Current Genetics and Epigenetics of Smoking/Tobacco-Related Cardiovascular Disease

Lutz P. Breitling

Abstract—Genetic and epigenetic factors are of great importance in cardiovascular biology and disease. Tobacco-smoking, one of the most important cardiovascular risk factors, is itself partially determined by genetic background and is associated with altered epigenetic patterns. This could render the genetics and epigenetics of smoking-related cardiovascular disease a textbook example of environmental epigenetics and modern approaches to multimodal data analysis. A pronounced association of smoking-related methylation patterns in the F2RL3 gene with prognosis in patients with stable coronary heart disease has recently been described. Nonetheless, surprisingly little concrete knowledge on the role of specific genetic variants and epigenetic modifications in the development of cardiovascular diseases in people who smoke has been accumulated. Beyond the current knowledge, the present review briefly outlines some chief challenges and priorities for moving forward in this field. (Arterioscler Thromb Vase Biol. 2013;33:1468-1472.)

Key Words: cardiovascular disease ■ epigenetics ■ genetics ■ methylation ■ tobacco-smoking

Genetic aspects of smoking-related disease have been a matter of interest since before smoking was universally accepted as one of the most important component causes of disease. Early assertions that lung cancer might predominantly occur in smoking subjects either because of common genetic factors predisposing simultaneously to the development of lung cancer and smoking or because of (pre)cancerous lesions leading to increased smoking rates could be greeted with smiles nowadays, but understanding the causality of observed associations remains one of the greatest challenges also in the modern times of genome-wide and epigenome-wide approaches to tobacco-smoking–related disease research.

Genetic Determination of Nicotine Dependence

Genetic predisposition to nicotine dependence has by now been firmly established. To some extent, all smoking/tobacco-related cardiovascular disease thus may contain a certain genetic component. Like other addictions, nicotine dependence is a complex phenotype resulting from a chain of behavioral events, and whereas environmental factors act especially during smoking initiation, genetic factors seem to be more important for the later progression to addiction. Intriguingly, one of the first studies investigating genetic determinants of smoking behavior on a genome-wide scale a priori included both a malignant and a cardiovascular phenotype as examples of smoking-related disease: single nucleotide polymorphisms in linkage disequilibrium with rs1051730 located on chromosome 15 in the vicinity of a cluster of genes coding for nico-}

Numerous other genes coding mostly for neurotransmitter receptors (eg, CHRNA3, GABRA4) or enzymes involved in neurotransmitter or nicotine metabolism (eg, DAO, FMO1, CYP2B6) have been implicated in nicotine dependence through candidate gene approaches, some of them showing sex-specific effects. Additional relevant genetic regions have also been identified in more extensive genome-wide analyses. The association of 15q24-25.1 with smoking continues to be refined, additional variants showing independent associations, for example, with the age at onset of daily and habitual smoking. Such aspects of smoking behavior are plausibly associated with severity of nicotine dependence, life-time cumulative tobacco–smoking exposure, and ultimately smoking-related cardiovascular disease risk. On a side note, recent evidence suggests that the association of rs1051730 and related polymorphisms with lung cancer is not exclusively because of their association with tobacco-smoking phenotypes, but rather acts substantially through additional pathophysiological connections. Such detailed studies on rs1051730 and peripheral arterial disease or other cardiovascular phenotypes are still lacking.

Interactions of Smoking With Cardiovascular Risk Genes

Classical gene–environment interactions are believed to play a role in tobacco-related cardiovascular disease and have been reported for various polymorphisms. One of the best studied examples is the interaction of smoking with apolipoprotein E variants. The apolipoprotein E isoform $\varepsilon 4$ is not only a worse antioxidant than the other variants, but it is also associated with a lipoprotein distribution shift favoring the production of oxidized lipoproteins. This genetic
variation in antioxidative resilience seems to have no clinical cardiovascular consequences by itself, but—in interaction with the excessive oxidative stress caused by smoking—it means that smoking-associated cardiovascular risks are much greater in e4 carriers than in other subjects.13 It has been realized that such gene–environment interactions could also interfere with the performance of lipid-lowering drugs, but the extent of this potentially far-reaching problem still needs to be clarified.15

