nature genetics

PERSPECTIVE

https://doi.org/10.1038/s41588-022-01103-1

Check for updates

From Mendel to quantitative genetics in the genome era: the scientific legacy of W. G. Hill

Brian Charlesworth[1,8], Michael E. Goddard[2,3,8], Karin Meyer[4,8], Peter M. Visscher[5,8] ✉, Bruce S. Weir[6,8] and Naomi R. Wray[5,7,8]

The quantitative geneticist W. G. ('Bill') Hill, awardee of the 2018 Darwin Medal of the Royal Society and the 2019 Mendel Medal of the Genetics Society (United Kingdom), died on 17 December 2021 at the age of 81 years. Here, we pay tribute to his multiple key scientific contributions, which span population and evolutionary genetics, animal and plant breeding and human genetics. We discuss his theoretical research on the role of linkage disequilibrium (LD) and mutational variance in the response to selection, the origin of the widely used LD metric r^2 in genomic association studies, the genetic architecture of complex traits, the quantification of the variation in realized relationships given a pedigree relationship and much more. We demonstrate that basic theoretical research in quantitative and statistical genetics has led to profound insights into the genetics and evolution of complex traits and made predictions that were subsequently empirically validated, often decades later.

The rediscovery of Mendel's laws in 1900 resulted in new scientific disciplines, including evolutionary and population genetics and, after the reconciliation of continuous ('complex') traits with Mendelian inheritance[1,2], quantitative genetics. Mendelian genetics gave a mechanism for Darwin's theory of natural selection, underlies the risk of rare and common disease in families and populations and is the foundation for efficient breeding programs in agriculture. Not many scientists have contributed to all these fields of research and their applications. One exception is W. G. (Bill) Hill, who died 17 December 2021. In this Perspective, we highlight key scientific discoveries and contributions made by Hill, summarized in Fig. 1, which all ultimately derived from Mendel's laws of assortment and segregation.

LD in finite populations and the metric r^2

When Mendel documented cosegregation of pairs of traits in hybridization studies of peas, he reported on trait pairs determined by unlinked genes, so that their segregation was independent[3,4]. Linkage between loci was soon recognized by geneticists, and theoretical studies introduced the concept of LD, denoted Δ by Robbins[5] and now known as D (ref. [6]). For a pair of loci, each segregating for a pair of alleles, A/a and B/b, respectively, with frequencies p_A versus $1-p_A$, and p_B versus $1-p_B$, D is defined as $p_{AB}-p_Ap_B$, where p_{AB} is the frequency of haplotypes carrying A and B and p_Ap_B is their frequency expected if the alleles at the two loci are combined randomly. It was Hill together with Alan Robertson[7] who introduced the r^2 measure of LD. r^2 is the squared correlation between the states of the alleles at the two loci, measured as $r^2=\frac{D^2}{p_A(1-p_A)p_B(1-p_B)}$ and has advantages over other measures of LD because of its range (0 to 1) and known statistical properties. The r^2 statistic is a metric that has shaped much of empirical population genetics research since the 1970s and human genetics research in the last 20 years, as evidenced by citations over time (Fig. 2).

By 1918, it was already recognized that, in a large randomly mating population, LD is broken down by recombination between the two loci (whose frequency is denoted by c), so that after t generations, $D_t=D_0(1-c)^t$ tends to zero as t increases, where D_0 is the value of D at time $t=0$. More than 50 years of theoretical studies followed that described LD under different mating and selection scenarios, all under the assumption of infinite population size[8]. It was Hill and Robertson in 1966[9] and 1968[7] who first investigated the impact of random sampling of genotype frequencies in finite populations (genetic drift) on LD, concluding that drift can induce (substantial) LD even when loci are on different chromosomes, if the population size is sufficiently small. Moreover, they demonstrated that the expected value of r^2 was inversely related to $4N_ec$, where N_e is the effective population size[7]. A few years later, Sved derived this relationship as $E(r^2)\approx 1/(1+4N_ec)$, which became a well-known expression[8]. The r^2 measure is fundamental to the design and analysis of genome-wide association studies (GWASs), because it determines how many markers are needed in the genome to detect association between a genotyped marker and an ungenotyped causal variant[10]. Single-nucleotide polymorphisms (SNPs) on GWAS arrays are selected to be representative of variation in the genome using the estimate of the r^2 statistic, and the power of the association test is directly proportional to the r^2 between a causal variant and the genotyped SNP. Methods for estimating genetic (co) variance from GWAS results also rely on the r^2 metric[11].

