Nature of QG traits 2

Authors: Author 1, Author 2, ..., H. Jorjani
Last update: 2019-12-20

Inside the black box

First, the predictions of quantitative genetics depend upon the structure of genome. In particular, we need to know the number of loci, their linkage relations, their allelic frequency distribution and the mutational and recombinational sources of new variation. Second, we need to know the relations between gene and organism, how gene action, in particular environments, is translated into phenotype. The knowledge about these questions can come to us only by opening up the black box whose outer shape we have so far been describing, and seeing what the machinery inside really looks like. This is the task of molecular and developmental genetics and some general knowledge is already available to us from the recent activity in these fields. Our models of quantitative genetics must either take cognizance of these findings or else show that they are, in fact, irrelevant because of the robustness of our theory.

Richard Charles Lewontin (1976)

The purpose of this chapter is to show

- The existence of genetic variation at all biological levels,
- The single locus model of genetics is unrealistic and inadequate,
- A more realistic model is based on involvement of hundreds to thousands of loci in phenotypic expression of any trait.

2.1 Reach of Single Locus Model

Single locus models, in any branch of genetics, have been the starting point of developing new theories, and they are excellent pedagogical first steps in teaching (as we will see in Part V). In every species, thousands (if not tens of thousands) of traits are attributed to a single nucleotide mutation, each in a single locus. At the first glance this seems to be the case. However, a deeper investigation of these traits reveal a complicated picture. Take for example, some easily accessible traits in humans, such as the pigmented iris (being blue or not), freckles, and diabetes. These examples of human traits are used here because they are so easy to observe.

In humans, eye color depends on the amount of the pigment Melanin in two layers of cells in the eye (one of which is only two rows of cells). There is variation in the number of such cells among individuals, and also during the life time of an individuals. The number and size of the Melanin particles also shows variation, and they have different "life

cycles". The color of the eye depends on the interaction between the shape of the eye ball and the existing Melanin pigments, and how the light is reflected from the eye because of this interaction. In fact, in humans, the blue color of eye is not related to any blue pigment at all, but due to the Tyndall effects (similar to the Rayleigh scattering which gives the sky its blue color). In other species a variety of mechanisms contribute to the eye color. For example, in birds there are oil droplets that affect the eye color irrespective of the background pigmentation.

In the case of freckles in humans, there is variation among individuals in the number of freckles per square centimeter. Also, there is variation with regard to the color contrast. Further, there is variation in the spread of freckles, limited to around the nose, or spread wider to other parts of the body (especially upper body).

Another frequent example is diabetes, for which many questions raise the doubt about single locus explanations. For example, is the onset of the disease at the same age for all patients? Do all the patients have the same level of blood sugar at the each incidence of disease? Does everyone need the same amount of insulin?

There are many details related to each of the above examples. One can easily draw the conclusion that all aspects of a *"simple"* trait such as eye color, freckles, or diabetes, cannot be explained just by the single mutation / single locus model. Single locus models are good pedagogical starting points, but their reach to explain the nature of the traits are short of being satisfactory.

Form your own opinion

No references have been provided in this section! Convince yourself that the arguments put forward in this section are valid. Search in the literature to find references providing enough evidence that show single mutation / single locus models are just the "*first answers*".

Do the searching for several traits until you can form your own general opinion about the so-called "*single mutation / single locus models"*.

2.2 All traits are polygenic

The entire field of quantitative genetics, and much of the field of population genetics, deal with continuous variation at the phenotypic level. The most important genetic model for studying continuous variation, the infinitesimal model (Fisher (1918), Wright (1921)[1]), maintains that the loci affecting continuous variation follow the Mendelian rules of transmission genetics. However, among other things, there are two main differences between the simple single locus models and the infinitesimal model [2] :

- a large number of loci affect the character/trait under consideration; and
- the effect of each locus on the phenotypic variance is small.

Combining these two properties of the infinitesimal model enables us to invoke central limit theorem, and consequently the normal distribution for genotypic and environmental values and their sum, i.e. phenotypic values. As an illustration Figure 2.1 assumes there is a trait affected by a large number of loci (*i.e.* 12 loci), each with a small effect on the genotypic value. The sum of the effect of these 12 simulated loci shows an approximate normal distribution of the genetic effects.

There are several lines of evidence that suggest the infinitesimal model is actually not far from being true.

