

GENOMICS

When the smoke clears ...

Stephen J. Chanock and David J. Hunter

Three studies identify an association between genetic variation at a location on chromosome 15 and risk of lung cancer. But they disagree on whether the link is direct or mediated through nicotine dependence.

The characterization of variation in the human genome has begun to pay dividends. These are mainly being delivered through genome-wide association studies (GWAS), provided that they are sufficiently large and are followed by replication of the notable findings. Since early 2007, variations at nearly 100 regions of the genome have been associated with an increased risk for diseases with a complex genetic background, such as diabetes, inflammatory bowel disease, cancer (most notably breast, colorectal and prostate) and heart disease.

The latest instalment in the GWAS story deals with lung cancer, and comes in the form of three papers — two published in this issue^{1,2}, on pages 633 and 638, and one in *Nature Genetics*³. All of them identify variation in the same region of the long arm of chromosome 15 (15q24/15q25.1) as the ‘top hit’ for genomic association with lung cancer. Among the genes in this region are those that encode subunits of nicotinic acetylcholine receptors, which as their name suggests have an affinity for nicotine. The genetic variation analysed was in the form of single nucleotide polymorphisms (SNPs). These are DNA sequence variations that arise from the substitution of one nucleotide base for another, and overall constitute approximately 90% of common variation in the human genome.

The three studies are all large and appropriately replicated, and provide strong evidence for an association between SNP variation at 15q24/15q25.1 and lung cancer. But they differ on whether the connection is direct or mediated via smoking behaviour — that is, characteristics such as the duration and ‘dose’ of lifetime smoking, and/or the propensity for nicotine addiction. Previous studies^{4,5} had identified the genes encoding subunits of nicotinic acetylcholine receptors as strongly associated with smoking behaviour. The association between smoking and lung cancer is among the strongest in the epidemiological firmament, and any gene variant that is modestly linked with smoking behaviour will seem to be associated with lung cancer unless the matching of cancer cases and controls by smoking behaviour is close to perfect.

Thorgerisson *et al.*² report an association

of SNP variation at 15q24/15q25.1 with the number of cigarettes smoked per day and a nicotine-dependence scale. They suggest that the link with lung cancer is primarily mediated through nicotine dependence and thus provides a case study of a gene–environment correlation in the pathogenesis of disease. The two other groups, Hung *et al.*¹ and Amos *et al.*³, examined SNP variation in patients with lung cancer and control subjects, and reach the opposite conclusion — namely, that the association is primarily with lung cancer and not with smoking.

In these kinds of studies, researchers distinguish between ‘ever-smokers’ and ‘never-smokers’. Ever-smokers are current or former smokers; never-smokers are usually defined as those who have smoked fewer than 100 cigarettes in a lifetime.

Amos and colleagues³ studied both cancer patients and controls mainly in ever-smokers. They describe the evidence for the association of their ‘lead’ SNPs with smoking behaviour (among ever-smokers) as “weak”. Hung *et al.*¹ did not match cases and controls by smoking, but carefully examined the relation between the lead SNPs and smoking behaviour. In contrast to Thorgerisson *et al.*, Hung *et al.* did not observe an association between their lead SNP and any of the measures of nicotine addiction, and thus suggest that the genetic association in their study is unlikely to be due to confounding effects of smoking behaviour.

The two studies^{1,3} that also examined risk of lung cancer in never-smokers reached different conclusions, but we should consider this discrepancy preliminary. Hung *et al.*¹, with 352 cases among never-smokers, conclude that there is an association with the risk SNPs; Amos *et al.*³ conclude that there is no such association, although with only 125 never-smoking cases in their series, statistical power was even more limited.

The extent of the disagreement in the conclusions of these studies is remarkable. Careful examination of the relation between the SNP variants in even larger data sets, with

detailed measures of both smoking patterns and addictive behaviour, is needed to discern whether these discoveries relate to the risk of smoking, the risk of lung cancer or the risk of both. These are the first GWAS to attempt to identify the genetic component of a disease that has such an overwhelmingly strong environmental cause. They signal the need for greater methodological rigour in attempts to account for both the genetic and the environmental causes that we think underlie most diseases.

