News & views

NOEL CELIS/AFP/GETTY

Figure 1 | Explaining variation in height. Yengo *et al.*[1] have performed one of the largest studies so far of genetic variants associated with height.

Genetics

Missing heritability found for height

Karoline Kuchenbaecker

A combined analysis of 281 genome-wide association studies finds 12,111 common DNA variants associated with a person's height — and shows that larger studies will not yield more variants in populations of European ancestry. **See p.704**

One of the largest human genetic studies so far has just been completed. Yengo *et al.*[1] describe on page 704 how they have used genetic data from almost 5.4 million people to identify DNA variants associated with differences in height between individuals. Remarkably, they show that they have reached the point at which more data will not reveal more variants — at least for some populations. This feat marks a milestone in our understanding of the contribution of genetics to complex traits. It also highlights the essential work still to be done to close the diversity gap in existing genetic data.

Some disorders, such as Huntington's disease, are caused by changes in one or only a few specific genes. Once the genes have been identified, the cause of the disorder is seen as resolved (although greater genetic complexity can be involved[2]). Genome-wide association studies (GWAS) were designed in 2005 to identify the genetic causes of more-complex diseases[3]. GWAS use an approach called genotyping to compare the DNA of thousands of individuals, enabling researchers to identify genetic variants associated with differences in a trait of interest. Early GWAS typically involved a few thousand participants[4]; these and subsequent studies identified many disease-associated variants and helped to uncover the biology underlying these associations[5].

However, variants from early GWAS could explain only 5% of the variation in height between people (Fig. 1), even though height is determined mainly by genetic factors. This finding sparked a debate about the 'missing heritability' in GWAS[6]. Hypotheses to explain the mismatch included the existence of rare variants not detected by GWAS, gene–gene interactions and overestimation of heritability estimates[7]. It emerged that missing heritability arose in part because many unidentified genetic variants have small effects that contribute to a complex trait[4] — some are so tiny that it seemed unlikely that researchers could ever gather enough samples to detect them. Consequently, a full picture of the genetics underlying any complex trait seemed out of reach.

However, technological advances have resulted in genotyping of millions of human genomes across thousands of GWAS[8]. Yengo *et al.* therefore asked whether a large enough sample size could reveal the missing heritability for height. In an immense collaborative effort, they combined data from 281 studies, predominantly including participants of European ancestry.

The researchers found 12,111 genetic variants associated with human height, which clustered into small genomic segments that cover 21% of the genome. This confirms earlier predictions[9] that a large proportion of the human genome is involved in shaping height. The heritability that can potentially be explained by all genetic variants detectable by GWAS has been estimated[10] to be about 50% — and indeed,

the group showed that the genomic sections they had identified could explain 50% of the variation in height.

Hence, the missing heritability of height has been found. Further genetic studies in people of European ancestry will not contribute any more information — 'saturation' has been achieved for this trait in this ancestry group. This astonishing result will now motivate efforts to achieve saturation for other traits and diseases. Once achieved, attention might shift more widely to the momentous task of identifying the genes that underlie these associations, and inferring mechanisms for some of them.

Efforts to achieve saturation for complex diseases will face challenges. Height is a relatively stable trait that can be measured easily and reliably. By contrast, it can be difficult to determine how to measure outcomes for complex diseases such as major depressive disorder. Low disease prevalence can make it hard to generate large-enough sample sizes, and some diseases might be affected by more genes than is the case for height[11]. But concerted international efforts to overcome these challenges are a likely outcome of Yengo and colleagues' work.

It is important to acknowledge, as the authors do, that saturation is currently limited to only one ancestral group. Yengo *et al.* used existing data sets, in which diversity is limited. Out of the 5.4 million samples that were included in the study, 4 million were from participants of European ancestry; fewer than 78,000 were from those of South Asian ancestry.

There is also a lack of geographical diversity. Among the studies contributing data from participants of African descent, to the best of my knowledge, only one is from the African continent, with the rest coming from the diaspora. This single study, conducted in Nigeria, had 1,188 participants, implying that only 0.4% of the participants of African descent included in the GWAS were from Africa. The genetic and linguistic diversity of Africa is immense[4,12,13]. Although participants from the African diaspora add valuable data, these groups do not capture Africa's diversity.

Thus, this study, like other GWAS, did not comprehensively cover human genetic variation[14]. Some of the genetic variants associated with height will have been missed. Yengo *et al.* showed that, for participants of African ancestry from the United States and the United Kingdom, the identified genetic variants could explain only 5–12% of variation in height (in comparison to 40% in people of European ancestry). This is in line with other research demonstrating that GWAS based on people of European descent can less accurately predict gene–trait relationships in other groups[15]. Moreover, a study published earlier this year demonstrated that, across different sub-Saharan African populations, the performance of genetic scores generated from GWAS of African Americans varied widely[16], suggesting that environment, as well as ancestry, can affect genetic associations. Understanding how environmental factors might affect gene–trait relationships[17] will require more global diversity in genetic research.

