

Genetics of Smoking Behaviour

Jennifer J. Ware and Marcus R. Munafò

Abstract It is now well-established that smoking-related behaviours are under a substantial degree of genetic influence. Efforts are now focused on identifying the specific genetic variants which underlie these behaviours. Within this chapter, we introduce a variety of established and emerging methods employed to identify such variants, ranging from candidate gene to whole genome sequencing approaches, and highlight what these techniques have taught us about the genetic architecture of smoking-related behaviours. Further, we discuss how phenotype refinement has developed our understanding of these relationships, affording us insight into the specific mechanisms linking genetic variants to smoking-related behaviours.

Keywords Genetics • Smoking • Heritability • Candidate gene • GWAS • Sequencing • Phenotype refinement • Recall-by-genotype

Contents

1	Introduction	20
2	Genome-Wide Association Studies	21
3	Phenotype Refinement	26
3.1	In Vitro Studies and Animal Models	26
3.2	Candidate Gene Follow-up Studies	28
3.3	Recall-by-Genotype Studies	30
4	New and Emerging Technologies	31
5	Conclusions and Future Directions	32
	References	33

M.R. Munafò
MRC Integrative Epidemiology Unit, UK Centre for Tobacco and Alcohol Studies,
and School of Experimental Psychology, University of Bristol, Bristol BS8 1TU, UK

J.J. Ware (✉)
MRC Integrative Epidemiology Unit and School of Social and Community Medicine,
University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK
e-mail: jen.ware@bristol.ac.uk

1 Introduction

Behaviour genetics has historically relied on the use of twin, family and adoption studies to determine the proportion of variance in a phenotype that can be attributed to genetic influences. The cornerstone of this approach is the ‘natural experiment’ of twinning—approximately 1 in 85 live births are twin births, and of these approximately one-third will be identical (monozygotic or MZ) twins and the remaining two-thirds non-identical or fraternal (dizygotic or DZ) twins. Since MZ twins have identical genotypes, whereas DZ twins share on average only 50 % of their variegated genotype, and yet in both cases twins will be raised in (presumably) identical environments, the logic of twin studies is that if a behavioural trait is more similar in pairs of MZ twins than it is in pairs of DZ twins, then that trait must presumably be under a degree of genetic influence. Family and adoption studies provide other means by which natural variation in genotypic similarity and environmental similarity can be associated with phenotypic variability to calculate the proportion of variation in the phenotype that can be accounted for by variation in genetic and environmental influences. The proportion of variation in phenotype that is due to variation in genotype is expressed as the *heritability* of a trait (h^2)—a heritability coefficient of 0.50 means that 50 % of the variation in that trait is due to genotypic variation.

Early evidence suggested that various aspects of smoking behaviour are under partial genetic influence; concordance rates are higher in MZ compared to DZ twins for persistent smoking and for successful cessation (Carmelli et al. 1992). Further evidence has supported this, suggesting a substantial heritability of smoking initiation (Heath et al. 1993), with this effect being partly independent of genetic effects on persistent smoking behaviour (Heath and Martin 1993), which itself shows a similar heritability. Twin studies have also allowed the dissection of different tobacco use phenotypes, using a conditional approach to understand the shared and unique genetic influences on initiation of tobacco use, progression to regular use, and development of dependence. This has indicated significant overlap in the contribution of genetic factors to these distinct phenotypes, but also important unique contributions, particularly to progression to regular tobacco use and the development of dependence (Maes et al. 2004). Similar approaches have enabled the investigation of shared and unique genetic influences on tobacco use by different routes of administration (e.g., cigarette smoking versus smokeless tobacco). This indicates that genetic influences may indeed vary by route of administration (Schmitt et al. 2005). More specific tobacco use phenotypes have also been investigated using these methods, such as initial reactions to tobacco use (both positive and negative), which also appear to be under a moderate degree of genetic influence (Agrawal et al. 2014). This is of particular interest given the evidence that initial reactions to tobacco use may predict the subsequent development of regular use and dependence (Sartor et al. 2010). Nevertheless, despite the valuable insights provided by twin, family and adoption studies, they are unable to identify the specific genetic variants that influence tobacco use.

However, since the mid-1990s it has been possible to directly genotype specific variants, and therefore look for association between measured genetic variation and phenotypic variation. Genetic association studies typically use a case-control methodology to compare the prevalence of genetic variants in these two unrelated groups (e.g., tobacco dependent and non-dependent individuals). Genetic association studies typically investigate candidate genes selected on the basis of a priori evidence of relevance, based on the known or presumed neurobiology of the phenotype of interest. For example, given the acknowledged role of dopamine in addictive behaviours, candidate genes related to variation in the dopamine pathway received considerable interest (Munafò et al. 2004). However, despite over a decade of candidate gene studies, very few associations proved reproducible. It now seems that this was principally because the effects of common genetic variants on complex behavioural phenotypes, such as smoking behaviour, are very small (typically $<0.1\%$ phenotypic variance). These studies were therefore simply underpowered to reliably detect these effects.

Replication of candidate gene studies has been uneven, at best. This may reflect the fact that initial findings were false positives (which therefore subsequently failed to replicate), or may instead reflect differences in phenotype definition across studies. Certainly some candidate gene findings have proven to be robust (notably those relating to genes encoding nicotine-metabolising enzymes, such as *CYP2A6* see chapter entitled [Pharmacogenetics of Nicotine and Associated Smoking Behaviors](#); this volume). However, there is growing acceptance that most have not, and it remains difficult to discern from the published candidate gene literature which findings are reliable (given differences in study design and phenotypic definition, as well as uneven statistical evidence). However, in the early 2000s, candidate gene methods were rapidly supplanted by genome-wide association studies (GWAS).

2 Genome-Wide Association Studies

Genome-wide association studies adopt the same approach to sample selection and analysis as candidate gene studies (e.g., comparison of allelic frequencies between groups selected on the basis of phenotype). However, instead of focusing on a pre-specified locus, microarrays (“gene chips”) are used to systematically genotype hundreds of thousands of single nucleotide polymorphisms (SNPs) across the entire genome. As such, and in contrast to candidate gene studies, they are strictly agnostic in their approach, requiring no a priori hypothesis. The multiple testing burden implicit in GWAS led to consensus that signals had to achieve a statistical significance threshold corrected for this (typically $p < 5 \times 10^{-8}$), which in turn necessitated the pooling of data across multiple studies to achieve the necessary sample size to detect the small effects associated with the common genetic variants captured by GWAS. As a direct consequence of these large-scale efforts, GWAS have been extremely successful in identifying genetic variants associated with a range of complex phenotypes, aiding identification of variants that would perhaps not have been considered previously on the basis of biological function.