1: **Wright (1921):** In this book Wright (1921) refers to five publications in the journal *Genetics* by Wright (1921b; 1921c; 1921d; 1921e; 1921f)

2: **The Infinitesimal Model:** For a full description of the infinitesimal model see Chapter 5

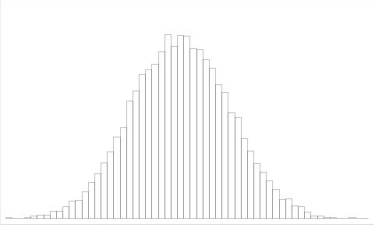

Figure 2.1: Distribution of sum of 12 randomly generated numbers from a uniform distribution in the range 0 to 1

2.3 Evidence from biochemistry

Each cell contains many types of molecules. Here, we concentrate on proteins, because some of them, namely enzymes, have unequivocally important roles in life. Techniques for detection of proteins have evolved a lot since 1950's, and consequently the number of detected proteins has also increased. Simple gel electrophoresis techniques could detect tens of different proteins in any biological sample (see *e.g.* Hubby and Lewontin (1966), and Lewontin and Hubby (1966)). With the advent of two-dimensional gel electrophoresis (O'Farrell, 1975) number of detectable proteins increased to thousands of proteins (Magdeldin *et al.*, 2014).

A little bit of history

The two papers by Hubby/Lewontin had a profound effect on genetics. The main *take home message* was the claim that most of the variation exists within populations. Two immediate effects of these two papers were:

A) The concept of *race* in humans, and *breed* in animals/ plants lost their genetic justification; and

B) The existence of huge amounts of variation within each population was one of the major igniting stuff of the Neutral Mutation Theory of Molecular Evolution by Kimura (1968).

Using mass spectrometry, and depending on how proteins are defined, between several hundreds of thousands to close a million proteins can be identified in humans (Wilhelm *et al.*, 2014). Incidentally, the number of proteins for human and *Arabidopsis thaliana* are similar (see *Proteomics Database*[*]).

Development of different *omic* techniques has made it possible to routinely detect several thousands of protein molecules in each cell line (see *e.g.* Zhu *et al.* (2008), and Chang *et al.* (2011)). Geiger *et al.* (2012) could show that the number of proteins in any cell line of human is, on average, more than 10,000. The majority of the cell proteins are *housekeeping* proteins, *i.e.* they have the role of maintaining basic cell functions and are found in all cells. The Human Protein Atlas[†], based the work of Uhlén *et al.* (2015), puts the number of housekeeping proteins close to 10,000. This is in agreement with the estimates by Geiger *et al.* (2012). The housekeeping proteins are coded by housekeeping loci (*e.g.* Eisenberg and Levanon (2013) and Zhang *et al.* (2015)) whose number, because of different mechanisms, is not as many as proteins. Nonetheless, the number of loci (housekeeping or otherwise) expressed in any cell line is in the order of several thousands.

From a biochemical perspective virtually all enzymes are bound together in (linear or branched) chains. It is easy to see the interaction of the enzymes just by looking at any biochemical chart. This means that all enzymes are virtually connected to each other. According to the theory of *"Metabolic Control Analysis"* (see e.g. Kacser and Burns (1979), Kacser and Burns (1981), Hofmeyr and Cornish-Bowden (1991), and Bagheri and Wagner (2004)) flux of any enzymatic chain is affected by all enzymes in that chain. However, effect of each enzyme follows a hyperbolic function. Following a suggestion by Hartl *et al.* (1985), it has also been shown that the control coefficient exerted by any enzyme is subject to change under selection (Dean *et al.* (1986); Dykhuizen *et al.* (1987)). Consequently, effects of some enzymes may be so small in certain generations that their effect might not be measurable in small sample sizes. Kacser (1989), without any reference to modern techniques of identifying proteins/enzymes, postulated that any quantitative trait may be under control of a very large number of enzymes, *i.e. "say 5000"*.

[*] Proteomics Database: `https://www.proteomicsdb.org/`. Accessed 2019-11-28.
[†] The Human Protein Atlas: `https://www.proteinatlas.org/humanproteome/tissue/housekeeping`. Accessed 2019-11-25.

