

## GENETIC VARIATION

# Searching for solutions to the missing heritability problem

**Rare genetic variants in yeast explain a large amount of phenotypic variation in a complex trait like growth.**

LUISA F PALLARES

**Related research article** Bloom JS, Boockock J, Treusch S, Sadhu MJ, Day L, Oates-Barker H, Kruglyak L. 2019. Rare variants contribute disproportionately to quantitative trait variation in yeast. *eLife* **8**:e49212. DOI: [10.7554/eLife.49212](https://doi.org/10.7554/eLife.49212)

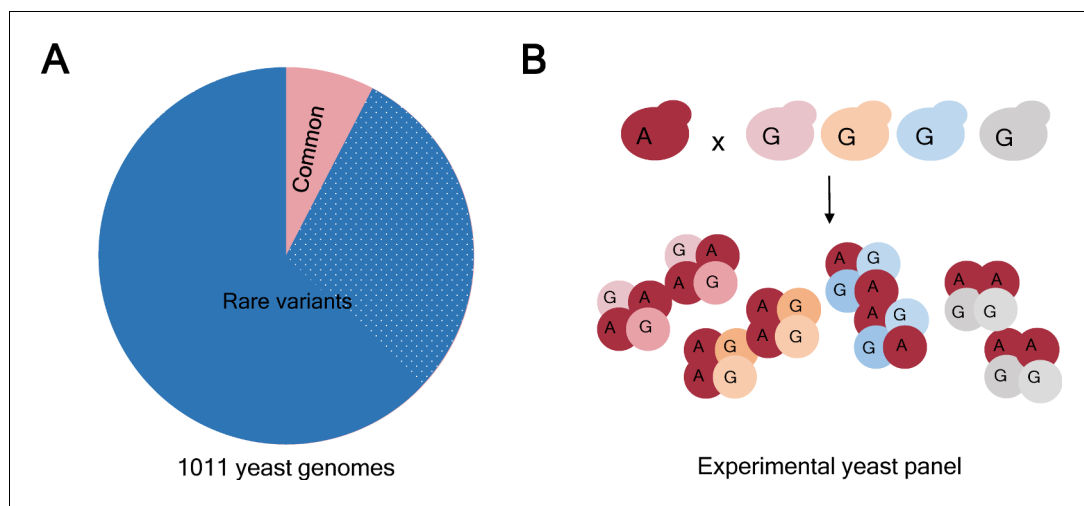
**Related research article** Fournier T, Abou Saada O, Hou J, Peter J, Caudal E, Schacherer J. 2019. Extensive impact of low-frequency variants on the phenotypic landscape at population-scale. *eLife* **8**:e49258. DOI: [10.7554/eLife.49258](https://doi.org/10.7554/eLife.49258)

Although most of the 3000 million nucleotides in the human genome are the same in every person on the planet, there are about 90 million sites that can vary between individuals. The source of all phenotypic variation in humans lies in these 90 million genetic variants, and in their interactions with each other and with the environment. Identifying the genetic variants that are involved in a specific trait (such as height or disease status) is a long-standing goal in biology.

Today, researchers rely on genome-wide association studies (GWAS) to find the genetic variants that are relevant to a specific trait. In GWAS the genomes of individuals are analyzed to see if particular genetic variants are correlated with variation in traits of interest. GWAS results have identified hundreds of variants underlying phenotypic variation in humans,

mice, fruit flies, rice, maize, and many other taxa. Yet, despite the large number of alleles that have been identified using this technique, the amount of phenotypic variation they explain is just a fraction of what twin and pedigree studies predict is heritable. For example, twin studies have shown that approximately 80% of variation in human height can be explained by genetic factors (*Silventoinen et al., 2012*). However, the results of the best powered GWAS only explain around 20% of such variation (*Wood et al., 2014*). This gap is known as the ‘missing heritability problem’.