The most robust finding to emerge from GWAS of smoking phenotypes is the association between a variant within the nicotinic acetylcholine receptor gene cluster *CHRNA5-A3-B4* on chromosome 15 and smoking quantity. An association between rs16969968 in *CHRNA5* and nicotine dependence (ND) was first reported in 2007 in a candidate gene study (Saccone et al. 2007), with the minor allele found to confer increased risk. The following year, the same locus (tagged by rs1051730 in *CHRNA3*, a variant highly correlated with rs16969968) was found to be associated with smoking quantity in a GWAS conducted by Thorgeirsson et al. (2008). This study also demonstrated an association between rs1051730 and nicotine dependence and two smoking-related diseases, lung cancer and peripheral arterial disease. Notably, whilst the candidate gene study was published first, it was the GWAS that made much more of an impact. This may have been because *CHRNA5* was not recognised as a particularly strong candidate at the time, given the then known neurobiology of tobacco dependence—more emphasis had been placed on genes encoding α_4 and β_2 subunits, based on evidence from animal models that these mediate the reinforcing properties of nicotine (e.g., Picciotto et al. 1998). Furthermore, the simultaneous demonstration of an association between this locus and two smoking-related diseases (i.e., lung cancer and peripheral arterial disease) lent further weight to this finding. The association between this locus and smoking quantity has since been replicated in several consequent independent studies (David et al. 2012; Liu et al. 2010; Thorgeirsson et al. 2010; Tobacco and Genetics Consortium 2010).

Early GWAS of smoking behaviours typically examined associations with directly genotyped SNPs only, which provide relatively sparse coverage of the genome, and include only common SNPs with a minor allele frequency (MAF) greater than 5 % (Caporaso et al. 2009; Thorgeirsson et al. 2008). In contrast, GWAS now commonly employ imputation to expand genomic coverage. Imputation allows us to predict ('impute') genotypes that have not been directly genotyped on a microarray. In order to do so, microarray data is matched to a genome reference panel, which consists of densely genotyped (or sequenced) genomic data from multiple individuals (see Fig. 1). The genomic reference panels typically used for imputation are HapMap (<http://hapmap.ncbi.nlm.nih.gov/>) and 1000 Genomes (<http://www.1000genomes.org/>). The most recent release of 1000 Genomes (Phase I) provides a much larger set of SNPs (~ 37 million) relative to HapMap (~ 4 million). Imputation is used in GWAS for a number of reasons. First, it can increase study power (Marchini and Howie 2010). Second, it provides us with a more detailed view of particular genomic regions which contain many associated SNPs. This fine-mapping approach may allow us to identify actual causal variants not captured by microarrays. Finally, imputation allows us to combine GWAS results across studies that have used different GWAS chips, thus enabling meta-analysis (Marchini and Howie 2010). Imputation has been used to great success in recent GWAS of smoking phenotypes, enabling the running of very large ($n > 75,000$), consortia-based GWAS meta-analyses, which have identified a number of new genomic regions (Liu et al. 2010; Thorgeirsson et al. 2010; Tobacco and Genetics Consortium 2010). Further, it has enabled fine-mapping of the 15q25 locus (Liu et al. 2010). These results will be discussed in turn.

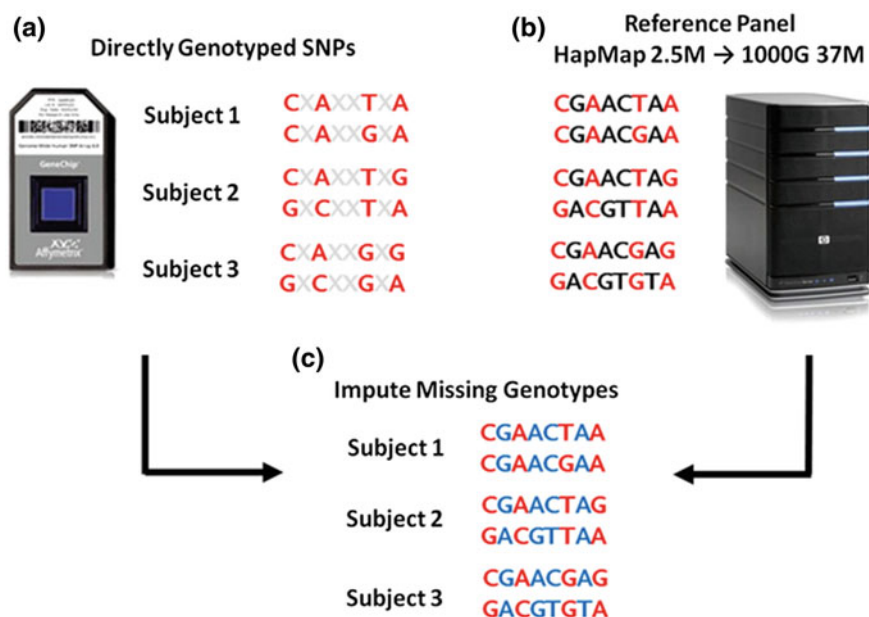


Fig. 1 Schematic overview of imputation. Imputation allows us to predict ('impute') genotypes which have not been directly genotyped on a microarray. In order to do so, microarray data is matched to a genome reference panel, which consists of densely genotyped (or sequenced) genomic data from multiple individuals. The genomic reference panels typically used for imputation are HapMap and 1000 Genomes. The most recent release of 1000 Genomes (phase I) provides a much larger set of SNPs (~37 million) relative to HapMap. This affords increased resolution of genomic regions. Imputation error increases as minor allele frequency decreases, as the imputation is based on an algorithm derived from a much smaller number of individuals (Marchini and Howie 2010). Accurate imputation of low-frequency SNPs requires genomic data from large numbers of individuals. HapMap 2 contains haplotype information on just 120 Europeans. However, HapMap 3 is based on data from 330 Europeans. Similarly, 1000 Genomes phase I includes far more haplotypes than 1000 Genomes pilot I, thus allowing for more accurate imputation of low frequency or rare SNPs. If rare variation is key to a particular trait, then 1000 Genomes Phase 1 (and later releases) may help identify these variants. Figure reproduced from Wood et al. (2013)

Consortia-based GWAS meta-analyses (typically imputed to HapMap 2) have identified a number of novel genomic regions associated with a range of smoking phenotypes, including smoking initiation and smoking cessation, as well as smoking quantity. Specifically, a nonsynonymous SNP (rs6265) in the brain-derived neurotrophic factor (*BDNF*) gene on chromosome 11 was found to be associated with smoking initiation. *BDNF* is involved in regulating synaptic plasticity and survival of cholinergic and dopaminergic neurons (Thorgeirsson et al. 2010; Tobacco and Genetics Consortium 2010). A variant near dopamine beta hydroxylase (*DBH*) gene on chromosome 9 was also identified in relation to smoking cessation (Thorgeirsson et al. 2010; Tobacco and Genetics Consortium 2010). *DBH* is an enzyme involved in the conversion of dopamine to noradrenaline. In relation to smoking quantity, genomic regions on chromosomes 8, 10 and 19

