GENETICS

Herit-Ability

A genome-wide association study reveals possible variants that influence the complex behavior of educational attainment.

Jonathan Flint and Marcus Munafò

Studies of the genetic basis of human behavior have a history of generating controversy. For example, when studies of identical and nonidentical twins were first used to estimate the proportion of variation in income and years of schooling that can be attributed to genetic variation (in other words, the heritability of going to university and becoming rich) (*1*), one response was that the estimates were "pointless" and that determining the heritability of socioeconomic achievement measures should be abandoned (*2*). Yet, on page 1467 of this issue, Rietveld *et al*. (*3*) claim progress toward identifying genes underlying variation in educational attainment. How should these findings be interpreted, given that similar claims in the past have often not borne out (*4*, *5*)?

Behavior, like other complex traits (such as height and weight), may arise in part from the combined action of hundreds, if not thousands, of individual tiny genetic effects across the genome (*6*). Efforts to locate the contributing variants have proceeded through (unsuccessful) family studies involving the search for the coinheritance of trait and genetic markers, and through candidate gene studies (also generally unsuccessful) focused on genes thought to be involved on the basis of known or presumed neurobiology. Genome-wide association studies (GWAS) test whether a trait is associated with variation at any one or more of hundreds of thousands, or even millions, of markers. There is a consensus that for nonbehavioral traits, GWAS works well, with the caveats that it only finds a small fraction of the loci involved and does not capture all forms of genetic variation (*7*, *8*). However, with notable exceptions (*9*), GWAS of behavior has not had the success achieved for GWAS of weight, height, and diseases.

Perhaps this is because behavior is not genetically regulated like other traits. It could be that a large proportion of the predisposition to autism and schizophrenia is due to extremely rare variants, some arising de novo, which would be undetectable with GWAS methods (because these use only known, relatively common types of genetic variation) (*10*). Alternatively, the genetic effects on behavior might be even smaller and more numerous than suspected and therefore evade detection (*11*).

Achievement genes? Analyses of a genome-wide association study reveal genetic variants that are linked to educational achievement, but each with a very small effect.

The results of Rietveld *et al*. provide some answers. The contribution of each genetic locus to variation in educational attainment (defined in the study as years of schooling and the completion of college) is particularly small. The effect attributable to a locus is much smaller than for physical traits, and only just detectable with the enormous sample employed (about 126,000 individuals). The largest estimated effect (0.02%) is an order of magnitude smaller than those reported for height (0.4%) (*12*) and weight (0.3%) (*13*). However, that finding alone cannot explain the overall contribution of small effect loci. For this purpose, another observation is important.

Suppose that there are only 10 loci contributing to the genetic variation, and that they are the top 10 signals in the GWAS, each contributing 0.02%. If all 10 signals are combined to create a new predictor, one can ask how much of the variation in a new cohort the combined 10 variants explain. This is known as a polygenic score method (*14*). If the correct 10 variants are selected (and some other assumptions are made about similar allele frequencies), the polygenic score in the new population would be expected to account for about 0.2% of the variance (10×0.02). Now suppose that there are hundreds, perhaps thousands, more variants that contribute. Adding these additional variants to the 10 will increase the ability to explain variation in the trait, whereas if only the 10 variants are contributing, then nothing will be gained, however, many additional variants are combined. As shown by Rietveld *et al*., the more variants are added, the better the prediction. In short, educational attainment looks to be a very polygenic trait.

But why are the effects so small? Is this just because of the poor quality of the measurement in this, and other behavioral studies (*15*)? To some extent, Rietveld *et al*. concede this point. They argue that the GWAS effort might progress by capitalizing on the wide availability of very crude phenotypic measures, such as years of education, to identify genes robustly associated with an outcome of broad social relevance. These genes can then be further analyzed with the knowledge that they are associated with something, even if it isn't known exactly what or how.