2.4 Evidence from Genomics

Results of the various genome projects put the total number of functional chromosome segments for the most important animal and plant species to an average of about 25000 (between 15000 and 50000). Viruses and bacteria have usually less than 1000 such segments, arthropods less than 15000, mammals more than 20000, and plants more than 25000 (e.g. Rogers (2017)).

Terminology: locus (pl. loci)

The term *"chromosome segment"* (whether functional or not) has had many equivalents since the days of re-discovery of Mendel's rules. Each of these has been connected to the prevailing understanding of their role at a specific time period. The most famous equivalent happens to be the most controversial (see *e.g.* Bromham (2016) and Portin and Wilkins (2017)).

In this book, the term locus, and its plural loci, are used.

The question of interest is the following: given the information that there are 15000 to 50000 functional segments, and given the information that up to 10000 of them are expressed in each cell line, how many of them have a *measurable* effect on *a* trait?

This question has received special attention in human genetics, where individuals with certain phenotype (usually a disease state) are compared to other individuals free from that phenotype. The techniques used in the published studies have been the most recent technique disposable to the investigators at the time of study. For example, in a study of seven brain tumors Watson *et al.* (2001), and Shannon *et al.* (2002) used expression arrays, and found differential expression for 196 loci out of 1013 loci studied.

More than a decade later, single nucleotide polymorphism (SNP) arrays were used by Wood *et al.* (2014) in a meta-analysis of studies in stature in humans (see also Marouli *et al.* (2017)). They found close to 10000 SNPs that collectively could explain about 1/3 of stature's heritability value[3] . The number of detected loci (in most cases, SNPs) affecting a trait (any trait) increased almost exponentially in the first 10 years of genome wide association studies (GWAS) from a few hundreds in 2008 to a few thousands 10 years later (Visscher *et al.* (2017)[4]). For intelligence, a trait difficult to define, Savage *et al.* (2018) found more than 12000 variants, clustered in 205 loci, that collectively explain up to 5.2% of the variance of intelligence.

Terminology: Variant

The term *"variant"*, often combined with other words to form various terms, such as *"sequence variant"*, *"common variant"*, *"rare variant"*, refers to any polymorphism, *i.e.* a mutation compared to the *"reference genome"*, at the nucleotide level. This term can be used irrespective of the effect that the polymorphism might have. In contrast, the term

3: **Heritability:** The concept of *heritability* will be discussed in Chapter 9.

4: **Estimated number of loci involved in a quantitative trait:** Figure 2 in Visscher *et al.* (2017) is a *must see* figure.

"causal variant" is a polymorphism that is assumed to have a certain effect on the phenotype of interest.

In human genetics, high levels of genome-wide significance (say, $p < 10^{-8}$) are important because the detected loci may be used in the design of diagnostic tools or therapeutics substances. Therefore, the number of loci with measurable effect for human traits is probably underestimated. In contrast, in animal breeding the significance level for any SNP is not important, and usually not reported, because the sum of the effects of all SNPs is of interest. There are, however, many exceptions to this rule (*e.g.* Bolormaa *et al.* (2010) and Bolormaa *et al.* (2014) for beef cattle). In a study in dairy cattle, Jiang *et al.* (2019), using data on almost 300000 Holstein cows, observed genome-wide significant effects ($p < 10^{-7}$) for between 15 loci (heifer conception rate) and 15215 loci (milk protein percentage). According to *Animal QTL Database*‡ in cattle alone more than 130000 QTLs, related to more than 634 traits have been detected.

2.5 Evidence from trait-based versus locus-based mutation rates

Close examination of natural and laboratory populations reveals that per trait rate of mutation is between 10^3 and 10^4 times larger than per locus rate of mutation (see *e.g.* Houle *et al.* (1992), Keightley and Hill (1992), Santiago *et al.* (1992), and Mackay *et al.* (1994)). By comparison of these two rates one can easily deduce that the number loci affecting a trait may be as high as 1000-10000.

2.6 Evidence from long-term selection experiments

As we will see in Chapter 27 selection on a single locus is very effective in changing the frequency of the favorable allele. We will also see that as the number of loci affecting a trait goes up, the change in the frequency of the favorable alleles becomes smaller and smaller (Crow and Kimura, 2009, Ch 5). Consequently, continued change of mean in an experimental population indicates the involvement of many loci. Of course, *de novo* mutations contribute to the maintenance of genetic variance in spite of selection (see Chapter 26).