Given the multitude of physiological phenomena involved in tobacco-smoking–related cardiovascular disease (ref. 12 for an informative overview), it is hardly surprising that similar interactions with smoking have been reported for sequence variants in candidate genes coding for proteins as diverse as lipoprotein lipase and interleukin-6,10 prothrombotic mutations like factor II G20210A,13 or transforming growth factor-β1.14 One study investigating an interaction of interleukin-18 polymorphisms with smoking regarding cardiovascular disease risk was exceptional for its laudable collaborative design emphasizing statistical power.15 Nonetheless, the level of evidence for the vast majority of currently suggested gene–smoking interactions ultimately remains very limited. The evolution of findings concerning an interaction between tobacco-smoking and variants in glutathione S-transferase genes with respect to cardiovascular disease risk—which started out as a seemingly plausible effect modification supported by multiple lines of evidence16 and was further suggested later meta-analysis,16 only to vanish with subsequent sample size escalation17—should be taken as a reminder of the danger of prematurely overinterpreting gene–smoking interactions or even higher order interactions (eg, gene–age–smoking).18 Thorough replication and follow-up studies need to become more of a norm in this statistically challenging field.

A New Era of Epigenetics

Studies on epigenetic aspects of health and disease have exploded in recent years, mainly because of technological advances allowing more and more epigenome-wide approaches.19 Definitions vary rather widely, but epigenetics generally refers to stable—yet impermanent and not necessarily heritable—regulatory mechanisms resulting in differences in gene expression that are not attributable to DNA sequence variation. The best-known components of this epigenome machinery are the methylation of cytosines at cytosine neighbored by guanine sites in the DNA sequence and various modifications of histone residues, such as acetylation or ubiquitination.20 Noncoding RNA with regulatory properties, such as microRNA, mostly act at the post-transcriptional level, but a subset is closely intertwined with the aforementioned mechanisms and themselves considered part of the epigenome machinery.21,22 Epigenetic differences present an additional layer of information that can be exploited to explain interindividual variation in disease phenotype development. There is also an obvious interface of environmental stimuli and genetic background, because they can be influenced by both kinds of factors.23,24 Especially DNA methylation is determined by sequence variants, so-called methylation quantitative trait loci,25,26 to a far greater extent than previously thought.27 The effect of methylation quantitative trait loci notably can differ between tissues, and this seems to be a more common phenomenon than for expression quantitative trait loci, that is sequence variants with an effect on gene expression levels.28 Methylation patterns altogether undergo major changes during development and tissue differentiation.29

The Encyclopedia of DNA Elements Consortium, which aims “to build a comprehensive parts list of functional elements in the human genome, including elements that act at the protein and RNA levels, and regulatory elements that control cells and circumstances in which a gene is active” (www.encodeproject.org), is becoming an important tool for better understanding the involvement of epigenetic mechanisms in the regulation of gene expression. For instance, it could be shown that differences in cytosine neighbored by guanine methylation levels closely correlate with cell type-specific variability of binding dynamics of CCCTC-binding factor, an important polyfunctional transcription factor.30

Epigenetics of Tobacco-Associated Cardiovascular Disease

The epigenetics of cardiovascular disease have received ample attention in the recent literature, and it has generally been accepted that epigenetic regulation plays an important role in cardiovascular biology. A great variety of epigenetic mechanisms are involved in physiological and pathophysiological vascular differentiation, proliferation, and related inflammatory processes.22 The importance of DNA methylation has been shown, for example, for the (dys)regulation of inducible nitric oxide synthase, whereas specific histone modifications have been identified that are involved in the fixation of proinflammatory gene expression patterns in diabetic disease models.31 Platelet function also seems to be crucially dependent on physiological epigenetic imprinting.32 All these processes are among the mechanisms considered important in smoking-associated cardiovascular pathology.32

Tobacco-smoking, on the other hand, has been investigated as a determinant of epigenetic patterns for quite some time. Associations of smoking behavior with crude epigenetic measures, such as surrogates of overall DNA methylation,33 or with modifications in specific gene promoters have been described.34 A certain focus in this line of research had been on candidate genes involved in malignant disease, rendering the findings of 2 recent epigenome-wide studies with single base resolution35,36 all the more interesting: of some 27 000 cytosine neighbored by guanine sites covered by the assay used, the one with the statistically most significant association with smoking behavior was located in F2RL3, a gene coding for protease-activated receptor 4. Proteinase-activated receptor 4 has been implicated in cardiovascular pathophysiology, especially inflammation,37 platelet function,38 and possibly perioperative myocardial injury after bypass surgery.39 Subsequently, a pronounced association of lower F2RL3 methylation with higher mortality (both cardiovascular and noncardiovascular) was demonstrated in a cohort of patients with stable coronary heart disease.40