Hill also considered the genetic sampling variation of LD[12]. In an algebraic tour-de-force, he explored the properties of the squared correlation by determining the variances and covariances of squared disequilibria[13]. Finding tractable approximations for eighth-order moments is a rare skill and was matched only by Hill's skill in the

[1]School of Biological Sciences, Institute of Ecology and Evolution, University of Edinburgh, Edinburgh, UK. [2]Centre for AgriBioscience, Department of Economic Development, Jobs, Transport and Resources, Bundoora, Victoria, Australia. [3]Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, Victoria, Australia. [4]AGBU, a joint venture of NSW Department of Primary Industries and University of New England, Armidale, New South Wales, Australia. [5]Institute for Molecular Bioscience, The University of Queensland, St Lucia, Queensland, Australia. [6]Department of Biostatistics, University of Washington, Seattle, WA, USA. [7]Queensland Brain Institute, The University of Queensland, St Lucia, Queensland, Australia. [8]These authors contributed equally: Brian Charlesworth, Michael E. Goddard, Karin Meyer, Peter M. Visscher, Bruce S. Weir, Naomi R. Wray. ✉e-mail: peter.visscher@uq.edu.au

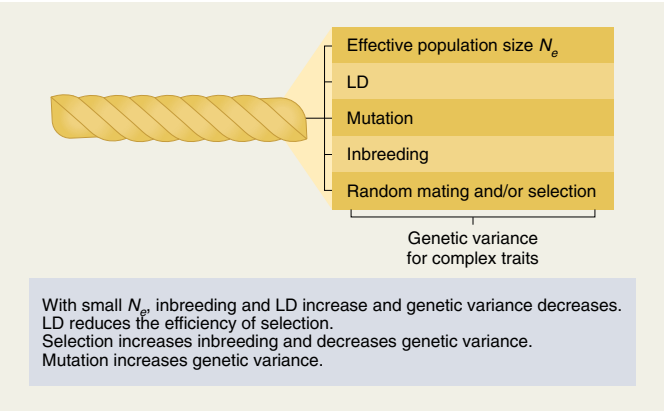

Fig. 1 | Five intertwined strands that shape genetic variance for complex traits within populations. Bill Hill's legacy encompasses a breadth of theoretical papers that address these different threads and their dependencies.

simulation and numerical evaluation of his theoretical results. In a highly cited paper published in 1974 (ref. [14]), Hill used 'chromosome counting' for estimating LD by estimating the numbers of coupling and repulsion heterozygotes at two loci and using these in maximizing the likelihood for two-locus gamete frequencies, which in turn were used to update heterozygote frequencies in an iterative process. This preceded the now-common and equivalent expectation–maximization (EM) algorithm (which was not named and published until 1977), which is widely used for maximum likelihood estimation of LD[15].

The term LD can be confusing. In his 'One hundred years of linkage disequilibrium' review coauthored with John Sved[8], Hill offers some key historical insights: "From a present-day point of view, it is difficult to appreciate the background of population genetics theory in the premolecular era. It was well known, from Drosophila for example, that there are many cases of very closely linked loci. What was less clear, however, was whether there are many cases of closely linked polymorphic loci in populations. In retrospect, the lack of thought given to this possibility seems surprising … One result of this history is the usage of the term LD. In modern usage it usually applies to closely linked loci, where the idea that linked SNPs within linkage blocks are somehow in 'disequilibrium' seems counterintuitive. The LD term is also used to describe the situation for unlinked loci" (e.g., correlation across chromosomes induced by population stratification) "where the term is especially inappropriate. In retrospect, the term 'allelic association' (see, e.g., Morton et al.[16]) would probably have been more suitable."

The Hill–Robertson effect

For evolutionary biologists, Hill's most famous contribution is his PhD work on the effects of linkage on the effectiveness of selection. This work showed that linkage between different loci subject to directional selection impedes the ability of the population to respond to selection, because randomly generated LD can cause a beneficial variant at one genomic location to become associated with a harmful variant nearby in the genome, reducing its chance of fixation in the population. This was dubbed the Hill–Robertson effect by Joe Felsenstein in his seminal paper on the evolutionary significance of genetic recombination[17]. The effect is also known as Hill–Robertson inference. The importance of the Hill–Robertson effect in explaining patterns of molecular evolution and variation in relation to the rate of genetic recombination has become increasingly apparent with the advent of population genomic studies, with increasing evidence that the ability of populations to incorporate beneficial mutations and eliminate deleterious mutations is impeded when recombination is rare or absent[18]. The large increase in citations over the past 20 years shown in Fig. 2 reflects the importance of the concept of Hill–Robertson inference for studies of molecular evolution and variation.