There are many old and new long-term selection experiments that have shown significant response to selection after many generations: *e.g.* in Drosophila (Yoo, 1980), mice (Bünger *et al.*, 1998; Holt *et al.*, 2005), maize (Dudley, 2007), chicken (Dunnington *et al.*, 2013), and many more.

To estimate the number of loci affecting a trait, different quantitative or molecular genetic methods can be employed. In the longest running selection experiment (the Illinois Long-Term Selection Experiment in Maize), the cumulative response to selection for oil and protein contents

‡ Animal QTL Database: https://www.animalgenome.org/cgi-bin/QTLdb/index. Accessed 2019-12-20.

have been used to arrive at an estimate of up to 200 loci involved in these traits (Dudley and Lambert, 2010). Different molecular methods have resulted in estimates ranging from tens of loci to a few hundreds SNP cites (Moose *et al.*, 2004, Laurie *et al.*, 2004, see also the commentary by Hill, 2005). In mice, Bünger *et al.* (1998) estimated about 92 loci to be involved for the trait body weight at six weeks of age. In case of chickens Johansson *et al.* (2010) estimated hundreds of loci to be involved in the body weight.

In this chapter, it was claimed that *"single locus models"* are inadequate. Further, abundant evidence from several lines of arguments were presented to support the claim that many loci affect each and every quantitative trait. The number of loci mentioned in this chapter are all conservative estimates. Given larger data sets and/or better analytical methods, would enable the detection of further loci that hitherto have not been detected.

Bibliography

Citations in alphabetical order.

Alberts, B., A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Walter (2015). *Molecular biology of the cell*. Sixth edition. New York, NY: Garland Science, Taylor and Francis Group. 1465 pp. (cited on page 5).

Bagheri, H. C. and G. P. Wagner (2004). Evolution of Dominance in Metabolic Pathways. *Genetics* **168**: 1713–1735. DOI: 10.1534/genetics.104.028696 (cited on page 11).

Bateson, W. (1901). Problems of heredity as a subject for horticultural investigations. *Journal of the Royal Horticultural Society* **25**: 54–61. DOI: 10.1017/CBO9780511693946 (cited on page 4).

Bolormaa, S., J. E. Pryce, B. J. Hayes, and M. E. Goddard (2010). Multivariate analysis of a genome-wide association study in dairy cattle. *Journal of Dairy Science* **93**: 3818–3833. DOI: 10.3168/jds.2009-2980 (cited on page 13).

Bolormaa, S., J. E. Pryce, A. Reverter, Y. Zhang, W. Barendse, K. Kemper, B. Tier, K. Savin, B. J. Hayes, and M. E. Goddard (2014). A Multi-Trait, Meta-analysis for Detecting Pleiotropic Polymorphisms for Stature, Fatness and Reproduction in Beef Cattle. *PLoS Genetics* **10**: ed. by J. Flint, e1004198. DOI: 10.1371/journal.pgen.1004198 (cited on page 13).

Bromham, L. (2016). What is a gene for? *Biology & Philosophy* **31**: 103–123. DOI: 10.1007/s10539-014-9472-9 (cited on page 12).

Bulmer, M. G. (2003). *Francis Galton: pioneer of heredity and biometry*. Baltimore: Johns Hopkins University Press. 357 pp. (cited on page 17).

Bünger, L., U. Renne, G. Dietl, and S. Kuhla (1998). Long-term selection for protein amount over 70 generations in mice. *Genetical Research* **72**: 93–109. DOI: 10.1017/S0016672398003401 (cited on pages 13, 14).

Casella, G. and R. L. Berger (2002). *Statistical inference*. 2nd Ed. Duxbury. 686 pp. (cited on page 6).

Castle, W. E. (1905). The Mutation Theory of Organic Evolution, from the Standpoint of Animal Breeding. *Science* **21**: 521–525. DOI: 10.1126/science.21.536.521 (cited on page 19).

Chang, C.-W., W.-C. Cheng, C.-R. Chen, W.-. Shu, M.-L. Tsai, C.-L. Huang, and I. C. Hsu (2011). Identification of Human Housekeeping Genes and Tissue-Selective Genes by Microarray Meta-Analysis. *PLOS ONE* **6**: e22859. DOI: 10.1371/journal.pone.0022859 (cited on page 11).