Table 1 Genomic loci identified through GWAS of smoking phenotypes

Author	Year	Phenotype	Chr	Gene	SNP	Notes
Thorgeirsson	2008	Smoking quantity	15	<i>CHRNA3</i>	rs1051730	
Liu	2009	Smoking status	4	near <i>IL15</i>	rs4956302	Marginal evidence for association ($p = 8.8 \times 10^{-8}$) in males only
Furberg	2010	Smoking quantity	15	<i>CHRNA3</i>	rs1051730	Primary signal at this locus
		Smoking quantity	15	<i>CHRNA5</i>	rs684513	Independent signal (after conditioning on rs1051730)
		Smoking quantity	10	<i>LOC100188947</i>	rs1329650	
		Smoking quantity	10	<i>LOC100188947</i>	rs1028936	
		Smoking quantity	19	<i>EGLN2</i>	rs3733829	
		Initiation	11	<i>BDNF</i>	rs6265	
		Cessation	9	<i>DBH</i>	rs3025343	
Thorgeirsson	2010	Smoking quantity	15	<i>CHRNA3</i>	rs1051730	Primary signal at this locus
		Smoking quantity	15	near <i>IREB2</i>	rs2869046	Independent signal (after conditioning on rs1051730)
		Smoking quantity	15	<i>AGPHD1</i>	rs2036534	Independent signal (after conditioning on rs1051730)
		Smoking quantity	19	<i>CYP2A6</i>	rs4105144	
		Smoking quantity	8	<i>CHRNA3</i>	rs6474412	SNP in shared LD block with <i>CHRNA6</i>
Liu	2010	Smoking quantity	15	<i>CHRNA3</i>	rs1051730	Original GWAS using HapMap (release 22)
		Smoking quantity	15	<i>CHRNA5</i>	rs55853698	1000 Genomes (Pilot 1). SNP is in high LD with rs1051730
		Smoking quantity	15	<i>CHRNA3</i>	rs6495308	Independent signal (after conditioning on rs55853698)
David	2012	Smoking quantity	15	near <i>CHRNA5</i>	rs2036527	African American sample

were identified, in addition to the 15q25 locus (Thorgerirsson et al. 2010; Tobacco and Genetics Consortium 2010). In particular, two genes with clear biological relevance were identified—two nicotinic receptor sub-unit genes (*CHRNA3/CHRNA6*) on chromosome 8 and *CYP2A6* on chromosome 19 (Thorgerirsson et al. 2010). The *CYP2A6* enzyme is principally responsible for the metabolism of nicotine (see chapter entitled [Pharmacogenetics of Nicotine and Associated Smoking Behaviors](#); this volume). However, the proportion of phenotypic variance explained by these SNPs is far less than the variance explained by rs16969968-rs1051730 at 15q25, which explains $\sim 1\%$ of the variation in smoking quantity.

High-density imputation has afforded additional benefits; for example, it has enabled high-resolution examination of the 15q25 region in relation to smoking quantity. Liu et al. (2010) used 1000 Genomes Pilot 1 data to fine-map the 15q25 region for association with smoking quantity. This allowed for analysis of virtually all common SNPs ($MAF > 5\%$) in the region, and offered a fivefold increase in marker density compared to HapMap 2. Using this imputation approach, combined with meta-analysis, Liu et al. (2010) identified rs55853698 as the variant with strongest evidence for association with smoking quantity at this locus. This variant, which is absent from HapMap 2, is located within a promoter region of *CHRNA5*, and is a plausible candidate for affecting mRNA transcription. However, this variant is also in very high LD with rs16969968. Conditioning on rs55853698 revealed a second independent signal at rs6495308, located within *CHRNA3*. Conditioning on both SNPs left no residual signal, suggesting that these two variants could explain the full signal at 15q25.1 in relation to smoking quantity (Liu et al. 2010).

In summary, GWAS has dramatically advanced our understanding of the genetic underpinnings of smoking initiation, cessation, and smoking quantity (see Table 1). Despite these successes, however, the total proportion of variance explained by all variants identified through GWAS to date is far less than the heritability estimates indicated by earlier twin studies. It is possible that rare variants ($MAF < 1\%$), which cannot be very accurately imputed using current genomic reference panels (given the limited numbers of haplotypes on which they are based) may account for this ‘missing heritability’. This is discussed in more detail in Text Box 1.

Text Box 1: The Missing Heritability Problem Genome-wide association studies (GWAS) have been extremely successful in identifying genetic variants associated with a range of complex phenotypes. As we discuss, several loci associated with various tobacco use phenotypes have been identified, through large consortium-based efforts (Tobacco and Genetics Consortium 2010). This is in stark contrast to the candidate gene literature, where few reliable signals emerged after almost two decades of effort. Despite this, variants identified to date via GWAS explain less than half the heritability of complex phenotypes such as smoking behaviour estimated by twin and family studies. This is the so-called “missing heritability” problem (Manolio et al. 2009).

It is now clear that some missing heritability will be accounted for by variants that have not yet been identified via GWAS (Yang et al. 2010). This is in part

because most common variant chips have relatively poor coverage in the minor allele frequency (MAF) spectrum below 5 %. Evolutionary theory predicts that mutations strongly affecting complex phenotypes will tend to occur at low allele frequencies (Visscher et al. 2012). There is, therefore, growing interest in the potential role of low frequency and rare variants (i.e., MAF <5 and <1 %, respectively). In other words, “missing heritability” is likely to be due to two factors (Visscher et al. 2012): (i) unidentified common variants with very small effects that cannot be detected by current GWAS using existing sample sizes, and (ii) rare variants not in sufficient linkage disequilibrium with variants on currently available genotyping chips and therefore undetectable via GWAS.

3 Phenotype Refinement

As we have discussed, GWAS approaches have proven highly successful in identifying genetic variants that are associated with a variety of tobacco use phenotypes (Thorgerirsson et al. 2010; Tobacco and Genetics Consortium 2010). For example, these approaches have confirmed the importance of the nicotinic receptor subunit gene cluster *CHRNA5-A3-B4* in smoking quantity. Identification of novel genes implicated in tobacco use through genome-wide approaches (e.g., *CHRNA5*) has led to renewed interest in these genes and their products. However, these studies are limited in two important ways. First, they rely on self-report measures of smoking behaviour, often ascertained retrospectively and in different ways across different studies. This combination of self-report, retrospective recall, and the need for phenotype harmonisation, may introduce considerable measurement error. Second, these studies have not taught us what the fundamental mechanisms linking these genes to these behaviours are. Phenotype refinement provides a means by which to understand the specific mechanisms linking genetic variants to disease. Carefully designed and well-characterised phenotypes also offer greater measurement precision, in principle providing a ‘cleaner’ genetic signal, and improving the likelihood of effect replication. In this section, we discuss how phenotype refinement has advanced our understanding of these relationships.