Yengo *et al.* have demonstrated that it is possible to achieve saturation for complex traits. Now, ancestrally, ethnically, globally and socio-economically diverse samples are needed to reap the full benefits of GWAS.

Karoline Kuchenbaecker is in the Division of Psychiatry and the UCL Genetics Institute, University College London, London WC1E 6BT, UK.
e-mail: k.kuchenbaecker@ucl.ac.uk

1. Yengo, L. *et al. Nature* **610**, 704–712 (2022).
2. Barton, A. R., Hujoel, M. L. A., Mukamel, R. E., Sherman, M. A. & Loh, P.-R. *Am. J. Hum. Genet.* **109**, 1298–1307 (2022).
3. Klein, R. J. *et al. Science* **308**, 385–389 (2005).
4. The Wellcome Trust Case Control Consortium. *Nature* **447**, 661–678 (2007).
5. *Nature* **456**, 18–21 (2008).
6. Visscher, P. M. *et al. Am. J. Hum. Genet.* **101**, 5–22 (2017).
7. Manolio, T. A. *et al. Nature* **461**, 747–753 (2009).
8. Loos, R. J. F. *Nature Commun.* **11**, 5900 (2020).
9. Shi, H., Kichaev, G. & Pasaniuc, B. *Am. J. Hum. Genet.* **99**, 139–153 (2016).
10. Yengo, L. *et al. Hum. Mol. Genet.* **27**, 3641–3649 (2018).
11. Zhang, Y., Qi, G., Park, J.-H. & Chatterjee, N. *Nature Genet.* **50**, 1318–1326 (2018).
12. Zeggini, E. & Morris, A. (eds) *Assessing Rare Variation in Complex Traits: Design and Analysis of Genetic Studies* (Springer, 2015).
13. Fan, S., Hansen, M. E. B., Lo, Y. & Tishkoff, S. A. *Science* **354**, 54–59 (2016).
14. Fatumo, S. *et al. Nature Med.* **28**, 243–250 (2022).
15. Martin, A. R. *et al. Nature Genet.* **51**, 584–591 (2019).
16. Kamiza, A. B. *et al. Nature Med.* **28**, 1163–1166 (2022).
17. Giannakopoulou, O. *et al. JAMA Psychiatry* **78**, 1258–1269 (2021).

The author declares no competing interests.
This article was published online on 12 October 2022.

Chemical synthesis

A leap forward in the quest for general catalysts

Manuel J. Scharf & Benjamin List

Truly general chemical reactions work well regardless of the structural features and functional groups in the starting molecule. A new screening protocol speeds up the identification of such reactions in the field of asymmetric catalysis. **See p.680**

Whether a newly developed chemical reaction finds its way into the chemist's toolbox for synthesis depends on various factors, such as the convenience of the experimental set-up and the toxicity of the reagents and by-products. Most importantly, an ideal reaction should be general — it must generate the expected products reliably for a wide range of previously unused substrate molecules. However, the generality of a reaction typically becomes apparent only after it has been used reproducibly in a variety of syntheses. On page 680, Wagen *et al.*[1] present a protocol for screening reactions that might accelerate the development of general catalytic methods for asymmetric synthesis, the field of chemistry that aims to make single mirror-image isomers of organic compounds.

Synthetic chemists strive to make complex molecules in the most efficient manner. Methods are therefore constantly being developed to enable the formation of previously unattainable chemical bonds and to prepare products in high yields. A subfield of chemical synthesis deals with a more intricate problem. Some molecules possess a property called chirality, which means that they can form as two possible isomers that are mirror images of each other. These isomers — known as enantiomers — have the same physical properties, but behave very differently when placed in chiral environments, such as the binding pockets of proteins in our bodies. In the case of pharmaceutical compounds, only one of the enantiomers might have the desired biological activity; in the worst-case scenario, the other enantiomer is harmful[2].

Medicinal chemists therefore need to be able to make enantiopure compounds — that is, single enantiomers — for biological testing. This is not straightforward, because most reactions generate chiral molecules as a one-to-one mixture of enantiomers. Enantiopure compounds can sometimes be made by starting a chemical synthesis from a single-enantiomer building block found in nature — most biomolecules, such as amino acids and sugars, are produced as single enantiomers. However, this approach is restricted by the limited set of enantiopure compounds found in biological systems[3].

Another strategy, known as asymmetric catalysis, is to use a chiral, enantiopure catalyst to transform a substrate into a chiral product that exists predominantly in one mirror-image