Cohen, I. B. (1985). *Revolution in science*. 1st Ed. Harvard University Press. 746 pp. (cited on page viii).

Collas, P., T. M. Liyakat Ali, A. Brunet, and T. Germier (2019). Finding Friends in the Crowd: Three-Dimensional Cliques of Topological Genomic Domains. *Frontiers in Genetics* **10**: 602. DOI: 10.3389/fgene.2019.00602 (cited on page 5).

Crow, J. F. and M. Kimura (2009). *An introduction to population genetics theory*. OCLC: 1027901624. Jodhpur; New Jersey: Scientific Publisher (India) ; The Blackburn Press (cited on pages 7, 13, 16).

Darwin, C. R. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London. 502 pp. (cited on pages 15, 16, 18, 19).

Dean, A. M., D. E. Dykhuizen, and D. L. Hartl (1986). Fitness as a function of ?-galactosidase activity in *Escherichia coli*. *Genetical Research* **48**: 1–8. DOI: 10.1017/S0016672300024587 (cited on page 11).

Dudley, J. W. (2007). From Means to QTL: The Illinois Long-Term Selection Experiment as a Case Study in Quantitative Genetics. *Crop Science* **47**: (Supplement_3), S–20. DOI: 10.2135/cropsci2007.04.0003IPBS (cited on page 13).

Dudley, J. W. and R. J. Lambert (2010). '100 Generations of Selection for Oil and Protein in Corn'. *Plant Breeding Reviews*. Oxford, UK: John Wiley & Sons, Inc., pp. 79–110. DOI: 10.1002/9780470650240.ch5 (cited on page 14).

Dunnington, E. A., C. F. Honaker, M. L. McGilliard, and P. B. Siegel (2013). Phenotypic responses of chickens to long-term, bidirectional selection for juvenile body weight: A Historical perspective. *Poultry Science* **92**: 1724–1734. DOI: 10.3382/ps.2013-03069 (cited on page 13).

Dykhuizen, D. E., A. M. Dean, and D. L. Hartl (1987). Metabolic Flux and Fitness. *Genetics* **115**: 25–31 (cited on page 11).