3.1 *In Vitro Studies and Animal Models*

The rs16969968 variant in *CHRNA5* is non-synonymous, resulting in an amino acid change (aspartate to asparagine) in the resultant α_5 nicotinic receptor subunit protein. Research has shown that this variant is functional. In vitro studies have demonstrated that nicotinic receptor complexes containing the α_5 receptor subunit featuring the aspartic acid variant exhibit a twofold greater maximal response to a nicotine agonist compared to α_5 receptor complexes containing the asparagine variant (i.e., the risk variant robustly associated with heavier smoking) (Bierut et al. 2008).

'Knockout' mouse models have enabled us to further explore the impact of specific genes on specific phenotypes. A knockout mouse is simply a mouse in which a specific target gene has been selectively disabled. Studying the physiological characteristics or behaviours of such mice (e.g., assessing their response to nicotine), provides insight into the mechanisms which link the knocked out gene to disease. The development of α_5 knockout mouse models has illustrated the role that the *CHRNA5* gene plays in determining response to nicotine. Such mice are considered a model of individuals with reduced *CHRNA5* gene function (i.e., carriers of the rs16969968 minor allele, associated with increased heaviness of smoking). Fowler et al. (2011) conducted an elegant series of experiments using the α_5 knockout mouse model. Based on a self-administration paradigm, they observed that knockout mice responded far more vigorously than wild-type mice (i.e., mice with no inactivated genes) for nicotine infusions at high doses (see Fig. 2a). Whilst wild-type mice self-titrated delivery of nicotine dose to achieve a consistent, desired level during each test session ($\sim 1.5 \text{ mg kg}^{-1}$), knockout mice did not, consuming greater amounts as dosage increased (see Fig. 2b). The authors proposed that deficient α_5 signalling attenuates the negative effects of nicotine that normally serve to limit its intake, a conclusion which fits well with human research (i.e., smokers carrying the rs16969968 risk allele are likely to smoke more heavily than their counterparts without the risk allele). Fowler and colleagues further demonstrated that this effect could be 'rescued' in α_5 knockout mice through injection of a lentivirus vector into the medial habenula (MHb), an area where α_5 subunits are typically densely expressed, rescuing expression of α_5 subunits in this region. The MHb projects to the interpeduncular nucleus (IPN). Notably, the pathway between the MHb and the IPN has previously been shown to regulate avoidance of noxious substances such as quinine (Donovick et al. 1970). Fowler and colleagues further

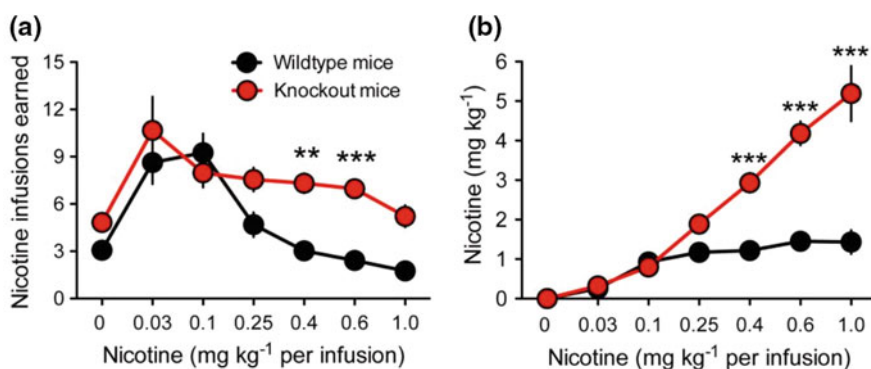


Fig. 2 Self-administration of nicotine in α_5 subunit knockout versus wild-type mice. *Panel a* Knockout and wild-type mice respond at an equivalent level for nicotine infusions at low doses. However, knockouts respond more vigorously than wild-types for high doses of nicotine. *Panel b* Wild-type mice self-titrate delivery of nicotine dose to achieve a consistent, desired level during test sessions ($\sim 1.5 \text{ mg kg}^{-1}$), whereas knockout mice do not, consuming greater amounts as dosage increases. Figure reprinted by permission from Macmillan Publishers Ltd: Nature (Fowler et al. 2011), copyright 2011

observed reduced activity in the IPN in response to nicotine in knockout animals. Additionally, disruption of IPN activity was found to increase nicotine self-administration. However, both knockout and wild-type mice were indistinguishable in terms of responding for low doses of nicotine (see Fig. 2), indicative of shared experience of the rewarding, stimulatory effects of nicotine at low doses in both groups. These observations are complemented by a previous study by Jackson (2010), which also illustrated the initially shared and later differential experience of reward in response to an increasing dose of nicotine between α_5 knockouts and wild-types, in this instance illustrated using a conditioned place preference task.

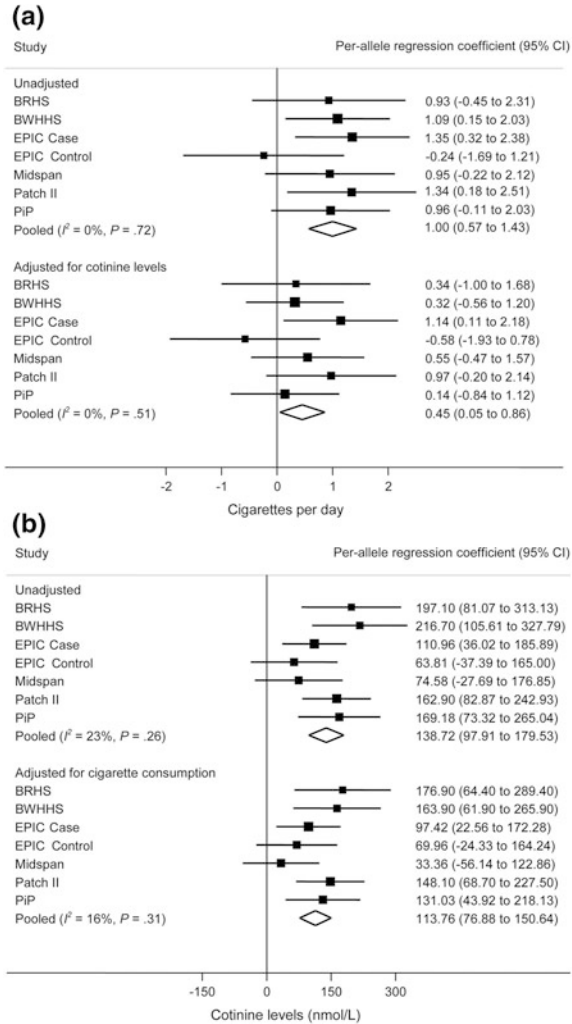
Fowler et al. (2011) concluded that high doses of nicotine stimulate the pathway between the MHB and the IPN pathway via nicotinic receptor complexes containing α_5 subunits. This results in the relay of an inhibitory motivational signal that serves to limit further nicotine intake. It is perhaps easiest to consider this pathway as a negative feedback loop, gated by α_5 containing nicotinic receptors, which is inadequately activated by high doses of nicotine in those with the rs16969968 risk allele.

