetic variations in the glucuronidation detoxification pathway of carcinogens could increase the risk of lung cancer, as a greater proportion of African Americans have a slow glucuronidation phenotype compared with European Americans. On the other hand, a greater proportion of African Americans are CYP2A6-reduced metabolizers compared with populations of European descent, which should confer protection from lung cancer. Thus, the impact of CYP2A6-reduced activity might be opposed by an increased risk conferred by impaired glucuronidation. Finally, among smokers CYP2A6 faster-metabolizing genotypes increase smoking intensity, thus increasing their exposure to carcinogens. This is another mechanism by which CYP2A6 may affect lung cancer risks.

**Nonmetabolic Genetic Risk Factors**

In addition to metabolic genes, genes involved with nicotinic receptor signaling have also been associated with lung cancer risk. GWASs of lung cancer have been conducted. GWASs examine many single nucleotide polymorphisms (SNPs) across the genome of thousands of lung cancer cases and thousands of controls and are useful in identifying SNPs associated with lung cancer and help rule out chance findings. Lung cancer GWASs have found more than two dozen common loci associated with this cancer.

**Chromosome 15p25.1**

GWASs have found a strong association between genetic variations in the chromosome 15q25.1 region and lung cancer. The 15q25.1 region encompasses the nicotinic receptor subunit gene cluster \( CHRNA5-CHRNA3-CHRNB4 \). Variations in this cluster are associated with multiple histological subtypes of lung cancer, including adenocarcinoma, squamous cell carcinoma, and small cell carcinoma and with an earlier age of cancer onset. Genetic association studies of the 15q25.1 region initially were conducted in European and European American populations. The tag SNP (genetic variant) most highly associated with increased lung cancer risk within \( CHRNA5-CHRNA3-CHRNB4 \) occurs at a frequency of 37% in European populations but only 10% in African Americans and ~1% in Asian populations. A two-fold elevated risk is associated with being homozygous for the risk alleles at rs588765-rs16969968 compared to being wildtype at those SNPs.

Subsequent association studies in Japanese, Chinese, and African American populations have also implicated the region in lung cancer risk. Many of the SNPs in the \( CHRNA5-CHRNA3-CHRNB4 \) gene cluster contribute to lung cancer susceptibility in African-Americans and Asian populations as well as European/American white populations. However, there are many other variants of potential importance in the \( CHRNA5-CHRNA3-CHRNB4 \) region among African Americans and Asians.

In addition to lung cancer risk, genetic variations in the 15q25.1 region are associated with a modest increase in cigarette consumption and an increased nitrosamine exposure, potentially increasing the risk of lung cancer through increased carcinogen exposure. However, cancer risk remains elevated after controlling for cigarette smoke exposure, which suggests that genetically determined alterations in nicotinic receptor signaling might contribute directly to the risk of lung cancer as well as indirectly through increased cigarette consumption. The \( CHRNA5-CHRNA3-CHRNB4 \) subunits are expressed in the lung and participate in nicotinic receptors with a high affinity for nitrosamines in addition to nicotine and the endogenous ligand acetylcholine. Nitrosamines could foster the development and progression of cancer by binding to nicotinic receptors, thereby disrupting the balance...
between inhibitory and stimulatory receptor signaling and resulting in increased cell proliferation, invasion, angiogenesis, and reduced apoptosis. \(^{212}\)

GWASs found that variations in the chromosome 15q25.1 region, but not specifically the \(CHRNA5-CHRNA3-CHRNB4\) subunits, were implicated in lung cancer risk. Thus, other genes in this region, such as \(IREB2\) (iron-sensing response element) and \(PSMA4\) (proteasome \(\alpha\) 4 subunit isoform 1), could contribute to lung cancer risk in addition to or instead of the nicotinic receptor subunit genes. \(^{107,277−279,290}\)

Because the \(CHRNA5-CHRNA3-CHRNB4\) region has been associated with both tobacco smoking behaviors and lung cancer, investigators have examined whether the region influences smoking behaviors and is thus associated with lung cancer, or whether the region is associated with lung cancer independent of its effect on smoking behaviors. Evidence on this question among African Americans is inconsistent. \(^{112,276,284}\)