East, E. M. (1910). A Mendelian Interpretation of Variation that is Apparently Continuous. *The American Naturalist* **44**: 65–82 (cited on page 19).
Eisenberg, E. and E. Y. Levanon (2013). Human housekeeping genes, revisited. *Trends in Genetics* **29**: 569–574. DOI: 10.1016/j.tig.2013.05.010 (cited on page 11).
Eldredge, N. and S. J. Gould (1972). 'Punctuated equilibria: an alternative to phyletic gradualism'. *Models in paleobiology*. Cooper and Co., San Francisco, pp. 82–115 (cited on page 19).
Ewens, W. J. (2004). *Mathematical population genetics. 1: Theoretical introduction*. 2. ed. Interdisciplinary applied mathematics Mathematical biology. New York, NY: Springer. 417 pp. (cited on page 7).
Falconer, D. S. (1960). *Introduction to Quantitative Genetics*. 1st edition (cited on page v).
Falconer, D. S. and T. F. C. Mackay (1996). *Introduction to Quantitative Genetics*. 4th ed. Longman Group Ltd. 480 pp. (cited on pages v, vi, viii).
Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **525**: 399–433 (cited on pages 4, 10, 20, 23, 26).
Galton, F. (1865). Hereditary talent and character. *Macmillan's Magazine* **12**: 157–166, 318–327 (cited on page 18).
Galton, F. (1877). Typical laws of heredity, 492–533 (cited on page 18).
Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland* **15**: 246–263 (cited on pages 4, 19).
Galton, F. (1889). *Natural inheritance*. Macmillan, London (cited on page 19).
Geiger, T., A. Wehner, C. Schaab, J. Cox, and M. Mann (2012). Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins. *Molecular & Cellular Proteomics : MCP* **11**: DOI: 10.1074/mcp.M111.014050 (cited on page 11).
Gould, S. J. (2002). *The structure of evolutionary theory*. Cambridge, Mass: Belknap Press of Harvard University Press. 1433 pp. (cited on page 19).
Griffiths, A. J., S. R. Wessler, J. Carroll, and J. Doebley (2015). *Introduction to genetic analysis*. 11th Ed. W. H. Freeman & Company (cited on pages 5, 15).
Haldane, J. B. S. (1924). A mathematical theory of natural and artificial selection. Part I. (cited on pages 23, 24).
Haldane, J. B. S. (1932). *The causes of evolution*. Longmans, Green and Co. 235 pp. (cited on page 23).
Hamilton, M. B. (2009). *Population genetics*. OCLC: ocn259716125. Chichester, UK ; Hoboken, NJ: Wiley-Blackwell. 407 pp. (cited on page 7).
Hartl, D. L. and A. G. Clark (2007). *Principles of population genetics*. 4th Ed. Sinauer Associates (cited on page 7).
Hartl, D. L., D. E. Dykhuizen, and A. M. Dean (1985). Limits of Adaptation: The Evolution of Selective Neutrality. *Genetics* **111**: 655–674 (cited on page 11).
Hill, W. G. (2005). A Century of Corn Selection. *Science* **307**: 683–685 (cited on page 14).
Hill, W. G. (1984a). *Quantitative Genetics: Explanation and analysis of continuous variation*. Van Nostrand Reinhold. 376 pp. (cited on page 21).
Hill, W. G. (1984b). *Quantitative genetics: Selection*. Van Nostrand Reinhold. 426 pp. (cited on page 21).
Hofmeyr, J.-H. S. and A. Cornish-Bowden (1991). Quantitative assessment of regulation in metabolic systems. *European Journal of Biochemistry* **200**: 223–236. DOI: 10.1111/j.1432-1033.1991.tb21071.x (cited on page 11).
Hogg, R. V., J. W. McKean, and A. T. Craig (2019). *Introduction to mathematical statistics*. Eighth edition. Boston: Pearson. 746 pp. (cited on page 6).
Holt, M., T. Meuwissen, and O. Vangen (2005). Long-term responses, changes in genetic variances and inbreeding depression from 122 generations of selection on increased litter size in mice. *Journal of Animal Breeding and Genetics* **122**: 199–209. DOI: 10.1111/j.1439-0388.2005.00526.x (cited on page 13).
Houle, D., D. K. Hoffmaster, S. Assimacopoulos, and B. Charlesworth (1992). The genomic mutation rate for fitness in Drosophila. *Nature* **359**: 58–60. DOI: 10.1038/359058a0 (cited on page 13).
Hubby, J. L. and R. C. Lewontin (1966). A Molecular Approach to the Study of Genic Heterozygosity in Natural Populations. I. the Number of Alleles at Different Loci in Drosophila Pseudoobscura. *Genetics* **54**: 577–594 (cited on page 11).
Hull, D. L. (1988). *Science as a process: an evolutionary account of the social and conceptual development of science*. Paperback ed., [Nachdr.] Science and its conceptual foundations. OCLC: 837696311. Chicago: University of Chicago Press. 586 pp. (cited on page 15).

Jensen, J. D., B. A. Payseur, W. Stephan, C. F. Aquadro, M. Lynch, D. Charlesworth, and B. Charlesworth (2019). The importance of the Neutral Theory in 1968 and 50 years on: A response to Kern and Hahn 2018: COMMENTARY. *Evolution* **73**: 111–114. DOI: 10.1111/evo.13650 (cited on page 18).

Jiang, J., L. Ma, D. Prakapenka, P. M. VanRaden, J. B. Cole, and Y. Da (2019). A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle. *Frontiers in Genetics* **10**: 412. DOI: 10.3389/fgene.2019.00412 (cited on page 13).

Johansson, A. M., M. E. Pettersson, P. B. Siegel, and Ö. Carlborg (2010). Genome-Wide Effects of Long-Term Divergent Selection. *PLoS Genetics* **6**: ed. by B. Walsh, e1001188. DOI: 10.1371/journal.pgen.1001188 (cited on page 14).

Kacser, H. (1989). 'Quantitative variation and the control analysis of enzyme systems.' *Hill W. G., and T.F.C. Mackay (1989) Evolution and Animal Breeding: Reviews on Molecular and Quantitative Approaches in Honour of Alan Robertson*. C.A.B. International, Wallingford, UK, pp. 219–226 (cited on page 11).

Kacser, H. and J. A. Burns (1979). Molecular Democracy: Who Shares the Controls? *Biochemical Society Transactions* **7**: 1149–1160. DOI: 10.1042/bst0071149 (cited on page 11).