Altogether, remarkably little is known about epigenetic patterns associated with cardiovascular disease in people who smoke. Even smoking-associated differential F2RL3 methylation was a chance finding, in the sense that it was identified in studies merely scanning for methylation differences according to smoking behavior.35,36 Both these
observational studies relied on agnostic epigenome-wide methylation assessment in DNA extracted from peripheral blood samples, which tend to be available in large epidemiological cohorts but—given the pronounced tissue-specificity of DNA methylation—possibly are not the most relevant sample matrix if epigenetic modifications relevant to cardiovascular disease are sought. Wet laboratory work, which could experimentally address the presence and consequences of methylation changes in cardiovascular candidate genes that could be caused by tobacco smoke exposure in appropriate cell cultures, seems to be altogether lacking. Successful proof-of-concept for such investigations can be found in the literature on malignant sequelae of smoking, including follow-up studies to support the functional relevance of hits from an epigenome-wide methylation screen.

**Challenges and Potentials of Epigenetics of Smoking-Related Cardiovascular Disease**

In the case of smoking and F2RL3, data published so far might justify the interpretation that tobacco-smoking results in slowly reversible changes in F2RL3 methylation, whatever the exact molecular mechanism. Plausible links of such an epigenetic pattern to the risk of adverse outcomes in patients with cardiovascular disease, however, are manifold (Figure). It will be difficult to elucidate what causal connection or mixture of connections is responsible for epigenetic associations described so far or to be discovered in the future.

Longitudinal studies, which were preferably conducted on different scales from populations to specific tissues and cells, with repeated assessments of exposure, methylation patterns, and carefully selected (endo)phenotypes, including gene expression levels, might be the most promising approach to resolve the question of causality in the relationships between smoking, differential DNA methylation, and cardiovascular disease. Mendelian randomization approaches, which make use of methylation quantitative trait loci as instrumental variables, could also help in some cases, but rely on the availability of sufficiently strong markers and a variety of other conditions that cannot necessarily always be satisfied.

Irrespective of causality, identifying simple correlational associations between interindividual epigenetic differences and clinical outcomes holds merit in itself, because it may help to improve cardiovascular risk prediction and stratification for the purpose of targeted prevention, intervention, and rehabilitation efforts. It seems promising that F2RL3 methylation was a much better predictor of prognosis in cardiovascular patients than was self-reported, cotinine-validated smoking behavior. Further studies are needed to clarify how well F2RL3 methylation or other epigenetic markers correlate with smoking-related cardiovascular risks, which decrease only slowly after smoking cessation and might be reflected better by epigenetic alterations than other markers. Epigenetics will not revolutionize cardiovascular risk prediction, but may contribute to multimodal risk stratification models in the future.

The greatest hope of the epigenomic approach nonetheless is to identify novel genes of importance for the cardiovascular consequences of smoking, that is, lying in or affecting the causal chain leading from smoking to cardiovascular clinical outcomes. Although this seems principally possible, it remains to be seen whether epigenetics will truly deliver in this regard and whether the epigenetics of cardiovascular disease actually differ substantially between smoking and nonsmoking subjects.

Apart from these most obvious objectives of epigenetic studies on cardiovascular diseases in people who smoke, it has also been suggested to use such methods to identify specific substances contained in tobacco smoke that are responsible for the development of smoking-related disease. This could be used (or abused) in the contentious area of tobacco product harm reduction, for example, by designing and promoting “healthier” tobacco products that do not affect F2RL3 methylation (but possibly maintain carcinogenic and other pathogenic potential). More or less on the other end of the spectrum of possible applications, information on tobacco-smoking–associated long-lasting modifications of the genetic material could be evaluated in the context of motivational interventions for smoking cessation. For instance, confronting expectant women who smoke with the fact that their behavior causes concrete alterations of the newborns’ DNA at F2RL3 and other specific loci could provide an additional incentive for quitting in this important target group. A different epigenetic mechanism intriguingly seems relevant and could be used on the paternal side: smoking is associated with

---

**Figure.** Some obvious causal structures that could explain the associations between tobacco-smoking, F2RL3 methylation, and cardiovascular disease (CVD) risk. A: Smoking might increase cardiovascular risk via a direct effect on F2RL3 methylation and subsequent protease-activated receptor 4 (PAR4) expression. B, F2RL3 methylation might be reversely influenced by smoking-associated changes in cardiovascular biology (eg, frequent transcription factor binding in the F2RL3 region attributable to a proinflammatory reaction could hamper passive F2RL3 methylation in smoking subjects). C, F2RL3 methylation might simply be a good proxy of smoking exposure and related risks. Hypotheses A through C are neither exhaustive nor mutually exclusive.
altered microRNA expression in spermatozoa, which might also contribute to adverse outcomes in the offspring.48