3.2 Candidate Gene Follow-up Studies

Despite the success of GWAS, the number of individual genes identified, and in particular the proportion of phenotypic variance explained, remains relatively modest (typically <1 %). However, as we have discussed, the phenotypes employed in studies of smoking behaviour and often quite crude, for example using self-reported daily cigarette consumption as a measure of total tobacco exposure (Thorgeirsson et al. 2008). Subsequent studies have sought to follow-up on associations identified through GWAS techniques, employing a candidate gene approach in conjunction with refined phenotypes.

Munafò et al. (2012) examined the association of rs1051730 with both self-reported smoking quantity, and circulating levels of cotinine (see Fig. 3). Cotinine is the primary metabolite of nicotine, and is a much more precise and objective measure of total tobacco exposure (see chapter entitled [Pharmacogenetics of Nicotine and Associated Smoking Behaviors](#); this volume). Perhaps unsurprisingly, they observed a much stronger association between this locus and cotinine relative to self-reported daily cigarette consumption. Furthermore, the relationship observed between rs1051730 and cotinine was robust to adjustments made for self-reported cigarette consumption. This suggests that even amongst equal cigarette consumers, there is still genetically influenced variation in total nicotine exposure. This finding is supported by a previous study conducted by Le Marchand et al. (2008), who noted that the association between the same locus and total nicotine equivalents (an extremely accurate measure of nicotine exposure, encompassing nicotine, cotinine, related glucuronides, and 3-hydroxycotinine) was also robust for adjustments made for self-reported daily cigarette consumption. Munafò et al. (2012) propose that this is the result of differences in smoking topography (i.e., how a cigarette is smoked: number of puffs per cigarette, volume of smoke inhaled and so on).

Fig. 3 Meta-analysis of association of rs1051730/rs16969968 risk allele with two measures of heaviness of smoking in current smokers. *Panel a* self-reported cigarette consumption. *Panel b* biochemically-assessed cotinine levels. Munafò et al. (2012), by permission of Oxford University Press



These findings replicate and extend previous work by Keskitalo et al. (2009), who found that rs16969968 accounts for $\sim 1\%$ of the variation in self-report measures of smoking quantity, compared to $\sim 4\%$ of the variation in levels of cotinine. The key observation in these studies is that carefully designed and well-characterised phenotypes offer greater measurement precision. They provide a ‘cleaner’ genetic signal, and improve the likelihood of effect replication. While it may not be feasible to conduct GWAS on the same scale using these biochemical phenotypes as it is using self-report measures, these findings suggest two possibilities. First, variants identified through conventional GWAS can be followed up, exploring the candidate gene so identified with greater phenotype precision. Second, smaller GWAS which use these more precise phenotypes may still have greater statistical power, and identify

novel variants, compared to larger studies using less-precise phenotypes. We have recently demonstrated the utility of this approach, applying GWAS techniques to examine genetic determinants of cotinine level in current smokers (Ware et al. [in preparation](#)). A clear extension of this line of research would be to repeat this approach using total nicotine equivalents as a phenotype.

3.3 Recall-by-Genotype Studies

As we have seen, follow-up candidate gene studies of variants identified through GWAS can prove valuable in further refining the nature of the observed association. Together with the use of precisely-assessed phenotypes, they can provide important insights into underlying mechanisms of genetic action. One drawback to this approach, however, is that the measurement of detailed, precise phenotypes in large, adequately powered samples can be costly and time consuming. One approach which circumvents these issues is the ‘recall-by-genotype’ method, which we outline below.

The recall-by-genotype approach involves the selection of participants (or their biological samples) on the basis of a specific genotype, for further detailed, clinically-relevant, phenotype assessment. This approach has two key strengths. First, it is efficient. For example, statistical power can be improved by using genotype-specific recall instead of random sampling; it is far more powerful to selectively phenotype 100 minor homozygotes and 100 major homozygotes (at opposite ends of a biological gradient) than to phenotype 200 individuals at random, because the latter design would recruit far fewer homozygotes. Consider rs16969968 as an example. Only 16 % of Europeans are minor homozygotes at this locus. Therefore, in an unselected sample of 200 individuals, we would only expect 32 minor homozygotes. A second, related advantage of the recall-by-genotype approach is that it enables the collection of extremely detailed phenotypic data that would be impractical to collect in a much larger sample for reasons of time and expense.

We have recently adopted a recall-by-genotype approach to further investigate the association between rs16969968 and total nicotine exposure (Ware et al. [2014](#)). We saw in the previous section that the relationship between this locus and cotinine levels was robust to adjustments made for self-reported daily cigarette consumption, suggesting that even among equal cigarette consumers there is genetically influenced variation in total nicotine exposure. Munafò et al. ([2012](#)) proposed that this is likely due to differences in the way a cigarette is smoked (e.g., number of puffs taken per cigarette, volume of smoke inhaled per puff). Smoking topography is the term used to describe these nuances in smoking behaviour, and it can be assessed precisely using a small hand-held device. It has previously been established that cigarette smokers modify their smoking behaviour to self-titrate nicotine to a level appropriate to individual need (McNeill and Munafò [2013](#); Strasser et al. [2007](#)). What remains to be established, however, is whether differences in smoking topography mediate the relationship between the rs16969968 and nicotine

metabolite levels. We are currently investigating this using a recall-by-genotype approach, recalling current smokers from a large cohort study on the basis of minor and major homozygote status at rs1051730-rs16969968, assessing smoking topography and cotinine levels in both groups.