“These are the first genome-wide association studies to attempt to identify the genetic component of a disease that has such an overwhelmingly strong environmental cause.”

As GWAS have become available there has been something of a ‘flight to quantity’ as the importance of large sample sizes to detect modest effects has become apparent. Such sample sizes have often been obtained at the expense of collecting even minimal environmental and lifestyle

information. Much GWAS research has tended to be retrospective, and thus this information could be biased; the requirements for public access to individuals’ data may be suppressing the availability of potentially high-quality data sets because of limitations in obtaining informed consent. These studies^{1–3} demonstrate the importance of balancing the undeniable need for quantity in GWAS with the need for quality of ancillary data, especially if we are to disentangle and understand the interplay between genes and environment.

The next round of research will involve resequencing 15q24/15q25.1 and incorporating any additional SNP variants into future epidemiological work. Moreover, further genome-wide scans are needed — especially in studies that can look at patients with lung cancer for whom detailed data on smoking exposure, nicotine dependence and duration and intensity of smoking are available. It is also essential to examine larger samples of non-smoking lung cancer patients and to perform analyses according to the specific type of cancer. Pooling of GWAS data will help, as will full disclosure of SNP rankings for the major smoking behaviours to permit a better understanding of the genetic associations with these behaviours (rankings based on

grouped data can be made fully available and disclosure poses no risk of loss of confidentiality for individual subjects).

Can findings such as these be applied in the clinic? Companies are already offering direct-to-consumer genetic testing for the risk of developing various diseases, using the same SNP chips as in the new studies¹⁻³. The rationale claimed for determining risk in individuals is that it will encourage them to change their lifestyle and/or undergo screening so that disease can be detected earlier. In the case of smoking, however, there is a danger that such 'personalized' medical advice will weaken the public-health message that everyone should avoid smoking. Even if a subset of people are deemed 'resistant' to the effects of smoking on lung cancer development, it is unlikely that they will also be protected against the adverse effects of heart disease and obstructive pulmonary disease, disorders that are also associated with smoking. On the other hand, we may be able to evaluate smoking-cessation treatments informed by knowledge of a person's genetic predisposition to start smoking or to nicotine addiction, and thus add new weapons to the anti-smoking arsenal.

Follow-up studies should clear the smoke clouding the differing conclusions of these papers³ and establish the biological rationale for the robust association of 15q24/15q25.1 with lung cancer. It is at least reassuring that all three groups point to the same region on chromosome 15. The sceptics have fretted that association studies would be riddled with false positive results; yet, because of the high standards that have been developed⁶, the evidence for the association between some genetic loci and certain complex diseases is now unequivocal. For most diseases, more loci will surely be discovered, adding to the first wave of results that have been primarily related to disease causation. On the horizon we can see the crest of studies reporting on disease outcomes, but progress, both in understanding the basic causes and in estimating personal risks, will require environmental and lifestyle factors to be taken into account. To quote Winston Churchill, it is "perhaps the end of the beginning" in the battle to understand complex diseases. ■

Stephen J. Chanock is in the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892-4605, USA. David J. Hunter is in the Department of Epidemiology, Harvard School of Public Health, and the Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA.

e-mails: chanocks@mail.nih.gov;
dhunter@hsph.harvard.edu

- Hung, R. J. *et al.* *Nature* **452**, 633-637 (2008).
- Thorgeirsson, T. E. *et al.* *Nature* **452**, 638-641 (2008).
- Amos, C. I. *et al.* *Nature Genet.* doi:10.1038/ng.109 (2008).
- Saccone, S. F. *et al.* *Hum. Mol. Genet.* **16**, 36-49 (2007).
- Berrettini, W. *et al.* *Mol. Psych.* **13**, 368-373 (2008).
- Chanock, S. J. *et al.* *Nature* **447**, 655-660 (2007).

EXTRASOLAR PLANETS

With a coarse-tooth comb

Gordon Walker

The search for Earth-like planets outside our Solar System is bedevilled by the lack of an adequate frequency standard for calibrating starlight. Tweaking existing laser 'frequency combs' could be a way forward.