Rare and low-frequency genetic variants (which have allele frequencies of <1% and <5% respectively) have been proposed as one explanation for the missing heritability problem (reviewed in *Gibson, 2012*). Such variants are routinely excluded from GWAS studies because when an allele is present in few individuals, the statistical analysis used to draw correlations between traits and alleles is not powerful enough to obtain significant results. As a consequence around 90% of genetic variation in humans and other organisms like yeast has so far gone unexplored (*Figure 1A; Auton et al., 2015; Peter et al., 2018*). The missing heritability might be hiding in plain sight, but until now, studying the effect of rare alleles on the variation of traits influenced by more than one gene was extremely challenging. Now, in eLife, two independent groups report the results of experiments on yeast which show that rare variants have a fundamental role in phenotypic variation at the population level.



**Figure 1.** Allele frequency in natural isolates of yeast and in the experimental populations. (A) Based on a study of 1011 genomes it is known that 93% of the genetic variants in the yeast *Saccharomyces cerevisiae* are rare (that is, they have a frequency <1%; blue). Moreover, just over 508000 variants (31% of the total; dotted blue) were found in just 1 of the 1011 genomes studied. However, genome-wide association studies (GWAS) tend to focus on the 7% of genetic variants that are common (that is, have a frequency >1%; pink). (B) The frequency of a rare allele can be increased by crossing a yeast isolate carrying the rare variant with an isolate with the alternative (more common) variant. To obtain a variety of isolates with a specific rare allele in different genetic backgrounds, the isolate carrying the rare variant (allele A, dark red) can be crossed with several different isolates with the alternative allele (allele G, pink, yellow, blue, grey). As a result, allele A is more frequent in the experimental panel than in the parental isolates, making it suitable for GWAS analysis. Importantly, regardless of the frequency that any allele reaches in the experimental panel, the real natural frequency can be looked up in the collection of 1011 yeast genomes (panel A).

In a monumental effort, the two groups independently selected a set of wild and domesticated yeast isolates from all over the world and crossed them to generate a genetically diverse panel of thousands of strains (Figure 1B). They then exposed each cross to more than 35 different media conditions and quantified their growth by measuring colony size. As a result of the crossing scheme, genetic variants that were present in just one or a few yeast isolates were now present in hundreds of samples in the experimental panels (Figure 1B). This allowed the groups to include a large number of rare variants (up to 28% of the total) in the GWAS analysis: many of these variants would have been excluded from traditional GWAS studies due to their low allele frequency.

Both groups independently identified thousands of genetic variants associated with growth, and estimated that over half of growth variance can be attributed to additive effects. To determine how variants with different frequencies contributed to phenotypic effects, variants were classified into either rare (<1%) and common (>1%) (Bloom et al., 2019), or rare (<1%), low frequency (1–5%) and common (>5%)

(Fournier et al., 2019). This classification was based on 1011 yeast genomes that represent global yeast diversity (Figure 1A; Peter et al., 2018). Strikingly, rare variants contributed a disproportionate amount to phenotypic variation in both studies.

In one study Joseph Schacherer and co-workers at the University of Strasbourg – including Téo Fournier as first author – found that 16% of the GWAS results were rare alleles even when they made up just 4% of all the variants used in the experiments (Fournier et al., 2019). In the other study Joshua Bloom, Leonid Kruglyak and colleagues at UCLA estimated that over half of the observed growth variation can be explained by rare variants, even when they represented only 28% of the variants used (Bloom et al., 2019). The UCLA team also found that the rare variants detected in GWAS tend to have larger effect sizes than common variants, tend to reduce growth ability, and tend to have arisen more recently in evolutionary time.