Population genetics theory

Hill contributed substantially to our understanding of the variation among evolutionary replicates of population genetic parameters, and variation in the estimates of these parameters. The key population genetic parameters needed to understand the behavior of neutral alleles between and within populations are the probabilities that sets of alleles are identical by descent because they have descended from a single ancestral allele. Genetic sampling over the generations between that ancestral allele and the current alleles means that there is variation in the actual identity state across the genome and among individuals with the same pedigree-based identity probabilities. In an important note[19], Hill predicted the extent of this variation for genomes with specified numbers of chromosomes and map lengths. This followed substantial treatments of variation in heterozygosity[20] and inbreeding[21] in finite populations. Variation in actual relatedness identity was covered in multiple papers[22–24]. Figure 3 (which is based on Fig. 5 of Hill and Weir[22]) illustrates Hill's observation of the difficulty in distinguishing among different classes of relatedness on the basis of single-marker statistics even when the predicted identity probabilities are different. Although the variances of actual identity decrease with the pedigree expectations, the coefficients of variation increase as relatedness decreases. Distinguishing among classes becomes more difficult for more distant relatives. In work that parallels methods currently used by direct-to-consumer testing companies and forensic investigations of cold cases, Hill later[23] showed that inference on relatedness classes can be made using the number and genomic locations of shared identical-by-descent segments.

Genetic architecture of complex traits

Long-standing questions in quantitative genetics concern the nature of genetic variation. How many loci contribute to standing variation, what are their frequency and effect sizes, how much of the genetic variance that they generate is additive versus nonadditive and why is genetic variation so ubiquitous, including for traits that are associated with fitness? Hill contributed to all of these fundamental questions in a series of theoretical studies. As we show, many of the predictions made in those studies have now been validated empirically.

The standard model for a single locus and a quantitative (complex) trait is the Fisher model, derived in detail in Fisher[1] and popularized in Falconer's well-known textbook[25]. This model relates descriptors of variation (additive and dominance variance components) to gene action (additive and dominance coefficients) and allele frequencies and reconciles the correlations between relatives for complex traits with Mendel's laws of single-locus inheritance. For a recent detailed explanation of this model with an online web application, see Hivert et al.[26]. Multilocus (polygenic) models generate additional (epistatic) genetic variance components[27]. For multilocus models, if the dominance effects at individual loci are on average in one direction ('directional dominance'), there will be a change of the mean phenotype with inbreeding. In particular, 'inbreeding depression', whereby inbred individuals have reduced mean trait values for traits associated with fitness (such as survival and fertility), is ubiquitous in nature[28].

Knowledge of genetic variance components is important, because they are informative about the nature of new mutations, the nature and strength of past natural selection, predicted responses to artificial and natural selection and the optimum experimental designs for mapping trait loci. Dominance variance was considered

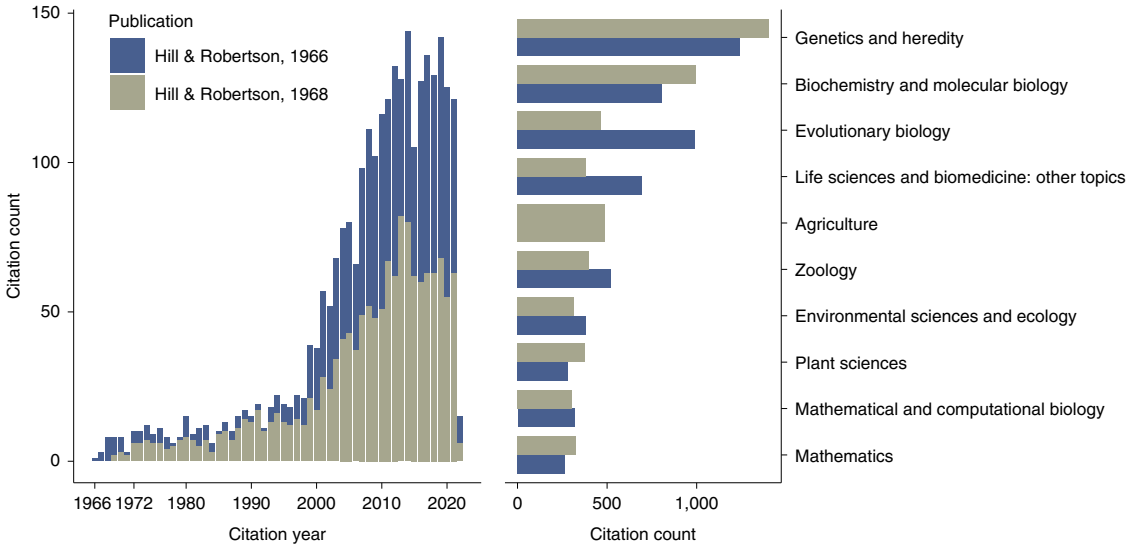

Fig. 2 | Citations of two key Hill papers on LD. There were 2,718 citations in total, 1,287 to Hill and Robertson[9] and 1,431 to Hill and Robertson[7]. Citations over time (left) and across disciplines (right). Source: Web of Science, accessed 2 April 2022.