The argument that whole-genome methods should inform candidate gene approaches (*16*) has demonstrated utility. Indeed, heaviness of smoking has been robustly associated with variation on the long arm of chromosome 15, using measures of number of cigarettes smoked that are widely available across a large number of cohorts (*9*). This phenotype can be very precisely measured, but is nevertheless a very imprecise measure of actual tobacco exposure, largely because of interindividual variability in how those cigarettes are smoked (*17*). Because nicotine consumption is most likely under the strongest genetic control (*18*), more direct measures of nicotine (and therefore tobacco) exposure provide a much more precise assessment of the strength of genetic association (*15*), while at the same time elucidating the likely mechanistic pathway.

The data from years of education make it possible for Rietveld *et al*. to test the importance of trait measurement, because educational attainment was measured in different ways, sometimes with a brief questionnaire, sometimes in more detail. There was no evidence that genetic effects are weaker in cohorts with coarser measures, which the authors argue is consistent with the view

CREDIT: K. SUTLIFF/*SCIENCE*

University of Oxford, Oxford OX3 7BN, UK. E-mail: jf@well.ox.ac.uk

that the "better" measures won't improve power to find genetic effects. However, cigarette consumption can be measured very reliably (i.e., precisely and accurately), but it still is not necessarily a reliable measure of actual tobacco exposure. Years of education can certainly be measured precisely, but it is still a noisy phenotype given the wide range of attainment and experiences that may result from the same number of years of education. Ultimately, in the case of the study by Rietveld *et al*., the question remains: What, exactly, is being measured? It seems that a genetic association has been observed for "something," but exactly what will require considerably more work. The nebulous nature of the phenotype makes this task considerably more difficult than in the case of heaviness of smoking.

A reasonable assumption is that educational attainment, and the years spent in schooling, partly reflects intellectual ability; those with a higher intelligence quotient (IQ) generally do better at school. So is this, by the backdoor, the first successful study of the genetics of IQ? That will certainly reignite some old disputes.

References

1. P. Taubman, *Am. Econ. Rev.* **66**, 858 (1976).
2. A. S. Goldberger, *Am. Econ. Rev.* **68**, 960 (1978).
3. C. A. Rietveld *et al*., *Science* **340**, 1467 (2013).
4. M. R. Munafò, C. Durrant, G. Lewis, J. Flint, *Biol. Psychiatry* **65**, 211 (2009).
5. M. R. Munafò *et al*., *Mol. Psychiatry* **14**, 119 (2009).
6. J. Yang *et al*., *Nat. Genet.* **42**, 565 (2010).
7. P. Donnelly, *Nature* **456**, 728 (2008).
8. T. A. Manolio *et al*., *Nature* **461**, 747 (2009).
9. T. E. Thorgeirsson *et al*., *Nature* **452**, 638 (2008).
10. J. McClellan, M. C. King, *Cell* **141**, 210 (2010).
11. P. M. Visscher *et al*., *Mol. Psychiatry* **17**, 474 (2012).
12. H. Lango Allen *et al*., *Nature* **467**, 832 (2010).
13. E. K. Speliotes *et al*., *Nat. Genet.* **42**, 937 (2010).
14. S. M. Purcell *et al*., *Nature* **460**, 748 (2009).
15. M. R. Munafò *et al*., *J. Natl. Cancer Inst.* **104**, 740 (2012).
16. J. Flint, M. R. Munafò, *Curr. Opin. Neurobiol.* **23**, 57 (2013).
17. A. McNeill, M. R. Munafò, *J. Psychopharmacol.* **27**, 13 (2013).
18. C. D. Fowler *et al*., *Nature* **471**, 597 (2011).