**Chromosome 5p15.33 and Other Loci**

Another region on chromosome 5, specifically 5p15.33, is near the gene telomerase reverse transcriptase (TERT) and is associated with lung cancer. \(^{291}\) Among African Americans, a variant located on 5p15.33, specifically at locus rs2853677, is associated with adenocarcinoma of the lung, but not other lung cancer types, with an odds ratio of 1.32 (95% CI 1.20–1.44).\(^{276}\) confirming previous studies of this SNP and lung adenocarcinoma in African Americans. \(^{276,288}\) Variants in \(CHRNA5\) (rs 2036527) are associated with lung cancer risk and smoking frequency in African Americans. \(^{276,288}\) In African Americans there are other loci in the nicotinic cholinergic receptor genes and in others, such as TERT. \(^{286,288,292}\) In studies of nonsmoking Asian women, associations were observed with multiple variants that are associated with longer telomere length and loci for lung cancer \(^{293}\) or lung adenocarcinoma \(^{294}\) that are different from those found in other populations. GWASs of lung cancer among Asians have identified a number of SNPs of importance to this population group. \(^{293−297}\)

Other common genes associated with lung cancer based on GWASs include \(PRPF6\) \(^{298}\) and \(NEXN-AS1\). \(^{299}\) In addition to common genes, rarer genes such as \(BRCA2\) and \(CHEK2\), are also associated with lung cancer risk. \(^{300}\)

Several other developments have provided further insights about the genetics of lung cancer. The GWAS approaches used to identify loci for specific cancers have been extended to examine whether some SNPs are associated with multiple cancers including lung cancer. For example, both lung cancer and breast cancer are associated with variants at 1q22. \(^{301}\) Risk prediction models that include tobacco smoking behaviors as well as genetic factors are being developed to help identify subgroups of people who are at greatest risk of lung cancer. For example, lung cancer risk prediction models have been developed, but the inclusion of top genetic hits in the model have not improved its utility for African Americans. \(^{302}\) And finally, research studies are exploring the functional implications of various risk alleles. \(^{303−305}\)

**Summary for Lung Cancer**

Genetic factors, along with smoking behaviors and socioeconomic and environmental influences, contribute to population disparities in the risk for developing tobacco-related cancers. This section introduced the general mechanisms by which genetic factors influence cancer risk among smokers, discussed lung cancer risk specifically, and provided key examples of genetic risk variants.
Variations in the genes involved with carcinogen processing and carcinogen signaling pathways modulate lung cancer risk. Variations in the CYP2A6 gene, which metabolically activates nitrosamines, are an important genetic risk factor for lung cancer, particularly in Asian populations in which the frequency of CYP2A6-slow metabolism is high. The evidence for other drug-metabolizing enzymes, such as CYP1A1 and GSTs, which metabolize PAHs, is more equivocal, possibly owing to fewer characterized functional variants. The association between CYP2A6 and cancer in European populations only became significant as more prevalent functional genetic variants were identified and evaluated. The association between lung cancer risk and variations in CYP2A6 has two potential mechanisms—altered carcinogen exposure through smoking behaviors and altered metabolic activation of nitrosamines.

The nicotinic receptor subunit genes CHRNA5-CHRNA3-CHRNB4 have also emerged as significant candidate genetic risk factors for lung cancer across multiple racial/ethnic populations. Recent GWASs have found loci in African Americans and nonsmoking Asian women that were previously identified in European populations; these studies have also found loci unique to those populations. Variations in these genes could influence lung cancer risk by modulating smoking behaviors by influencing the degree to which nitrosamines and/or nicotine interfere with endogenous signaling pathways. A number of other loci have emerged as important in lung cancer risk, but limited evidence about them in populations other than people of European origin is available. Some genetic loci are more important for adenocarcinoma than for other histologic forms of lung cancer. Until the specific cancer-causing genetic variants are determined and the frequencies of these variants (as opposed to tag SNPs) are assessed in each racial/ethnic population, it is difficult to determine whether variations in nicotinic receptor subunits could be contributing to population differences in lung cancer risk.

As more genetic variants are identified and characterized within different genes and among diverse populations and are investigated in the context of smoking behaviors, a clearer picture of disparities in the genetic risk for smoking-related cancers should emerge.

**Genetics and TRHD: Current Knowledge and Future Directions**

Genetic factors can contribute to tobacco use and its consequences by affecting:

- Use of tobacco including amounts used, tobacco dependence, age of onset, and reduced cessation success, and
- Risk of lung and other cancers.

Some major classes of genes that have been found through candidate gene and GWASs and affect use of tobacco are involved in:

- Nicotine metabolism, especially CYP2A6
- The dopaminergic system, especially relating to smoking initiation
- The nicotinic cholinergic system, especially the CHRNA5-CHRNA3-CHRNB4 gene cluster, related to tobacco quantity and dependence as well as cessation.