to be important by Fisher because of Mendel's experiments and because it causes the genetic correlation between full sibs to exceed that between parents and offspring. Hill made predictions about the role of additive versus nonadditive genetic variation in a number of theoretical studies. First, in a throwaway comment in Robertson & Hill[29], which concerns the effect of linkage on the response to directional selection in finite populations, the authors state "A rather surprising consequence of these arguments is that it is not possible to define an infinitesimal model with directional dominance in which there is linkage equilibrium and both the inbreeding depression and dominance variance are finite." In other words, highly polygenic traits cannot both show inbreeding depression and dominance variance. The relationship between inbreeding depression, dominance variance and the number of loci was in fact highlighted earlier in analyses of *Drosophila* experiments[30]. We illustrate the relationships among inbreeding depression, dominance variance and polygenicity in Fig. 4. It shows that dominance variance is expected to be negligible, unless the number of causal loci is small (<100) and inbreeding depression is very large (e.g., 0.5 standard deviations per 6.25% inbreeding). This would correspond in humans to a reduction of ~3.5 cm in height and 7.5 IQ points in offspring of first cousin matings. The estimated inbreeding depression reported for human traits is typically less than 0.25 standard deviations per 6.25% inbreeding[31–33].

Second, Hill presented theoretical results on the expected ratio of additive to total genetic variance by modeling the distribution of allele frequencies under a range of gene action models. He concluded that additive genetic variance is expected to account for most of the genetic variation of complex traits[34]. Third, Hill then expanded this theoretical research to include higher-order epistatic interactions and again concluded that most genetic variance will be additive[35]. Fourth, Hill and coworkers investigated highly nonlinear biological models (enzyme flux) that display strong dominance coefficients and again showed that most genetic variance is expected to be additive[34,36]. These results all have in common the property that strong interactions of loci within and between genes do not tend to result in much nonadditive genetic variation.

How do the data fit these theoretical predictions? We focus here on human complex traits, even though the results apply to other outbred populations such as *Drosophila* and were indeed obtained there much earlier in the context of viability and other

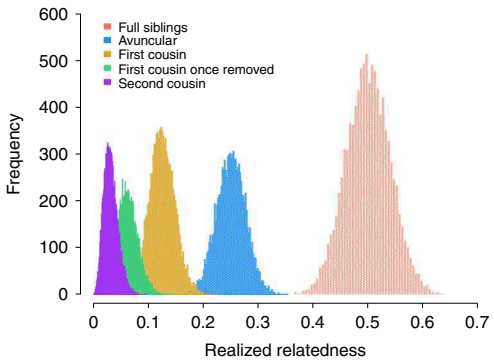

Fig. 3 | Distribution of the proportion of the genome shared identical-by-descent in a selected number of relative pairs. Based upon Figure 5 in Hill and Weir[22], using their approximation of the genome length in humans.

fitness components[30,37]. In human genetics, twin studies have been used extensively over many decades to estimate genetic variation for quantitative traits and diseases and partition it into additive and dominance components of variance. A review of nearly all twin studies in the last 50 years concluded that on average (across traits), results from twin studies were consistent with mostly additive genetic variation and an absence of nongenetic sources of twin similarity[38]. Estimation of inbreeding depression can be done easily by correlating trait values with inbreeding coefficients, either with pedigree data or with population data using GWASs. The results provide widespread evidence for inbreeding depression[31–33]. Moreover, we also know from GWASs that most complex traits are highly polygenic[39]. It follows, therefore, that there cannot be much dominance variance for human complex traits. Consistent with this purely theoretical prediction, the absence of dominance variance has recently been reported for many human traits[40–42]. Even in model organisms such as yeast, where gene frequencies in experimental line crosses are ½, which is the best-case scenario for observing nonadditive variance, most variation is additive[43,44].

