Guest Editorial

Epidemiological studies carried out as early as the 1950s established an association between smoking and lung cancer that literally jumped out at us: smokers accounted for 70–80 per cent of lung cancer patients.[1] More recent epidemiological studies have reliably replicated this finding.[2] All of the epidemiological studies, both old and new, also found that a minority (10–15 per cent) of smokers contract this often fatal disease. Consequently, the medical and scientific communities have been infatuated with the question: why do a minority of smokers develop lung cancer? Clearly, this is an important question, given that lung cancer accounted for approximately 30 per cent (160,390) of all cancer-related deaths reported in the USA in 2007.[3]

The famous statistician, Sir Ronald Fisher, was the first to suggest that genes regulate susceptibility to smoking-associated lung cancer in a 1958 study that determined concordance rates for lung cancer and smoking in monozygotic (MZ) and dizygotic (DZ) twins.[4] Fisher concluded that genetic factors do not influence lung cancer because the MZ and DZ concordances for lung cancer are the same. This conclusion has been ignored, in part, because more recent twins studies that included larger sample sizes have detected significantly higher concordance for lung cancer in MZ twins versus DZ twins[5,6] and because lung cancer shows familial aggregation.[7,8] The generally held opinion seems to be that only some smokers contract lung cancer because genes, probably more than one, play vital roles in regulating susceptibility to developing lung cancer. While the generally held opinion may be correct, it is clear from those studies that have attempted to quantify it that genetic influence on lung cancer is moderate to low. For example, 8 per cent heritability for lung cancer was calculated from the 9.6 million subject Swedish Family-Cancer database,[9] and heritability estimates derived from twins and family studies are in the 10 per cent range.[10,11]

Candidate gene and genome-wide association studies of lung cancer

Identifying genes that play a causal role in the development of lung cancer has been a major source of entertainment and frustration for geneticists for the past 50 years. Studies that evaluate candidate genes have been especially popular because they are often based on what appears to be sound biology. If the basic science branch of cancer research identifies a gene product that seems to influence the cancer process, geneticists have frequently attempted to identify polymorphisms in the human genes and have then designed studies to determine whether these polymorphisms influence susceptibility to lung cancer. Hundreds, if not thousands, of studies have attempted to determine whether a long list of candidate genes might influence the development of lung cancer. In a 2008 review of the literature, Risch and Plass conclude that lung cancer is a complex disease, influenced by low-penetrance polymorphisms in multiple genes, and that epigenetic factors may also be important.[12]

Three recently published studies used the genome-wide association study (GWAS) approach to identify genes that contribute to lung cancer.[13–15] All three studies used large panels of single nucleotide polymorphisms (SNPs) and large subject pools, affording sample sizes that should have provided adequate power to detect a genetic influence that contributes to as little as 2–3 per cent of the variance. It is somewhat surprising that a GWAS identified anything, given that the heritability for lung cancer seems to be so very low (<10 per cent), but all of these studies detected significant associations between lung cancer and the same cluster of genes resident on chromosome 15. This gene cluster encodes three

neuronal nicotinic cholinergic receptor (nAChR) subunit genes: α3 (*CHRNA3*), α5 (*CHRNA5*) and β4 (*CHRNB4*). The strongest association with lung cancer was found with *CHRNA5*. The fact that three different studies, using large sample sizes, yielded identical results must be viewed as provocative, if not downright exciting.

The GWAS approach might be described as a wide-eyed innocence approach, in that it is not burdened by preconceptions; all genes are fair game. This strength can also be a weakness, because once a gene has been discovered, researchers must suggest a mechanism that might explain how a polymorphism associated with the gene contributes to variation in a phenotype of interest. Providing a mechanism often involves obtaining answers to questions, such as: (1) Is the gene product expressed in the right place (e.g. is a lung cancer gene expressed in the lung)? (2) Does the polymorphism affect expression level or function? Obviously, it is not absolutely necessary for a lung cancer gene to be expressed in the lung, or that a polymorphism affects expression level or function, but it certainly would be convenient. Excitement concerning the *nAChR* gene cluster finding is enhanced by the observations that α5 mRNA and protein are expressed in lung epithelial cells, and by the demonstration that one of the *CHRNA5* polymorphisms results in an aspartic acid–asparagine switch at position 398 (second cytoplasmic loop of the α5 nAChR subunit gene product) that affects receptor function when expressed as α4α5β2 nAChRs in HEK cells.[16] We seem to have nearly everything: replication of the finding in three large studies and demonstration that the gene product is found in the lung and that one of strongest association polymorphisms alters receptor function. What more could we ask?