Many of the earlier types of studies involved selecting genes thought to be involved in a pathway and studying that set (i.e., candidate gene approach). Technological advances enabled GWASs that compare loci (specifically SNPs) across hundreds of thousands of loci in groups of people with particular diseases
or characteristics and those without them to identify SNP alleles that differ in the compared groups. These studies need to be very large, involving thousands of people, to help rule out chance as a reason for differences, given the very large number of statistical comparisons being made. Many loci have been identified through this approach. However, most of the evidence has come from populations of European origin. The relatively few studies of African American populations and studies of Asian women who did not smoke find both similarities and differences in the genetic regions involved.

As of this writing (2017), the contribution of genetic factors to TRHD, specifically tobacco-related cancers, cannot be estimated precisely because of the insufficient evidence available on this subject particularly about non-European-origin populations. Progress has been made in identifying and characterizing specific genetic variants that influence stages along the tobacco use continuum as well as in estimating the risk of developing tobacco-related cancers, as detailed in this section, but the inventory of genetic risk factors is far from complete for different population groups. In addition, these genetic studies have been undertaken primarily to determine disease etiology rather than to address TRHD; therefore, these studies have typically been performed within a restricted range of sociodemographic groups, which has limited the ability to translate findings to TRHD. Also, genetic investigations have favored methodological approaches that circumvent the heterogeneity of smoking behaviors and cancer risk in order to be able to detect genetic signals—an approach that minimizes the complexity of interactions between genes and environmental factors that lead to disease as opposed to investigating these interactions.

As of 2017, genetic factors do not readily explain TRHD because genetic variants account for a modest proportion of the heritable variance in smoking behaviors, and because of the scarcity of genetic studies in relevant populations such as racial/ethnic and LGBT groups. Large, replicated studies conducted primarily with populations of European origin have clearly established genetic susceptibility loci as involved in tobacco dependence, quantity of use, cessation, and lung cancer risk, both tobacco-related and not tobacco-related. These genetic factors have a modest effect on risk of tobacco behaviors and cancer risk, yet help elucidate the biological mechanisms underlying tobacco behaviors and lung cancer risk and may have important implications for stratifying groups for interventions, such as cessation treatments, that may be most effective for them.

Some large-scale studies of African American populations and studies of Asian women who did not smoke find both similarities and differences compared to populations of European origin in the genetic regions involved in both smoking behaviors and lung cancer risk. Similarities include the importance of loci in CYP2A6 involved in nicotine metabolism and the nicotinic cholinergic system, especially the CHRNA5-CHRNA3-CHRNB4 gene cluster. The latter is related to tobacco quantity and dependence as well as cessation and may also be involved in lung cancer risk independent of its role in tobacco use behaviors. However, African Americans and Asians each have other distinct loci, both in those and other genetic regions. These genetic factors are only part of a complex web of behavioral, biological, environmental, social, and other characteristics that contribute to TRHD.

The state of knowledge about the genetic risk for smoking behaviors and lung cancer continues to be characterized by significant gaps and a need for further research. To more fully understand TRHD, more research is needed, as described below.
More Large Studies in a Wider Range of Populations

Data on the numerous genetic regions and genes that have been investigated as potential risk factors in the etiology of lung cancer come primarily from Asian, European, and European American populations; additional genetic association studies are needed for minority populations.\textsuperscript{308}

Many of the studies of African Americans have achieved the large sample sizes needed by pooling data from multiple studies. The associations observed in a number of studies of nonsmoking women in Asia, specifically China,\textsuperscript{293,294} may or may not apply to women in the United States. It would be very difficult to obtain an adequate sample size of Asian Americans for GWASs. There are very few studies in Hispanic populations or in populations defined by sexual orientation. It is important to conduct studies in other ethnic groups because they account for the majority of the human population (Asians) and genetic variation (Africans). There are also few studies in these populations that have examined in detail smoking phenotypes such as age of initiation and cessation. And most studies that break ground in new areas, such as the functional implications of these variants, are done in populations of European origin first because the large number of study participants needed are easier to assemble, and studies in other populations come much later.