4 New and Emerging Technologies

As we discussed at the beginning of the chapter, smoking-related behaviours are highly heritable. Evidence from twin studies suggests that up to 75 % of the variation in nicotine dependence, for example, can be accounted for by genetic factors. However, findings from GWAS account for only a tiny proportion of such estimates. Why might this be? Have twin studies overestimated trait heritability? Do gene $\times$ gene and gene $\times$ environment interactions account for the shortfall? These are possible explanations. However, it is likely that a substantial proportion of ‘missing heritability’ remains undetected by GWAS because (a) the individual effects of common SNPs studied are too small to reach strict thresholds for genome-wide significance, and (b) rare and novel genetic variants, which are not captured by GWAS approaches, also contribute to the genetic architecture of these traits. Evidence in support of the former point is provided by genome-wide complex trait analysis (GCTA), a novel method for calculating the proportion of phenotypic variance explained by genome-wide SNPs for complex traits. Using this approach, Yang et al. (2010) illustrated that *common* SNPs explain 45 % of the variance in human height. This stands in stark contrast to the SNPs identified through height GWAS, which explain ~ 5 % variation, a far cry from the 80 % heritability estimates calculated from twin studies. As a more relevant example, Lubke et al. (2012) illustrated that 19 % of the variation in smoking initiation could be explained by ~ 1.1 M common SNPs. This is a substantially larger proportion of heritability than that explained by SNPs identified in GWAS of this phenotype.

It is also likely that rare and novel genetic variants, which are not captured by GWAS approaches, account for some of this additional unexplained heritability. In contrast to GWAS, genome sequencing, involving the use of high-throughput sequencing technologies to determine the order of nucleotide bases present in a particular individual, provides high-quality rare variant genotype calls. Given the size of the human genome (which consists of ~ 3 billion nucleotide base pairs), and the formerly prohibitive costs of sequencing whole genomes, much research in this field has focused on exome sequencing. Exome analysis involves the study of genetic variants within the exome only (i.e., protein-coding regions of the genome), which comprises ~ 1 % of the human genome (Ng et al. 2009). This approach allows for the identification of rare, non-synonymous variants associated with specific phenotypes. Of note, however, is that non-coding regions of the genome (e.g., introns) may also be of importance in determining disease. Indeed, the majority of trait-associated variants identified by GWAS fall in non-coding genome regions (Hindorff et al. 2009). For this reason, whole genome sequencing (covering ~ 95 %

of the genome) is preferable. Expense has, up until very recently, prohibited widespread use of this technique (the first sequenced human genome cost ~\$3 billion); however, recent advances in technology have resulted in a huge drop in cost in recent years, and the \$1,000 genome is now a reality (Hayden 2014).

Sequencing techniques are now being employed to further our understanding of tobacco dependence and other smoking-related phenotypes. An example of such a project is the GWAS and Sequencing Consortium of Alcohol and Nicotine use (GSCAN; <https://gscan.sph.umich.edu/>), led by Scott Vrieze at the University of Michigan. This is a collaborative effort comprising around 30 individual studies, with the aim of identifying genes associated with various tobacco- and alcohol-related phenotypes, including cigarettes per day, pack years, smoking initiation, and age of smoking initiation. The GSCAN project comprises three complimentary approaches: GWAS meta-analysis, exome meta-analysis, and whole genome sequencing meta-analysis. All contributing studies are to be imputed using a new haplotype reference panel, affording relatively high-density genotype imputation and accurate imputation of low frequency variants. Further, given the size of the collaboration effort, the sample size promises to be extremely large (affording increased power to detect small genetic effects). The exome meta-analysis should theoretically allow for the identification of rare, non-synonymous, exonic variants associated with nicotine-related phenotypes. Finally, the sequencing arm of the GSCAN project will be focused on following up variants identified through GWAS analyses, allowing evaluation at greatly increased genomic resolution. The complimentary use of GWAS and sequencing techniques has recently been used to great success (Holm et al. 2011). In short, this large, collaborative project holds the potential to substantially advance our understanding of the genetic basis of smoking-related behaviours.

5 Conclusions and Future Directions

We have known that smoking behaviours are heritable for some time, but the candidate gene era generated findings that proved difficult or impossible to replicate reliably, and taught us little about the specific genetic underpinnings of smoking behaviour. However, the introduction of genome-wide technologies in the mid-2000s has transformed our understanding of the genetic architecture of smoking behaviour, and identified a number of common variants associated with a range of smoking-related phenotypes. In particular, these approaches have implicated five nicotinic receptor subunit genes (*CHRNA5*, *CHRNA3*, *CHRNA6*, *CHRNA3* and *CHRNA4* see chapter entitled [Structure of Neuronal Nicotinic Receptors](#); this volume), all specifically in relation to smoking quantity (Thorgeirsson et al. 2010; Tobacco and Genetics Consortium 2010). The genomic region encoding the nicotine- and cotinine-metabolising enzyme CYP2A6 has further been implicated in smoking quantity (Thorgeirsson et al. 2010). *BDNF* has been highlighted in relation to smoking initiation and *DBH* in relation to smoking cessation (Thorgeirsson et al. 2010). These exciting initial findings have led to in vitro and preclinical research

focused on determining the mechanisms underlying these associations. Follow-up studies in humans employing carefully refined phenotypes have further developed our understanding of these relationships.

However, despite these successes, we can still only account for relatively little of the heritability of smoking behaviours. New sequencing approaches, which are increasingly becoming more affordable, hold promise and may offer the potential to discover rare or novel variants associated with smoking-related behaviours. However, as Loukola et al. (2014) stress, identification is “merely the first step in the discovery process”. Investigating the function of newly identified genetic variants must follow, through careful employment of the techniques discussed within this chapter (amongst others), which may ultimately lead to the development of novel treatments for tobacco dependence, and, as is the goal, an improvement in health outcomes.