As Darwin recognized, a change in mean trait value under selection requires the existence of heritable variation, which we now understand means additive genetic variance, which is abundantly

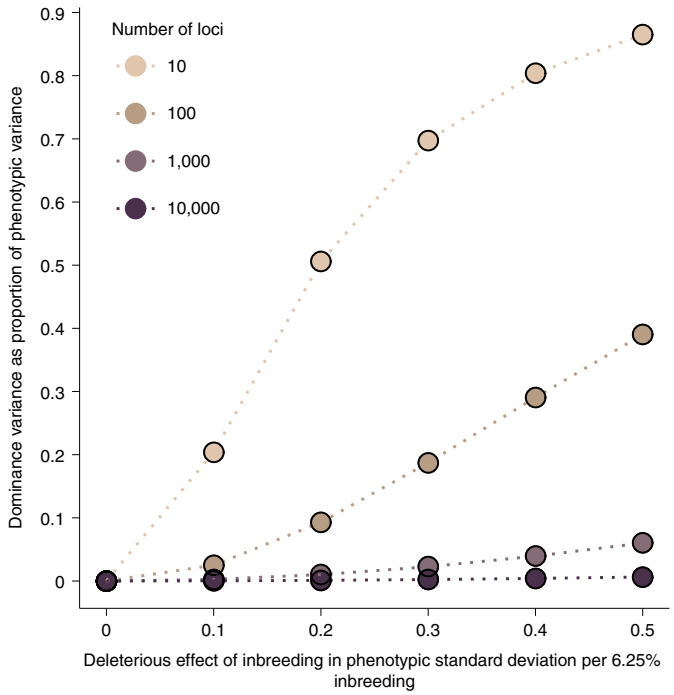

Fig. 4 | Relationships among polygenicity, inbreeding depression and dominance variance. Dominance variance (*y*-axis) is given as a function of the number of trait loci (10, 100, 1,000 and 10,000) and the effect of inbreeding depression on the trait (*x*-axis).

available for most quantitative traits, as we have just described. Hill contributed several important theoretical studies of how additive genetic variance can be maintained in populations as the result of the interactions among various forms of selection, mutation and genetic drift (e.g., ref. [45]). But selection of the intensity used in artificial selection experiments, and by animal and plant breeders, is expected to rapidly exhaust additive genetic variance. Hill was the first to quantify the role of new mutations in producing the variability needed for long-term continued responses to selection[46,47]. These responses are a remarkable feature of many artificial selection experiments and animal and plant breeding programs, as he was fond of pointing out[48].

Translation and impact of quantitative genetics theory in animal breeding

Animal breeding is the area of applied quantitative genetics dealing with selection of animals (mainly livestock) to improve their genetic potential for productivity. Among other tasks, this involves the design of breeding programs, the prediction of genetic ('breeding') values from available data and estimation of genetic covariances and other parameters. The last two are also relevant for evolutionary[49] and human genetic studies; polygenic risk predictors for disease in humans are equivalent to estimated breeding values in livestock[50]. Coming from a farming background, Hill took a keen interest in these areas and had a profound impact, not only through his own contributions but also by using them as training ground for numerous students (including authors K.M., P.M.V. and N.R.W.) and postdoctoral students and serving as a consultant for various breeding companies. We consider selected topics.

Some of Hill's earlier work addressed the response to selection in breeding programs when generations overlap. He advocated the use of 'transition matrices' to model gene flow in predicting genetic gain[51]. This elegant and versatile method has practical applications because it allows the economic optimization of breeding programs by predicting genetic gain over time while discounting future returns. Furthermore, he demonstrated the effects of sampling variation on the efficacy of selection indices, that is, linear combinations of information from various sources such as different traits or traits measured on relatives. Sales and Hill[52,53] showed that, due to sampling errors in estimates of genetic parameters the achieved response to selection is always less than expected. Subsequently, Hayes and Hill[54] proposed a method termed 'bending' to reduce the sampling variation of estimates of genetic covariance matrices by regressing eigenvalues to their mean. The original idea was to 'borrow strength' from the phenotypic covariance matrix by considering the canonical eigenvalues of the two matrices. Although their paper is often misquoted (it involved the product of two covariance matrices), bending generally has become synonymous with modifying the eigenvalues of single matrices and is widely used to 'fix' nonpositive definite estimates, not only in animal breeding but also in human genetics[55].

Estimation of genetic parameters using pedigree data relies on equating covariances between relatives to their expectations. Classic estimators, derived from analysis of variance, were restricted to considering few types of relatives simultaneously. In a pioneering study, Hill and Nicholas[56] showed that a maximum likelihood framework of estimation could overcome such limitations. In the 1980s, this was superseded by restricted maximum likelihood (REML)[57] estimation, fitting the so-called animal model. Hill championed the developments and uptake of this methodology, with many of his protégés performing REML analyses of data from livestock[58,59] and selection experiments[60,61], later applying it to evolutionary ecology[62] and human genetics[63,64].

An important prerequisite for quantitative genetic applications is adequate modeling of covariance structures. Hill was particularly interested in heterogeneity of variation, for example between herds of dairy cattle[58,65,66] and its effects on the response to selection. He showed that, if ignored, this could lead to selection favoring individuals in more variable groups and described pragmatic methods to correct for it[67]. Later, Hill focused on heterogeneity of environmental variation, proposing an extension of the standard additive model that allowed for partial genetic control of this component and considering the scope for selecting on it[68,69]. Genomic loci that are associated with trait variance have now been mapped in many species, including humans[70,71].