Genetics and Epigenetics of Smoking-Related Cardiovascular Disease: What Next?
Clearly, we have hardly even started to realize the potential of modern genetic and epigenetic approaches in this area. The foundation for changing this has largely been laid by the rapid spread of single base resolution large-scale assays.19 However fascinating even better coverage by whole (epi)genome sequencing methods may be, a reasonable and promising priority for rapidly advancing our understanding of genetics and epigenetics of smoking-related cardiovascular disease could be to improve the exploitation of data that can be acquired with currently widely available technologies.

It is likely that significant insights can still be extracted from data already collected. For example, a meta-analysis of genome-wide studies of respiratory phenotypes, which included gene–smoking interactions, identified multiple functional variants that had been missed in analyses neglecting interaction effects.49 A secondary methodological study found some evidence for an interaction of smoking with a variant located in the proximity of FMO4 with respect to incident coronary heart disease.50 These results should motivate secondary analyses of the evergrowing number of (epi)genome-wide data sets of cardiovascular disease, which usually will feature at least crude smoking exposure data, but have neither been compiled nor analyzed with tobacco-related disease in mind. Repeating the original data analysis separately for smoking and nonsmoking participants may already be worthwhile.

Further layers of complexity are added by attempts to integrate genetic and epigenetic information with data on the transcriptomic, proteomic, and metabolomic level. Even fairly straightforward approaches can yield interesting insights closer to functional relevance than pure (epi)genome-wide screens. Making use of gene expression in addition to genome-wide methylation data, it could, for example, be shown that altered epigenetic patterns are a prominent feature of myelodysplastic syndrome and correlate with the silencing or underexpression of specific apoptosis and immune response genes.51 Diverse data on sequence variation, regulatory mechanisms, and gene expression can also be combined to put forward testable hypotheses about the mechanistic relationships between genetic variants and disease phenotypes. Using Encyclopedia of DNA Elements data on gene regulation, such as DNase hypersensitivity peaks and transcription factor binding sites, an additionally probably functional polymorphism could be identified, that might explain the association of a nonfunctional polymorphism (rs1330349 on chromosome 9p21) with cardiovascular disease phenotypes, which had previously been found through a genome-wide screen.52 Indeed, the full potential of current multimodal data possibly can only be realized by complex bioinformatics in the context of systems biological analyses, which are also being established in the Cardiovascular field.53 The conscious incorporation of smoking phenotypes into these approaches certainly would be warranted and should significantly advance our understanding of the genetics and epigenetics of smoking/tobacco-related cardiovascular disease.

References
5. A secondary methodological study found some evidence for an interaction of smoking with a variant located in the proximity of FMO4 with respect to incident coronary heart disease. 50 These results should motivate secondary analyses of the evergrowing number of (epi)genome-wide data sets of cardiovascular disease, which usually will feature at least crude smoking exposure data, but have neither been compiled nor analyzed with tobacco-related disease in mind. Repeating the original data analysis separately for smoking and nonsmoking participants may already be worthwhile.

Further layers of complexity are added by attempts to integrate genetic and epigenetic information with data on the transcriptomic, proteomic, and metabolomic level. Even fairly straightforward approaches can yield interesting insights closer to functional relevance than pure (epi)genome-wide screens. Making use of gene expression in addition to genome-wide methylation data, it could, for example, be shown that altered epigenetic patterns are a prominent feature of myelodysplastic syndrome and correlate with the silencing or underexpression of specific apoptosis and immune response genes. 51 Diverse data on sequence variation, regulatory mechanisms, and gene expression can also be combined to put forward testable hypotheses about the mechanistic relationships between genetic variants and disease phenotypes. Using Encyclopedia of DNA Elements data on gene regulation, such as DNase hypersensitivity peaks and transcription factor binding sites, an additionally probably functional polymorphism could be identified, that might explain the association of a nonfunctional polymorphism (rs1330349 on chromosome 9p21) with cardiovascular disease phenotypes, which had previously been found through a genome-wide screen.52 Indeed, the full potential of current multimodal data possibly can only be realized by complex bioinformatics in the context of systems biological analyses, which are also being established in the Cardiovascular field.53 The conscious incorporation of smoking phenotypes into these approaches certainly would be warranted and should significantly advance our understanding of the genetics and epigenetics of smoking/tobacco-related cardiovascular disease.

Disclosures
None.