Tobacco addiction rears its ugly head

Establishing whether genetic factors influence lung cancer has been complicated by the finding that genetic factors affect individual differences in vulnerability to tobacco addiction. Fisher introduced this issue in his 1958 study,[4] when he noted that concordance for smoking status was significantly higher in MZ than in DZ twins, thereby suggesting that genes might influence whether a given individual does, or does not, smoke. A recent review of the genetics of smoking literature notes that smoking is a trait with high heritability (h^2 estimates are 0.28–0.84) that is most probably influenced by multiple genes.[17] Modern studies are attempting to identify genes that might modulate vulnerability to tobacco addiction. Of particular note, three recently published studies detected significant associations between *CHRNA5* gene markers and one, or more, components of tobacco use. One of these studies used the GWAS[18] method, and two candidate gene studies evaluated the *D398N* polymorphism in *CHRNA5*.[16,19] If we apply the 'location' and 'activity' questions to the nicotine addiction and α5 data we get provocative answers. The brain expresses α5-containing nAChRs, most often as the high-affinity α4α5β2 type, in dopaminergic neurones[20] that play vital roles in drug reinforcement, and, as noted previously, the *D398N* α5 polymorphism affects receptor function.[16] Once again, what more could we ask? The gene product is expressed in the right place and it affects activity, so it must be involved in the addiction process!

All three of the genome-wide association studies of lung cancer analysed their data in ways that would, hopefully, provide an answer to the 'direct effect on cancer versus an effect on addiction' issue. Unfortunately, the answer seems to differ, depending on the study. The study reported by Hung *et al.*[13] detected an increased risk for lung cancer that was associated with the chromosome 15 locus in six samples, where smokers'-only data were analysed and when the combined sample was divided into groups of former smokers, current smokers and never smokers. These authors concluded that the *CHRNA5*-associated locus exerts a direct effect on lung cancer because a significant association was also found between the *CHRNA5* locus and lung cancer in non-smokers and because the chromosome 15 locus was not associated with variance in three frequently used measures of tobacco addiction (time to first cigarette each day, number of cigarettes per day and the Fagerstrom Tolerance Questionnaire). Amos and colleagues[14] detected a

significant association between chromosome 15 *nAChR* gene cluster SNP markers and increased risk for lung cancer, but also found that SNPs in this region seem to influence smoking behaviour, particularly among former smokers. These findings led the authors to conclude,'...although nAChR genes may have a role in smoking behavior, variation in chromosome 15 *nAChR* gene cluster markers directly contributes to lung cancer susceptibility'. By contrast, Thorgeirsson *et al.*[15] concluded that an association between the 15q5.4 region and nicotine addiction is a key component in the aetiology of lung cancer because the apparent effect of the chromosome 15 SNPs on lung cancer was substantially reduced after correcting for smoking quantity. Thus, the question: 'does *CHRNA5* affect lung cancer directly?' was answered with a 'yes', a 'maybe' and a 'no'.

Lung cancer is a phenotype that is readily measured; an individual does, or does not, have the disease. In comparison, tobacco addiction is a messy phenotype and tobacco researchers have struggled for years to define the characteristics of a tobacco addict. The studies that were designed to identify genes that influence lung cancer used different criteria when defining tobacco addiction. The studies also differed in the populations studied and other important criteria. Consequently, it is not surprising that different studies that addressed the same question could yield different answers.

Further complications

Given that the heritability of lung cancer is low and the heritability of tobacco use/addiction is high, it seems most likely that the path between *CHRNA5* and lung cancer has a behavioural intermediate. That intermediate may not involve the actions of nicotine on the brain. This possibility is suggested by our recent finding,[21] obtained from two separate young adult samples, that age of initiation of tobacco (and alcohol) use is significantly associated with the *CHRNA5-CHRNA3-CHRNB4* gene cluster. This result may mean that the gene cluster influences an underlying psychological characteristic (eg impulsivity or behavioural disinhibition) that increases the likelihood that an adolescent will experiment with drugs such as tobacco and alcohol. Early-onset tobacco use will increase life-long exposure to tobacco, a factor that is known to increase the risk for lung cancer.

Suggestions for resolving these issues

Establishing a causal connection between a polymorphism in the *CHRNA5* gene and lung cancer, tobacco addiction or behavioural disinhibition may be next to impossible using currently available genetic strategies. Human genetic studies can detect nothing more than significant associations (correlations) between genetic markers and phenotypes. It may be that additional human genetic studies will be required to provide a more rigorous test of the hypotheses that have evolved from the recent genetic analyses of lung cancer and smoking. Such studies, if done, would increase, or decrease, our enthusiasm for the gene cluster hypotheses. They cannot, unfortunately, bridge the huge gap between correlation and causality. We suggest that researchers who work in the lung cancer area consider mounting a transgenic research programme. Mice have been engineered to express null mutations for all of the known *nAChR* subunit genes, and gain-of-function mutants have been developed for several. These mice have been used to test hypotheses concerning nicotine addiction[22–24] and could be used to test hypotheses concerning lung cancer. It is absolutely the case that animal models may not mimic the human condition exactly, and it is possible that molecular and biochemical compensation may influence results obtained with genetically engineered animals. Nonetheless, experiments carried out using this alternative approach might serve as a good test of hypotheses generated by human genetic studies.