Longitudinal traits are those that are measured repeatedly along a trajectory (e.g., body weight at various ages or daily milk yield) and expected to change gradually and continuously with time (or its equivalent), both in means and in covariance structure. The latter can be modeled parsimoniously through so-called covariance functions, allowing for potentially infinitely many dimensions. This concept was first developed in evolutionary biology. Hill recognized its importance and encouraged applications to animal breeding[72,73]. It turned out to be equivalent to fitting individual regression equations for random effects. As 'random regression models', these have become a standard part of the quantitative genetics arsenal and form the basis of many routine genetic evaluation schemes for dairy cattle and research applications in evolution[74] and human genetics[75].

Concluding remarks

Mendel's discoveries heralded the discipline of genetics and spawned many subdisciplines therein. Evolutionary, population and quantitative genetics, and their applications in plant and animal breeding and human genetics, are all anchored in the fundamentals of Mendel's laws. In this Perspective we have highlighted the contributions of one scientist, W.G. Hill, across all of these disciplines and demonstrated his remarkable contributions to a better understanding of genetics, selection and evolution of complex traits.

Hill's direct scientific legacy has been highlighted here. His indirect and behind the scenes presence in shaping the entire field of quantitative genetics has also been remarkable, not just through

research by his students, postdocs and collaborators, but also by influencing major graduate textbooks in the field, as evidenced by the acknowledgements therein for example[27,76].

We end with quote from Hill about the discipline of quantitative genetics. It was put on paper in 2010 (ref. [48]), but we were privileged to hear it verbally many times before that. We cite it because it is still true today and, we suspect, for many years to come: "For many decades claims have been made that quantitative genetics was dead or dying but, condescendingly, perhaps still useful until the contents of the black box were revealed, a feat which would be 'just round the corner' … In view of its complexity, it therefore seems likely that the black box will remain cloudy for a while, even though fed information on, inter alia, myriads of genetic markers, levels of gene expressions and trait phenotypes."

Data availability

The code used to create Fig. 3 is available at https://github.com/loic-yengo/Hill-and-Weir-2011—revisited.