10.1126/science.1240684

CHEMISTRY

Seeing the Reaction

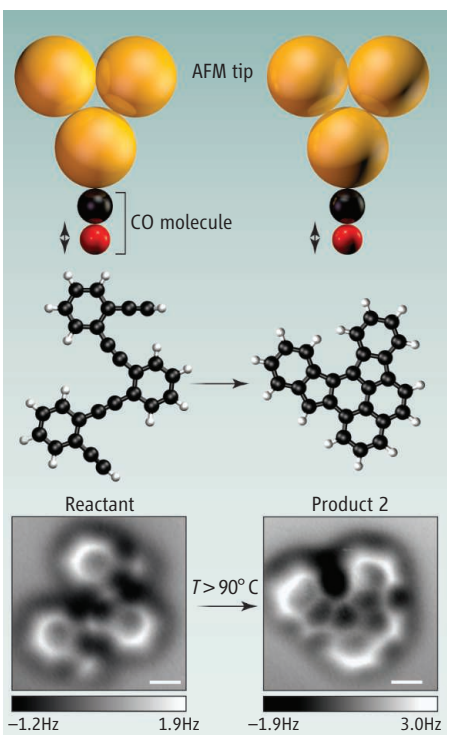
A molecule is imaged at atomic resolution as it undergoes a chemical reaction.

Franz J. Giessibl

What happens as a molecule goes through a chemical reaction? Model studies have provided important insights into these processes, but it remains extremely difficult to follow all the atomic rearrangements of a chemical reaction experimentally. In many cases, a reaction cannot be observed directly in real space, for example, because the reactants are in the gas state, zooming around at the speed of sound. On page 1434 of this issue, de Oteyza *et al*. (*1*) report atomically resolved imaging of a complex molecule as it undergoes a chemical reaction on a metal surface.

The authors investigated the molecule phenylene-1,2-ethynylene ($C_{26}H_{14}$) adsorbed on a silver surface. Upon heating above 90°C, the molecule split into several different chemical products (see the figure). The authors imaged several different products with scanning tunneling microscopy and identified them with the help of noncontact atomic force microscopy (*2*). They calculated the reaction pathway by density functional theory.

Three challenges had to be met to successfully perform this experiment. First, the authors had to find and synthesize a molecule that undergoes a reaction within the experimentally accessible temperature range, with both reactants and products remaining attached to the metal substrate surface within a viewing field accessible by the microscope.



Before and after. Schematic view of the reactant phenylene-1,2-ethynylene molecule (left) and product 2, one of three different products imaged by de Oteyza *et al*. with atomic force microscopy (AFM). Previous AFM studies focused on imaging individual molecules. Several challenges had to be overcome to extend the method to imaging both the reactant and the reaction products.

Second, they had to overcome the "fat and sticky finger problem" (*3*), which arises because the metal atoms of the probe are large relative to the hydrogen and carbon atoms in organic molecules and exert relatively large attractive forces (*4*). Meyer and co-workers have found a fine solution to this problem by attaching a carbon monoxide molecule to the end of the metal tip, which enabled unprecedented spatial resolution of pentacene, an organic molecule (*5*). This is the imaging method used by de Oteyza *et al*.

Third, the force sensor that holds the metal tip had to be sufficiently sensitive to probe the tiny forces between the carbon monoxide molecule and the organic molecule to be probed. To obtain their results, de Oteyza *et al*. had to operate their atomic force microscope in the small-amplitude mode and perform highly precise frequency measurements (*6*). Further challenges were associated with the complex sample and tip preparation and with heating the sample to initiate the reaction and then cooling it to image acquisition temperature. Overall, the work is a masterful experimental achievement.

The interpretation of image data is comparatively simple for flat hydrocarbons [such as pentacene (*5*) and phenylene-1,2-ethynylene studied by de Oteyza *et al*.], where the submolecular contrast is due to Pauli repulsion between the oxygen atom that terminates the tip and the carbon and hydrogen atoms that constitute the molecule (*7*). The flat orientation of the molecule results in a

Department of Physics, University of Regensburg, 93053 Regensburg, Germany. E-mail: franz.giessibl@physik.uni-regensburg.de