Allan C. Collins, Matthew McQueen and Marissa Ehringer
Institute for Behavioral Genetics
University of Colorado
Boulder, CO
USA

References

1. Doll, R. and Hill, A.B. (1952), 'A study of the aetiology of carcinoma of the lung', *Br. Med. J.* Vol. 2, pp. 1271–1286.
2. Alberg, A.J., Ford, J.G. and Samet, J.M. (2007). 'Epidemiology of lung cancer: ACCP evidence-based clinical practice guidelines', *Chest* Vol. 132, pp. 29S–55S.
3. Surveillance Epidemiology and End Result (SEER) statistics (2007), http://seer.cancer.gov/.
4. Fisher, R.A. (1958), 'Cancer and smoking', *Nature* Vol. 182, p. 596.
5. Hemminki, K., Dong, C. and Vaittinen, P. (2001), 'Cancer risks to spouses and offspring the Family-Cancer Database', *Genet. Epidemiol.* Vol. 20, pp. 247–257.
6. Braun, M.M., Caporaso, N.E., Page, W.F. and Hoover, R.N. (1995), 'A cohort study of twins and cancer', *Cancer Epidemiol. Biomarkers Prev.* Vol. 4, pp. 469–473.
7. Sellers, T.A., Chen, P.L., Potter, J.D. *et al.* (1994), 'Segregation analysis of smoking-associated malignancies: Evidence for Mendelian inheritance', *Am. J. Med. Genet.* Vol. 52, pp. 308–314.
8. Bermejo, J.L. and Hemminki, K. (2005), 'Familial lung cancer and aggregation of smoking habits: A simulation of the effect of shared environmental factors on the familial risk of cancer', *Cancer Epidemiol. Biomarkers Prev.* Vol. 14, pp. 1738–1740.
9. Czene, K., Lichtenstein, P. and Hemminki, K. (2001), 'Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer database', *Int. J. Cancer* Vol. 99, pp. 260–266.
10. Hemminki, K., Lonnstedt, I., Vaittanin, P. and Lichtenstein, P. (2001), 'Estimation of genetic and environmental components in colorectal and lung cancer and melanoma', *Genet. Epidemiol.* Vol. 20, pp. 107–116.
11. Lichtenstein, P., Holm, N.V., Verkasalo, P.K. *et al.* (2000), 'Environmental and heritable factors in the causation of cancer — Analyses of cohorts of twins from Sweden, Denmark, and Finland', *N. Engl. J. Med.* Vol. 343, pp. 78–85.
12. Risch, A. and Plass, C. (2008), 'Lung cancer epigenetics and genetics', *Int. J. Cancer* Vol. 123, pp. 1–7.
13. Hung, R.J., McKay, J.D., Gaborieau, V. *et al.* (2008), 'A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25', *Nature* Vol. 452, pp. 633–637.
14. Amos, C.I., Wu, X., Broderisck, P. *et al.* (2008), 'Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1', *Nat. Genet.* Vol. 40, pp. 616–622 .
15. Thorgeirsson, T.E., Geller, F., Sulem, P. *et al.* (2008), 'A variant associated with nicotine dependence, lung cancer and peripheral arterial disease', *Nature* Vol. 452, pp. 638–642.
16. Bierut, L.J., Hesselbrock, V., Stitzel., J.A. *et al.* (2008), 'Variants in nicotinic receptors and risk for nicotine dependence', *Am. J. Psychiatry*, In press.
17. Ho, M.K. and Tyndale, R.F. (2007), 'Overview of the pharmacogenomics of cigarette smoking', *Pharmacogenomics J.* Vol. 7, pp. 81–98.
18. Berrittini, W., Yuan, X., Tozzi, F. *et al.* (2008), 'Alpha-5/α-3 nicotinic receptor subunit alleles increase risk for heavy smoking', *Mol. Psychiatry* Vol. 13, pp. 368–373.
19. Saccone, S.F., Hinrichs, A.L., Saccone, N.L. *et al.* (2007), 'Cholinergic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs', *Hum. Mol. Genet.* Vol. 16, pp. 36–49.
20. Salminen, O., Murphy, K.L., McIntosh, J.M., Drago, J., Marks, M.J. and Collins, A.C. (2004), 'Subunit composition and pharmacology of two classes of striatal presynaptic nicotinic acetylcholine receptors mediating dopamine release in mice', *Mol. Pharmacol.* Vol. 65, pp. 1526–1535.
21. Schlaepfer, I.R., Hoft, N.R., Collins, A.C. *et al.* (2008), 'The CHRNA5/A3/B4 gene cluster variability as an important determinant of early alcohol and tobacco initiation in young adults', *Biol. Psychiatry* Vol. 63, pp. 1039–1046.
22. Picciotto, M.R., Zoli, M., Rimondini, R. *et al.* (1998), 'Acetylcholine receptors containing the β2 subunit are involved in the reinforcing properties of nicotine', *Nature* Vol. 391, pp. 173–177.
23. Tapper, A.R., McKinney, S.S., Nashmi, R. *et al.* (2004), 'Nicotine activation of α4* receptors: Sufficient for reward, tolerance and sensitization', *Science* Vol. 306, pp. 1029–1032.
24. Nashmi, R., Xiao, C., Deshpande, P. *et al.* (2007), 'Chronic nicotine cell specifically upregulates functional α4* nicotinic receptors: Basis for both tolerance in midbrain and enhanced LTP in perforant path', *J. Neurosci.* Vol. 2, pp. 8202–8218.