Received: 21 March 2022; Accepted: 18 May 2022;
Published online: 11 July 2022

References

1. Fisher, R. A. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* **53**, 399–433 (1918).
2. Provine, W. B. *The Origins of Theoretical Population Genetics* (University of Chicago Press, 1971).
3. Blixt, S. Why didn't Gregor Mendel find linkage? *Nature* **256**, 206 (1975).
4. Fairbanks, D. J. & Rytting, B. Mendelian controversies: a botanical and historical review. *Am. J. Bot.* **88**, 737–752 (2001).
5. Robbins, R. B. Some applications of mathematics to breeding problems III. *Genetics* **3**, 375–389 (1918).
6. Lewontin, R. C. & Kojima, K. The evolutionary dynamics of complex polymorphisms. *Evolution* **14**, 458–472 (1960).
7. Hill, W. G. & Robertson, A. Linkage disequilibrium in finite populations. *Theor. Appl Genet* **38**, 226–231 (1968).
8. Sved, J. A. & Hill, W. G. One hundred years of linkage disequilibrium. *Genetics* **209**, 629–636 (2018).
9. Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet Res* **8**, 269–294 (1966).
10. Lande, R. & Thompson, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**, 743–756 (1990).
11. Bulik-Sullivan, B. K. et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
12. Weir, B. S. & Hill, W. G. Effect of mating structure on variation in linkage disequilibrium. *Genetics* **95**, 477–488 (1980).
13. Hill, W. G. & Weir, B. S. Variances and covariances of squared linkage disequilibria in finite populations. *Theor. Popul Biol.* **33**, 54–78 (1988).
14. Hill, W. G. Estimation of linkage disequilibrium in randomly mating populations. *Heredity* **33**, 229–239 (1974).
15. Excoffier, L. & Slatkin, M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.* **12**, 921–927 (1995).
16. Morton, N. E. et al. The optimal measure of allelic association. *Proc. Natl Acad. Sci. USA* **98**, 5217–5221 (2001).
17. Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).
18. Charlesworth, B., Betancourt, A. J., Kaiser, V. B. & Gordo, I. Genetic recombination and molecular evolution. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 177–186 (2009).
19. Hill, W. G. Variation in genetic identity within kinships. *Heredity* **71**, 652–653 (1993).
20. Avery, P. J. & Hill, W. G. Variability in genetic parameters among small populations. *Genet. Res.* **29**, 193–213 (1977).
21. Weir, B. S., Avery, P. J. & Hill, W. G. Effect of mating structure on variation in inbreeding. *Theor. Popul. Biol.* **18**, 396–429 (1980).
22. Hill, W. G. & Weir, B. S. Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res (Camb.)* **93**, 47–64 (2011).
23. Hill, W. G. & White, I. M. S. Identification of pedigree relationship from genome sharing. *G3-Genes Genomes Genet.* **3**, 1553–1571 (2013).
24. Hill, W. G. & Weir, B. S. Variation in actual relationship among descendants of inbred individuals. *Genet. Res.* **94**, 267–274 (2012).
25. Falconer, D. S. *Introduction to Quantitative Genetics* (Oliver and Boyd, 1960).
26. Hivert, V., Wray, N. R. & Visscher, P. M. Gene action, genetic variation, and GWAS: a user-friendly web tool. *PLoS Genet.* **17**, e1009548 (2021).
27. Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer, 1998).
28. Charlesworth, D. & Willis, J. H. The genetics of inbreeding depression. *Nat. Rev. Genet.* **10**, 783–796 (2009).
29. Robertson, A. & Hill, W. G. Population and quantitative genetics of many linked loci in finite populations. *Proc. R. Soc. Ser. B Biol. Sci.* **219**, 253–264 (1983).
30. Mukai, T., Cardellino, R. A., Watanabe, T. K. & Crow, J. F. The genetic variance for viability and its components in a local population of *Drosophila melanogaster*. *Genetics* **78**, 1195–1208 (1974).
31. Joshi, P. K. et al. Directional dominance on stature and cognition in diverse human populations. *Nature* **523**, 459–462 (2015).
32. Yengo, L. et al. Detection and quantification of inbreeding depression for complex traits from SNP data. *Proc. Natl Acad. Sci. USA* **114**, 8602–8607 (2017).
33. Clark, D. W. et al. Associations of autozygosity with a broad range of human phenotypes. *Nat. Commun.* **10**, 4957 (2019).
34. Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* **4**, e1000008 (2008).
35. Maki-Tanila, A. & Hill, W. G. Influence of gene interaction on complex trait variation with multilocus models. *Genetics* **198**, 355–367 (2014).
36. Keightley, P. D. Models of quantitative variation of flux in metabolic pathways. *Genetics* **121**, 869–876 (1989).
37. Charlesworth, B. Causes of natural variation in fitness: evidence from studies of *Drosophila* populations. *Proc. Natl Acad. Sci. USA* **112**, 1662–1669 (2015).
38. Polderman, T. J. et al. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
39. Visscher, P. M. et al. 10 years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet* **101**, 5–22 (2017).
40. Zhu, Z. et al. Dominance genetic variation contributes little to the missing heritability for human complex traits. *Am. J. Hum. Genet* **96**, 377–385 (2015).
41. Hivert, V. et al. Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *Am. J. Hum. Genet* **108**, 786–798 (2021).
42. Pazokitoroudi, A., Chiu, A. M., Burch, K. S., Pasaniuc, B. & Sankararaman, S. Quantifying the contribution of dominance deviation effects to complex trait variation in biobank-scale data. *Am. J. Hum. Genet* **108**, 799–808 (2021).
43. Bloom, J. S. et al. Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. *Nat. Commun.* **6**, 8712 (2015).
44. Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T. L. & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. *Nature* **494**, 234–237 (2013).
45. Zhang, X. S. & Hill, W. G. Genetic variability under mutation selection balance. *Trends Ecol. Evol.* **20**, 468–470 (2005).
46. Hill, W. G. Predictions of response to artificial selection from new mutations. *Genet. Res.* **40**, 255–278 (1982).
47. Hill, W. G. Rates of change in quantitative traits from fixation of new mutations. *Proc. Natl Acad. Sci. USA* **79**, 142–145 (1982).
48. Hill, W. G. Understanding and using quantitative genetic variation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 73–85 (2010).
49. Hill, W. G. & Kirkpatrick, M. What animal breeding has taught us about evolution. *Annu. Rev. Ecol. Evol. Syst.* **41**, 1–19 (2010).
50. Wray, N. R., Kemper, K. E., Hayes, B. J., Goddard, M. E. & Visscher, P. M. Complex trait prediction from genome data: contrasting EBV in livestock to PRS in humans: genomic prediction. *Genetics* **211**, 1131–1141 (2019).
51. Hill, W. G. Prediction and evaluation of response to selection with overlapping generations. *Anim. Prod.* **18**, 117–139 (1974).
52. Sales, J. & Hill, W. G. Effect of sampling errors on efficiency of selection indexes. 1. Use of information from relatives for single trait improvement. *Anim. Prod.* **22**, 1–17 (1976).
53. Sales, J. & Hill, W. G. Effect of sampling errors on efficiency of selection indexes. 2. Use of information on associated traits for improvement of a single important trait. *Anim. Prod.* **23**, 1–14 (1976).
54. Hayes, J. F. & Hill, W. G. Modification of estimates of parameters in the construction of genetic selection indexes (bending). *Biometrics* **37**, 483–493 (1981).
55. de Vlaming, R. et al. Multivariate analysis reveals shared genetic architecture of brain morphology and human behavior. *Commun. Biol.* **4**, 1180 (2021).
56. Hill, W. G. & Nicholas, F. W. Estimation of heritability by both regression of offspring on parent and intra-class correlation of sibs in one experiment. *Biometrics* **30**, 447–468 (1974).
57. Patterson, H. D. & Thompson, R. Recovery of interblock information when block sizes are unequal. *Biometrika* **58**, 545–554 (1971).
58. Visscher, P. M., Thompson, R. & Hill, W. G. Estimation of genetic and environmental variances for fat yield in individual herds and an investigation into heterogeneity of variance between herds. *Livest. Prod. Sci.* **28**, 273–290 (1991).