As we look for planets orbiting other suns, it is often tiny, periodic shifts in the spectrum of light coming from a star — a tell-tale 'Doppler wobble' — that reveals the presence of one or more smaller, unseen companions. But when it comes to finding Earth-like planets, this technique reaches a hurdle: the lack of a suitable frequency standard with which to measure the truly tiny spectral shifts caused by such very small planets. On page 610 of this issue, Li *et al.*¹ take this matter in hand, adapting the optical comb — a technology developed to measure very high frequencies — to the purposes of astronomical measurements.

The Austrian physicist and mathematician Christian Doppler first suggested in 1842 that the difference in colour observed between the two stars in some binary systems could be explained if the stars were moving in opposite directions along our line of sight. The wavelength of the light from the star moving towards us would be shifted to shorter, bluer wavelengths, and that of the star moving away from us to longer, redder wavelengths.

In this case, Doppler was wrong: such a shift is much too small for the naked eye to detect, and the different colours of binary stars actually correspond to different surface temperatures. But he was right in principle. We are well attuned to the effect now named after him when it involves sound waves from bodies moving towards and away from us: the speed of sound is

a million times slower than that of light, allowing our senses to pick up the difference in their tones. And spectrographs allow us to measure the relative displacement of the spectral lines of binary stars (Fig. 1a). The difference between the stars' velocities can easily be calculated from this wavelength shift, and plotting how this relative velocity changes with time gives the period of revolution of the stars and, with a little additional information, their relative masses and the size of their mutual orbit.

When the Doppler technique is applied to the search for extrasolar planets, only a single spectrum is measured — that of the putative parent star. If an orbiting planet is present, its gravitational attraction will at times pull the star towards us, and at times pull it away from us. The result will be a slow, periodic oscillation in the star's velocity relative to us, and an accompanying shift in its spectrum. The effect is tiny: Jupiter, the giant of our Solar System, has only one-thousandth of the mass of the Sun, and its gravity changes the Sun's radial velocity by just $\pm 13 \text{ m s}^{-1}$ in a cycle that takes almost 12 years, the period of Jupiter's orbit. Earth, which is closer to the Sun, but 300 times smaller than Jupiter, affects the Sun's velocity to the tune of just $\pm 10 \text{ cm s}^{-1}$ over the course of a year.

Planet-searching astronomers need to measure, with confidence and over timescales of years, the truly minuscule wavelength (equivalently, frequency) shifts that accord with

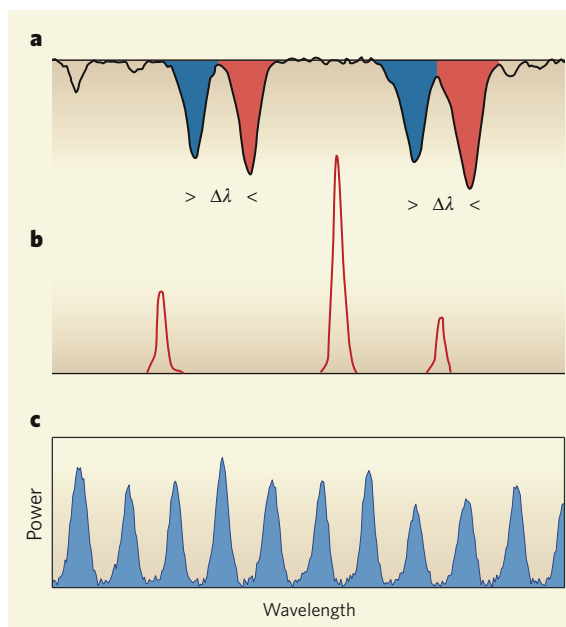


Figure 1 | Shifting stars. **a**, The motion of two stars in a binary system along our line of sight will shift the frequency of light observed by us: to longer, redder wavelengths if a star is moving away from us; to shorter, bluer wavelengths if the motion is towards us. **b**, Comparison with accurately calibrated reference lines (here from an iron arc) allows the degree of this wavelength shift, $\Delta\lambda$, to be measured. A similar measurement of the 'Doppler wobble' of a single star, induced by an unseen companion, is used in the search for extrasolar planets. **c**, Li *et al.*¹ provide a new reference standard based on a 'comb' of evenly spaced frequency lines that should allow these measurements to be made more accurately — perhaps opening up the possibility of spotting an Earth-like planet.