Some ways to help achieve the large sample sizes needed for assessing genetic risks in these relevant populations include fostering pooling data whenever possible. Although pooling projects have been successfully conducted and are critical to research progress, racial and ethnic minority individuals are still significantly under-represented in existing human population studies. Large-scale cohorts and case-control studies with good representations by race/ethnicity and SES are necessary to further our understanding of racial/ethnic disparities in the risk for lung cancer. One such cohort, described in the Southern Community Cohort Study, offers the potential to assess genetic risk factors among African Americans and European Americans of similar socioeconomic backgrounds while incorporating biomarkers of cigarette exposure and accounting for environmental factors such as menthol smoking.\textsuperscript{309}

Furthermore, to understand underlying risk differences among populations and the specific role of genetic factors, analyses should address interacting and interrelated environmental factors, such as SES, education level, diet, smoking behaviors, and other carcinogen exposures.\textsuperscript{310} Most importantly, a concrete understanding of the prevalence, amount, and intensity of smoking within each high-risk group (e.g., race/ethnicity, SES, sexual orientation) is necessary because smoking is the predominant risk factor for lung cancer; thus all other risk factors (genetic or environmental) need to be understood in the context of smoking. However, it will be challenging to obtain the large sample sizes needed for such studies.

By furthering our understanding of the biological basis of smoking behaviors and by disaggregating high-risk populations by interindividual risk, genetic factors will help guide novel treatments and help tailor intervention strategies. Although there is a small body of research that suggests that having certain genotypes may influence smoking cessation success with different cessation treatment approaches, this has not been examined within different population groups.

Most of the effects of individual gene variants on lung cancer risk found to date have been modest. However, it is possible that risk models incorporating genetic variations across multiple genes could capture more of the variance in lung cancer susceptibility than is captured by any of these factors acting alone. For example, a greater proportion of lung cancer risk could be accounted for by the combined effect of: (1) two metabolic-activating enzymes, \textit{CYP1A1} and \textit{MPO}; or (2) a metabolic-activating...
enzyme, CYP1A1, and a metabolic-inactivating enzyme, GSTM1; or (3) a metabolic-activating enzyme, CYP2A6, and nicotinic receptors, CHRNA5-CHRNA3-CHRNB4.

**Lung Cancer Heterogeneity**

The heterogeneity of lung cancer might also confound investigations. Lung cancer is a disease with multiple histological subtypes, each having a different relationship with smoking behaviors. For example, adenocarcinoma is currently the most prevalent lung cancer type among smokers, but adenocarcinoma shows a more modest association with cigarettes smoked per day and with years of smoking relative to other subtypes. As such, the smoking patterns of a population under study and the proportion of a given histological lung cancer subtype could influence a researcher’s ability to detect genetic and nongenetic associations, and future research on TRHD would benefit from separate studies of cancer subtypes or from correcting for cancer subtypes by using procedures analogous to those employed to correct for population stratification. Fortunately, investigators are increasingly publishing genetic data associations for specific lung cancer histologic types and finding that some susceptibility loci are specific for certain histologic types.

Although lung cancer has been the focus of this chapter, a better understanding of genetic contributions to other cancers would be valuable. Studies on never-smokers could also shed light on understanding disparities in racial/ethnic, sex, and LGBT groups.

**Gene–Environment Interactions**

The importance of environmental factors is highlighted by migration studies, which find that the cancer risks of generations born after the original immigration resemble the risks found in the adopted country rather than the ancestral land.

In addition to combined gene risk models, further progress in understanding the genetic factors associated with cancer risk requires a diligent assessment of gene–environment interactions. Such studies are challenging to conduct because they require large sample sizes. The degree of interaction between variations in metabolic genes and cigarette smoke exposure typically varies with the level of exposure. For instance, at lower versus higher levels of cigarette exposure, variations in genes such as CYP2A6, CYP2A13, CYP1A1, GSTM1, and MPO are more prominently associated with lung cancer risk. In contrast, the association of variations in the CHRNA5-CHRNA3-CHRNB4 subunit genes with lung cancer changes little with the level of cigarette smoke exposure. Future genetic association studies employing biomarkers of smoking dose and quantitative records of historical cigarette exposure will help characterize gene–environment interactions and their relationship to lung cancer risk. In addition to smoking dose, environmental factors that are unique to high-risk groups such as LGBT and lower socioeconomic groups must be investigated.

In the future, as more genetic association studies employ analytical approaches that incorporate and evaluate gene–environment interactions, population heterogeneity, biomarkers, and more precise behavioral measures, more of the genetic variance in smoking behaviors will be explained. Multigene models that assess the combined effect of multiple genetic variants on smoking characteristics may be more informative than studying genetic factors one at a time.
Biomarkers

Biomarkers of smoking dose could offer insight into the paradox of worse health outcomes among African American smokers despite their greater proportion of light smoking. Differences in lung cancer risk between African Americans, Native Hawaiians, and other racial/ethnic groups are greatest at lower levels of cigarette consumption (e.g., no more than 10 cigarettes per day).\(^2\) However, there is emerging evidence that African Americans smoke more intensively at lower levels of daily cigarette consumption than other racial/ethnic groups.\(^{96,316}\) The higher rates of lung cancer could be explained by greater carcinogen exposure despite consuming fewer cigarettes daily, coupled with the fact that a greater proportion of African Americans have slower carcinogen detoxification capabilities.