Genetics/Epigenetics of Smoking-Related CVD

1471

Breitling

References
5. A secondary methodological study found some evidence for an interaction of smoking with a variant located in the proximity of FMO4 with respect to incident coronary heart disease. 50 These results should motivate secondary analyses of the evergrowing number of (epi)genome-wide data sets of cardiovascular disease, which usually will feature at least crude smoking exposure data, but have neither been compiled nor analyzed with tobacco-related disease in mind. Repeating the original data analysis separately for smoking and nonsmoking participants may already be worthwhile.

Further layers of complexity are added by attempts to integrate genetic and epigenetic information with data on the transcriptomic, proteomic, and metabolomic level. Even fairly straightforward approaches can yield interesting insights closer to functional relevance than pure (epi)genome-wide screens. Making use of gene expression in addition to genome-wide methylation data, it could, for example, be shown that altered epigenetic patterns are a prominent feature of myelodysplastic syndrome and correlate with the silencing or underexpression of specific apoptosis and immune response genes. 51 Diverse data on sequence variation, regulatory mechanisms, and gene expression can also be combined to put forward testable hypotheses about the mechanistic relationships between genetic variants and disease phenotypes. Using Encyclopedia of DNA Elements data on gene regulation, such as DNase hypersensitivity peaks and transcription factor binding sites, an additionally probably functional polymorphism could be identified, that might explain the association of a nonfunctional polymorphism (rs1330349 on chromosome 9p21) with cardiovascular disease phenotypes, which had previously been found through a genome-wide screen.52 Indeed, the full potential of current multimodal data possibly can only be realized by complex bioinformatics in the context of systems biological analyses, which are also being established in the Cardiovascular field.53 The conscious incorporation of smoking phenotypes into these approaches certainly would be warranted and should significantly advance our understanding of the genetics and epigenetics of smoking/tobacco-related cardiovascular disease.

References
5. A secondary methodological study found some evidence for an interaction of smoking with a variant located in the proximity of FMO4 with respect to incident coronary heart disease. 50 These results should motivate secondary analyses of the evergrowing number of (epi)genome-wide data sets of cardiovascular disease, which usually will feature at least crude smoking exposure data, but have neither been compiled nor analyzed with tobacco-related disease in mind. Repeating the original data analysis separately for smoking and nonsmoking participants may already be worthwhile.

Further layers of complexity are added by attempts to integrate genetic and epigenetic information with data on the transcriptomic, proteomic, and metabolomic level. Even fairly straightforward approaches can yield interesting insights closer to functional relevance than pure (epi)genome-wide screens. Making use of gene expression in addition to genome-wide methylation data, it could, for example, be shown that altered epigenetic patterns are a prominent feature of myelodysplastic syndrome and correlate with the silencing or underexpression of specific apoptosis and immune response genes. 51 Diverse data on sequence variation, regulatory mechanisms, and gene expression can also be combined to put forward testable hypotheses about the mechanistic relationships between genetic variants and disease phenotypes. Using Encyclopedia of DNA Elements data on gene regulation, such as DNase hypersensitivity peaks and transcription factor binding sites, an additionally probably functional polymorphism could be identified, that might explain the association of a nonfunctional polymorphism (rs1330349 on chromosome 9p21) with cardiovascular disease phenotypes, which had previously been found through a genome-wide screen.52 Indeed, the full potential of current multimodal data possibly can only be realized by complex bioinformatics in the context of systems biological analyses, which are also being established in the Cardiovascular field.53 The conscious incorporation of smoking phenotypes into these approaches certainly would be warranted and should significantly advance our understanding of the genetics and epigenetics of smoking/tobacco-related cardiovascular disease.
Genetics and epigenetics play an important role not only in cardiovascular physiology, but also in nicotine dependence and tobacco-smoking, one of the most important cardiovascular risk factors. Trying to summarize the current understanding of genetics and epigenetics of smoking-associated cardiovascular disease among people who smoke, and outlines some chief challenges and priorities for future research.
Current Genetics and Epigenetics of Smoking/Tobacco-Related Cardiovascular Disease
Lutz P. Breitling

Arterioscler Thromb Vasc Biol. 2013;33:1468-1472; originally published online May 2, 2013;
doi: 10.1161/ATVBAHA.112.300157
Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://atvb.ahajournals.org/content/33/7/1468

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Arteriosclerosis, Thrombosis, and Vascular Biology can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Arteriosclerosis, Thrombosis, and Vascular Biology is online at:
http://atvb.ahajournals.org//subscriptions/