References

- Agrawal A, Madden PA, Bucholz KK, Heath AC, Lynskey MT (2014) Initial reactions to tobacco and cannabis smoking: a twin study. *Addiction* 109:663–671
- Bierut LJ, Stitzel JA, Wang JC, Hinrichs AL, Gruzza RA, Xuei X, Saccone NL, Saccone SF, Bertelsen S, Fox L, Horton WJ, Breslau N, Budde J, Cloninger CR, Dick DM, Foroud T, Hatsukami D, Hesselbrock V, Johnson EO, Kramer J, Kuperman S, Madden PAF, Mayo K, Nurnberger J Jr, Pomerleau O, Porjesz B, Reyes O, Schuckit M, Swan G, Tischfield JA, Edenberg HJ, Rice JP, Goate AM (2008) Variants in the nicotinic receptors alter risk for nicotine dependence. *Am J Psychiatry* 165:1163–1171
- Caporaso N, Gu F, Chatterjee N, Sheng-Chih J, Yu K, Yeager M, Chen C, Jacobs K, Wheeler W, Landi MT, Ziegler RG, Hunter DJ, Chanock S, Hankinson S, Kraft P, Bergen AW (2009) Genome-wide and candidate gene association study of cigarette smoking behaviors. *PLoS ONE* 4:e4653
- Carmelli D, Swan GE, Robinette D, Fabsitz R (1992) Genetic influence on smoking—a study of male twins. *N Engl J Med* 327:829–833
- David SP, Hamidovic A, Chen GK, Bergen AW, Wessel J, Kasberger JL, Brown WM, Petruzella S, Thacker EL, Kim Y, Nalls MA, Tranah GJ, Sung YJ, Ambrosone CB, Arnett D, Bandera EV, Becker DM, Becker L, Berndt SI, Bernstein L, Blot WJ, Broeckel U, Buxbaum SG, Caporaso N, Casey G, Chanock SJ, Deming SL, Diver WR, Eaton CB, Evans DS, Evans MK, Fornage M, Franceschini N, Harris TB, Henderson BE, Hernandez DG, Hitsman B, Hu JJ, Hunt SC, Ingles SA, John EM, Kittles R, Kolb S, Kolonel LN, Le Marchand L, Liu Y, Lohman KK, McKnight B, Millikan RC, Murphy A, Neslund-Dudas C, Nyante S, Press M, Psaty BM, Rao DC, Redline S, Rodriguez-Gil JL, Rybicki BA, Signorello LB, Singleton AB, Smoller J, Snively B, Spring B, Stanford JL, Strom SS, Swan GE, Taylor KD, Thun MJ, Wilson AF, Witte JS, Yamamura Y, Yanek LR, Yu K, Zheng W, Ziegler RG, Zonderman AB, Jorgenson E, Haiman CA, Furberg H (2012) Genome-wide meta-analyses of smoking behaviors in African Americans. *Transl Psychiatry* 2:e119
- Donovick PJ, Burright RG, Zuromski E (1970) Localization of quinine aversion within the septum, habenula, and interpeduncular nucleus of the rat. *J Comp Physiol Psychol* 71:376–383
- Fowler CD, Lu Q, Johnson PM, Marks MJ, Kenny PJ (2011) Habenular alpha5 nicotinic receptor subunit signalling controls nicotine intake. *Nature* 471:597–601
- Hayden EC (2014) Is the \$1,000 genome for real? *Nature* doi:10.1038/nature.2014.14530
- Heath AC, Martin NG (1993) Genetic models for the natural history of smoking: evidence for a genetic influence on smoking persistence. *Addict Behav* 18:19–34

- Heath AC, Cates R, Martin NG, Meyer J, Hewitt JK, Neale MC, Eaves LJ (1993) Genetic contribution to risk of smoking initiation: comparisons across birth cohorts and across cultures. *J Subst Abuse* 5:221–246
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA (2009) Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 106:9362–9367
- Holm H, Gudbjartsson DF, Sulem P, Masson G, Helgadóttir HT, Zanon C, Magnusson OT, Helgason A, Saemundsdóttir J, Gylfason A, Stefansdóttir H, Gretarsdóttir S, Matthiasson SE, Thorgeirsson GM, Jonasdóttir A, Sigurdsson A, Stefansson H, Werge T, Rafnar T, Kiemeneý LA, Parvez B, Muhammad R, Roden DM, Darbar D, Thorleifsson G, Walters GB, Kong A, Thorsteinsdóttir U, Arnar DO, Stefansson K (2011) A rare variant in MYH6 is associated with high risk of sick sinus syndrome. *Nat Genet* 43:316–320
- Jackson KJ, Marks MJ, Vann RE, Chen X, Gamage TF, Warner JA, Damaj MI (2010) Role of alpha5 nicotinic acetylcholine receptors in pharmacological and behavioral effects of nicotine in mice. *J Pharmacol Exp Ther* 334:137–146
- Keskitalo K, Broms U, Heloidie, vaara M, Ripatti S, Surakka I, Perola M, PitkÄrniemi J, Peltonen L, Aromaa A, Kaprio J (2009) Association of serum cotinine level with a cluster of three nicotinic acetylcholine receptor genes (CHRNA3/CHRNA5/CHRNA4) on chromosome 15. *Hum Mol Genet* 18:4007–4012
- Le Marchand L, Derby KS, Murphy SE, Hecht SS, Hatsukami D, Carmella SG, Tiiirikainen M, Wang H (2008) Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. *Cancer Res* 68:9137–9140
- Liu JZ, Tozzi F, Waterworth DM, Pillai SG, Muglia P, Middleton L, Berrettini W, Knouff CW, Yuan X, Waeber G, Vollenweider P, Preisig M, Wareham NJ, Zhao JH, Loos RJ, Barroso I, Khaw KT, Grundy S, Barter P, Mahley R, Kesaniemi A, McPherson R, Vincent JB, Strauss J, Kennedy JL, Farmer A, McGuffin P, Day R, Matthews K, Bakke P, Gulsvik A, Lucae S, Ising M, Brueckl T, Horstmann S, Wichmann HE, Rawal R, Dahmen N, Lamina C, Polasek O, Zgaga L, Huffman J, Campbell S, Kooner J, Chambers JC, Burnett MS, Devaney JM, Pichard AD, Kent KM, Satler L, Lindsay JM, Waksman R, Epstein S, Wilson JF, Wild SH, Campbell H, Vitart V, Reilly MP, Li M, Qu L, Wilensky R, Matthaí W, Hakonarson HH, Rader DJ, Franke A, Wittig M, Schafer A, Uda M, Terracciano A, Xiao X, Busonero F, Scheet P, Schlessinger D, St Clair D, Rujescu D, Abecasis GR, Grabe HJ, Teumer A, Volzke H, Petersmann A, John U, Rudan I, Hayward C, Wright AF, Kolcic I, Wright BJ, Thompson JR, Balmforth AJ, Hall AS, Samani NJ, Anderson CA, Ahmad T, Mathew CG, Parkes M, Satsangi J, Caulfield M, Munroe PB, Farrall M, Dominiczak A, Worthington J, Thomson W, Eyre S, Barton A, Wellcome Trust Case Control C, Mooser V, Francks C, Marchini J (2010) Meta-analysis and imputation refines the association of 15q25 with smoking quantity. *Nat Genet* 42:436–440
- Loukola A, Hallfors J, Korhonen T, Kaprio J (2014) Genetics and smoking. *Curr Addict Rep* 1:75–82
- Lubke GH, Hottenga JJ, Walters R, Laurin C, de Geus EJ, Willemsen G, Smit JH, Middeldorp CM, Penninx BW, Vink JM, Boomsma DI (2012) Estimating the genetic variance of major depressive disorder due to all single nucleotide polymorphisms. *Biol Psychiatry* 72:707–709
- Maes HH, Sullivan PF, Bulik CM, Neale MC, Prescott CA, Eaves LJ, Kendler KS (2004) A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. *Psychol Med* 34:1251–1261
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA, Visscher PM (2009) Finding the missing heritability of complex diseases. *Nature* 461:747–753
- Marchini J, Howie B (2010) Genotype imputation for genome-wide association studies. *Nat Rev Genet* 11:499–511