In addition to amassing genetic data on the relevant populations of interest, it is critical to incorporate better biomarkers and surrogates for cigarette smoke exposure. As smoking is the predominant risk factor for lung cancer, smoking biomarkers must be properly validated among different groups of interest and incorporated and/or controlled for to investigate the contribution of other risk factors, genetic or nongenetic, to TRHD. For example, African American smokers would be predicted to have lower cigarette exposure compared with European American smokers based on self-reported number of cigarettes smoked; however, using such biomarkers as total nicotine equivalents, it becomes apparent that African American smokers can achieve comparable levels of cigarette exposure despite smoking fewer cigarettes per day.\(^{96,316}\) The fact that this amount of cigarette exposure can be achieved despite smoking fewer cigarettes suggests differences in smoking topography. Topography could also interact with genetic factors to influence disease risk, given that the depth of inhalation, an aspect of topography, is believed to contribute to disease risk.\(^{317}\) Thus, without properly accounting for cigarette exposure, relevant genetic and nongenetic factors could be obscured by differing self-reported smoking behaviors among populations.

Expanded Genetic Characterization and Other High-Throughput Characterization Approaches (“-Omic”)

Cigarette smoke exposure is causally associated with numerous cancers and is the leading risk factor for lung cancer.\(^{188-191}\) Cigarette smoke contains more than 69 carcinogens, among which PAHs and nitrosamines have the strongest causal association with lung carcinogenesis.\(^{207}\) Variations in the genes that influence smoking behaviors might indirectly influence lung cancer risk by altering carcinogen exposure (i.e., tobacco intake). Variations in genes such as DNA repair genes can also increase individuals’ underlying susceptibility to the damage caused by carcinogens.

Increasingly, more sophisticated data is being generated: exome sequencing, gene expression, epigenetic data from shared online databases and consortia. It will be important to study the role of epigenetic changes and gene expression in addition to the genetic changes. Also, functional studies are helping unravel the basis for the effect of the genes involved. It will be very important that these studies are performed in multiple population groups to better understand TRHD. Epigenetics is another promising research avenue to better understand the environmental component of gene–environmental interactions influencing TRHD as both smoking and social environmental stressors appear to influence epigenetic patterns.\(^{318-320}\) These and many other examples of -omics and biomarkers approaches are shown in Figure 3.6.
Chapter 3: Genetics, Physiological Processes, and Tobacco-Related Health Disparities

Figure 3.6 Types of Biomarkers and -Omis Technologies That Could Help Understanding of TRHD

Complex Interrelationships

It is important to understand the impact of all other environmental and host factors of smoking. Sexual orientation is important and understudied. Poverty and many associated exposures (diet, obesity, work/home environment, geographic location), pollution, health care access, education, and many other factors contribute as well. Comorbidities, such as chronic obstructive pulmonary disease and immune-related conditions such as HIV, and family history of cancers and other conditions are also important.

The challenges in applying genetic findings from one population to another, even when dealing with a causative variant as opposed to a tag SNP, are numerous. There is also the issue of the potential contribution of correlated and as yet poorly defined risk factors. For example, race/ethnicity, which genetic researchers typically use to represent geographic/ancestral origin and genomic variability, also encompasses correlated environmental, economic, and sociocultural factors that can influence disease risk. Likewise, the effect of a gene variant on a particular outcome (e.g., smoking persistence) is likely influenced by multiple interacting genetic and environmental components that could differ across populations. Similarly, other populations experiencing TRHD, such as lower SES groups and LGBT groups, may experience environmental risk factors that could interact with genetic risk factors. Thus, without a better understanding of genetic and nongenetic risk factors and their potential interactions in the manifestation of tobacco-related diseases, it is not straightforward to extrapolate genetic findings...
from one population to another, even when the impact of a genetic variant on the function of a gene is known.

As our knowledge of genetic and nongenetic factors and their interactions with each other and with smoking behaviors increases, a clearer picture of TRHD will emerge. In addition, as key genetic and nongenetic risk factors become uncoupled from race/ethnicity, SES, and LGBT status, it should become feasible to predict whether a particular individual will suffer disproportionately from TRHD without having to rely on these demographic categories.
Chapter 3: Genetics, Physiological Processes, and Tobacco-Related Health Disparities

References

112


Chapter 3: Genetics, Physiological Processes, and Tobacco-Related Health Disparities