These results join recent efforts exploring the effect of rare variants on complex traits. For human height it has been shown that rare variants have effect sizes ten times larger than

common variants (*Marouli et al., 2017*), and that together they account for most of the missing heritability in this trait (*Wainschtein et al., 2019*). In parallel, it was estimated that at least a quarter of gene expression heritability in humans is accounted for by rare variants (*Hernandez et al., 2019*). The fact that in humans, as well as yeast, the contribution of rare variants to complex traits is now beyond doubt suggests that it may be the same in other species. However, addressing this question in organisms with larger genomes and not amenable to crossing schemes remains challenging. But rest assured, researchers will find a way.

**Luisa F Pallares** is at the Lewis-Sigler Institute for Integrative Genomics and the Department of Ecology and Evolutionary Biology, Princeton University, Princeton, United States  
pallares@princeton.edu

 <https://orcid.org/0000-0001-6547-1901>

**Competing interests:** The author declares that no competing interests exist.

**Published** 04 December 2019

## References

- Auton A**, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR, 1000 Genomes Project Consortium. 2015. A global reference for human genetic variation. *Nature* **526**:68–74. DOI: <https://doi.org/10.1038/nature15393>, PMID: 26432245
- Bloom JS**, Boocock J, Treusch S, Sadhu MJ, Day L, Oates-Barker H, Kruglyak L. 2019. Rare variants contribute disproportionately to quantitative trait variation in yeast. *eLife* **8**:e49212. DOI: <https://doi.org/10.7554/eLife.49212>, PMID: 31647408
- Fournier T**, Abou Saada O, Hou J, Peter J, Caudal E, Schacherer J. 2019. Extensive impact of low-frequency variants on the phenotypic landscape at population-scale. *eLife* **8**:e49258. DOI: <https://doi.org/10.7554/eLife.49258>, PMID: 31647416
- Gibson G**. 2012. Rare and common variants: twenty arguments. *Nature Reviews Genetics* **13**:135–145. DOI: <https://doi.org/10.1038/nrg3118>, PMID: 22251874
- Hernandez RD**, Uricchio LH, Hartman K, Ye C, Dahl A, Zaitlen N. 2019. Ultrarare variants drive substantial cis heritability of human gene expression. *Nature Genetics* **51**:1349–1355. DOI: <https://doi.org/10.1038/s41588-019-0487-7>, PMID: 31477931
- Marouli E**, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, Fine RS, Lu Y, Schurmann C, Highland HM, Røe S, Thorleifsson G, Justice AE, Lamparter D, Stirrups KE, Turcot V, Young KL, Winkler TW, Esko T, Karaderi T, et al. 2017. Rare and low-frequency coding variants alter human adult height. *Nature* **542**:186–190. DOI: <https://doi.org/10.1038/nature21039>, PMID: 28146470
- Peter J**, De Chiara M, Friedrich A, Yue JX, Pflieger D, Bergström A, Sigwalt A, Barre B, Freel K, Llored A, Cruaud C, Labadie K, Aury JM, Istace B, Lebrigand K, Barbry P, Engelen S, Lemainque A, Wincker P, Liti G, et al. 2018. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* **556**:339–344. DOI: <https://doi.org/10.1038/s41586-018-0030-5>, PMID: 29643504
- Silventoinen K**, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, de Lange M, Harris JR, Hjelmborg JVB, Luciano M, Martin NG, Mortensen J, Nisticò L, Pedersen NL, Skytthe A, Spector TD, Stazi MA, Willemsen G, Kaprio J. 2012. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Research* **6**:399–408. DOI: <https://doi.org/10.1375/136905203770326402>
- Wainschtein P**, Jain DP, Yengo L, Zheng Z, Cupples LA. 2019. Recovery of trait heritability from whole genome sequence data. *bioRxiv*. DOI: <https://doi.org/10.1101/588020>
- Wood AR**, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan Jian'an, Kutalik Z, Amin N, Buchkovich ML, Croteau-Chonka DC, Day FR, Duan Y, Fall T, Fehrmann R, Ferreira T, Jackson AU, Karjalainen J, et al. 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics* **46**:1173–1186. DOI: <https://doi.org/10.1038/ng.3097>