- McNeill A, Munafò MR (2013) Reducing harm from tobacco use. *J Psychopharmacol* 27:13–18
- Munafò MR, Clark T, Johnstone E, Murphy M, Walton R (2004) The genetic basis for smoking behavior: a systematic review and meta-analysis. *Nicotine Tob Res* 6:583–597
- Munafò MR, Timofeeva MN, Morris RW, Prieto-Merino D, Sattar N, Brennan P, Johnstone EC, Relton C, Johnson PC, Walther D, Whincup PH, Casas JP, Uhl GR, Vineis P, Padmanabhan S, Jefferis BJ, Amuzu A, Riboli E, Upton MN, Aveyard P, Ebrahim S, Hingorani AD, Watt G, Palmer TM, Timpson NJ, Davey Smith G (2012) Association between genetic variants on chromosome 15q25 locus and objective measures of tobacco exposure. *J Natl Cancer Inst* 104:740–748
- Ng SB, Turner E, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J (2009) Targeted capture and massively parallel sequencing of twelve human exomes. *Nature* 461:272–276
- Picciotto MR, Zoli M, Rimondini R, Lena C, Marubio LM, Pich EM, Fuxe K, Changeux JP (1998) Acetylcholine receptors containing the beta2 subunit are involved in the reinforcing properties of nicotine. *Nature* 391:173–177
- Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PAF, Breslau N, Johnson EO, Hatsukami D, Pomerleau O, Swan GE, Goate AM, Rutter J, Bertelsen S, Fox L, Fugman D, Martin NG, Montgomery GW, Wang JC, Ballinger DG, Rice JP, Bierut LJ (2007) Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet* 16:36–49
- Sartor CE, Lessov-Schlaggar CN, Scherrer JF, Bucholz KK, Madden PA, Pergadia ML, Grant JD, Jacob T, Xian H (2010) Initial response to cigarettes predicts rate of progression to regular smoking: findings from an offspring-of-twins design. *Addict Behav* 35:771–778
- Schmitt JE, Prescott CA, Gardner CO, Neale MC, Kendler KS (2005) The differential heritability of regular tobacco use based on method of administration. *Twin Res Hum Genet* 8:60–62
- Strasser AA, Lerman C, Sanborn PM, Pickworth WB, Feldman EA (2007) New lower nicotine cigarettes can produce compensatory smoking and increased carbon monoxide exposure. *Drug Alcohol Depend* 86:294–300
- Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsater A, Flex A, Aben KK, de Vegt F, Mulders PF, Isla D, Vidal MJ, Asin L, Saez B, Murillo L, Blondal T, Kolbeinnsson H, Stefansson JG, Hansdottir I, Runarsdottir V, Pola R, Lindblad B, van Rij AM, Dieplinger B, Haltmayer M, Mayordomo JI, Kiemeny LA, Matthiasson SE, Oskarsson H, Tyrfinngsson T, Gudbjartsson DF, Gulcher JR, Jonsson S, Thorsteinsdottir U, Kong A, Stefansson K (2008) A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature* 452:638–642
- Thorgeirsson TE, Gudbjartsson DF, Surakka I, Vink JM, Amin N, Geller F, Sulem P, Rafnar T, Esko T, Walter S, Gieger C, Rawal R, Mangino M, Prokopenko I, Magi R, Keskitalo K, Gudjonsson IH, Gretarsdottir S, Stefansson H, Thompson JR, Aulchenko YS, Nelis M, Aben KK, den Heijer M, Dirksen A, Ashraf H, Soranzo N, Valdes AM, Steves C, Uitterlinden AG, Hofman A, Tonjes A, Kovacs P, Hottenga JJ, Willemsen G, Vogelzang N, Doring A, Dahmen N, Nitz B, Pergadia ML, Saez B, De Diego V, Lezcano V, Garcia-Prats MD, Ripatti S, Perola M, Kettunen J, Hartikainen AL, Pouta A, Laitinen J, Isohanni M, Huei-Yi S, Allen M, Krestyaninova M, Hall AS, Jones GT, van Rij AM, Mueller T, Dieplinger B, Haltmayer M, Jonsson S, Matthiasson SE, Oskarsson H, Tyrfinngsson T, Kiemeny LA, Mayordomo JI, Lindholt JS, Pedersen JH, Franklin WA, Wolf H, Montgomery GW, Heath AC, Martin NG, Madden PA, Giegling I, Rujescu D, Jarvelin MR, Salomaa V, Stumvoll M, Spector TD, Wichmann HE, Metspalu A, Samani NJ, Penninx BW, Oostra BA, Boomsma DI, Tiemeier H, van Duijn CM, Kaprio J, Gulcher JR, Consortium E, McCarthy MI, Peltonen L, Thorsteinsdottir U, Stefansson K (2010) Sequence variants at CHRNA6 and CYP2A6 affect smoking behavior. *Nat Genet* 42:448–453

- Tobacco-and-Genetics-Consortium (2010) Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat Genet* 42:441–447
- Visscher PM, Goddard ME, Derks EM, Wray NR (2012) Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol Psychiatry* 17:474–485
- Ware JJ, Timpson N, Davey Smith G, Munafò MR (2014) A recall-by-genotype study of CHRNA5-A3-B4 genotype, cotinine and smoking topography: study protocol. *BMC Med Genet* 15:13
- Ware JJ, Chen X, Vink J, Loukola A, Minica C, Pool R, Milaneschi Y, Mangino M, Menni C, Chen J, Peterson R, Auro K, Lyytikäinen L, Wedenoja J, Stiby AI, Hemani G, Willemsen G, Hottenga JJ, Korhonen T, Heliövaara M, Perola M, Rose R, Paternoster L, Timpson N, Wassenaar CA, Zhu AZX, Davey Smith G, Raitakari O, Lehtimäki T, Kähönen M, Koskinen S, Spector T, Penninx BWJH, Salomaa V, Boomsma D, Tyndale RF, Kaprio J, Munafò MR (in preparation) Genome-wide meta-analysis of cotinine levels in cigarette smokers identifies locus at 4q13.2.
- Wood AR, Perry JR, Tanaka T, Hernandez DG, Zheng HF, Melzer D, Gibbs JR, Nalls MA, Weedon MN, Spector TD, Richards JB, Bandinelli S, Ferrucci L, Singleton AB, Frayling TM (2013) Imputation of variants from the 1000 Genomes Project modestly improves known associations and can identify low-frequency variant-phenotype associations undetected by HapMap based imputation. *PLoS ONE* 8:e64343
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42:565–569

The Neurobiology and Genetics of Nicotine and
Tobacco

Balfour, D.J.; Munafò, M.R. (Eds.)

2015, VII, 231 p. 11 illus., 5 illus. in color., Hardcover

ISBN: 978-3-319-13664-6