59. Meyer, K. Maximum likelihood procedures for estimating genetic parameters for later lactations of dairy cattle. *J. Dairy Sci.* **66**, 1988–1997 (1983).
60. Meyer, K. & Hill, W. G. Mixed model analysis of a selection experiment for food intake in mice. *Genetical Res.* **57**, 71–81 (1991).
61. Keightley, P. D. & Hill, W. G. Quantitative genetic-variation in body size of mice from new mutations. *Genetics* **131**, 693–700 (1992).
62. Kruuk, L. E. Estimating genetic parameters in natural populations using the 'animal model'. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **359**, 873–890 (2004).
63. Visscher, P. M., Haley, C. S., Heath, S. C., Muir, W. J. & Blackwood, D. H. Detecting QTLs for uni- and bipolar disorder using a variance component method. *Psychiatr. Genet* **9**, 75–84 (1999).
64. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
65. Brotherstone, S. & Hill, W. G. Heterogeneity of variance amongst herds for milk production. *Anim. Prod.* **42**, 297–303 (1986).
66. Hill, W. G., Edwards, M. R., Ahmed, M. K. A. & Thompson, R. Heritability of milk yield and composition at different levels and variability of production. *Anim. Prod.* **36**, 59–68 (1983).
67. Hill, W. G. On Selection among groups with heterogeneous variance. *Anim. Prod.* **39**, 473–477 (1984).
68. Hill, W. G. & Mulder, H. A. Genetic analysis of environmental variation. *Genet. Res.* **92**, 381–395 (2010).
69. Mulder, H. A., Bijma, P. & Hill, W. G. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. *Genetics* **175**, 1895–1910 (2007).
70. Young, A. I., Wauthier, F. L. & Donnelly, P. Identifying loci affecting trait variability and detecting interactions in genome-wide association studies. *Nat. Genet.* **50**, 1608–1614 (2018).
71. Yang, J. et al. FTO genotype is associated with phenotypic variability of body mass index. *Nature* **490**, 267–272 (2012).
72. Kirkpatrick, M., Hill, W. G. & Thompson, R. Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genetical Res.* **64**, 57–69 (1994).
73. Meyer, K. & Hill, W. G. Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. *Livest. Prod. Sci.* **47**, 185–200 (1997).
74. Wilson, A. J., Kruuk, L. E. & Coltman, D. W. Ontogenetic patterns in heritable variation for body size: using random regression models in a wild ungulate population. *Am. Nat.* **166**, E177–E192 (2005).
75. Ni, G. et al. Genotype-covariate correlation and interaction disentangled by a whole-genome multivariate reaction norm model. *Nat. Commun.* **10**, 2239 (2019).
76. Walsh, B. & Lynch, M. *Evolution and Selection of Quantitative Traits* (Oxford University Press, 2018).

Acknowledgements

P.M.V. is supported by grants from the National Health and Medical Research Council (1113400) and the Australian Research Council (FL180100072). N.R.W. is supported by the National Health and Medical Research Council (1113400 and 1173790). K.M. was supported by Meat and Livestock Australia (grant L.GEN.2204). We thank J. Sidorenko and L. Yengo for help with the figures. We acknowledge the many friends and colleagues of Bill Hill who would have loved to contribute to this paper and who would have been qualified to do so.

Author contributions

All authors contributed to the writing of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to Peter M. Visscher.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature America, Inc. 2022