Kacser, H. and J. A. Burns (1981). The Molecular Basis of Dominance. *Genetics* **97**: 639–666 (cited on pages 5, 11).

Keel, B. N., D. J. Nonneman, A. K. Lindholm-Perry, W. T. Oliver, and G. A. Rohrer (2019). A Survey of Copy Number Variation in the Porcine Genome Detected From Whole-Genome Sequence. *Frontiers in Genetics* **10**: 737. DOI: 10.3389/fgene.2019.00737 (cited on page 5).

Keightley, P. D. and W. G. Hill (1992). Quantitative Genetic Variation in Body Size of Mice From New Mutations. *Genetics* **131**: 693–700 (cited on page 13).

Kempthorne, O. (1976). 'Status of quantitative genetics'. *Proceedings of the International Conference on Quantitative Genetics*. Iowa State University Press, pp. 719–760 (cited on page 4).

Kern, A. D. and M. W. Hahn (2018). The Neutral Theory in Light of Natural Selection. *Molecular Biology and Evolution* **35**: ed. by S. Kumar, 1366–1371. DOI: 10.1093/molbev/msy092 (cited on page 18).

Kimura, M. (1968). Evolutionary Rate at the Molecular Level. *Nature* **217**: 624–626. DOI: 10.1038/217624a0 (cited on pages 11, 17).

Kuhn, T. S. (1996). *The structure of scientific revolutions*. 3rd Ed. Chicago, IL: University of Chicago Press. 212 pp. (cited on page viii).

Larsen, R. J. and M. L. Marx (2017). *An Introduction to Mathematical Statistics and Its Applications* (cited on page 6).

Laurie, C. C., S. D. Chasalow, J. R. LeDeaux, R. McCarroll, D. Bush, B. Hauge, C. Lai, D. Clark, T. R. Rocheford, and J. W. Dudley (2004). The Genetic Architecture of Response to Long-Term Artificial Selection for Oil Concentration in the Maize Kernel. *Genetics* **168**: 2141–2155. DOI: 10.1534/genetics.104.029686 (cited on page 14).

Lehmann, E. L. and G. Casella (1998). *Theory of point estimation*. 2nd Ed. Springer texts in statistics. New York: Springer. 589 pp. (cited on page 6).

Lehninger, A. L., D. L. Nelson, and M. M. Cox (2013). *Lehninger principles of biochemistry*. 6th ed. OCLC: ocn820352899. New York: W.H. Freeman. 1119 pp. (cited on page 5).

Lewontin, R. C. and J. L. Hubby (1966). A Molecular Approach to the Study of Genic Heterozygosity in Natural Populations. Ii. Amount of Variation and Degree of Heterozygosity in Natural Populations of Drosophila Pseudoobscura. *Genetics* **54**: 595–609 (cited on page 11).

Liu, Y., X. Chang, J. Glessner, H. Qu, L. Tian, D. Li, K. Nguyen, P. M. A. Sleiman, and H. Hakonarson (2019). Association of Rare Recurrent Copy Number Variants With Congenital Heart Defects Based on Next-Generation Sequencing Data From Family Trios. *Frontiers in Genetics* **10**: 819. DOI: 10.3389/fgene.2019.00819 (cited on page 5).

Liu, Y. and X. Li (2014). Has Darwin's Pangenesis Been Rediscovered? *BioScience* **64**: 1037–1041. DOI: 10.1093/biosci/biu151 (cited on page 17).

Lynch, M. and B. Walsh (1998). *Genetics and analysis of quantitative traits*. Sunderland, Mass: Sinauer. 980 pp. (cited on pages vi, 25).

Mackay, T. F. C., J. D. Fry, R. F. Lyman, and S. V. Nuzhdin (1994). Polygenic mutation in Drosophila melanogaster: estimates from response to selection of inbred strains. *Genetics* **136**: 937–951 (cited on page 13).

Magdeldin, S., S. Enany, Y. Yoshida, B. Xu, Y. Zhang, Z. Zureena, I. Lokamani, E. Yaoita, and T. Yamamoto (2014). Basics and recent advances of two dimensional- polyacrylamide gel electrophoresis. *Clinical proteomics* **11**: 16. DOI: 10.1186/1559-0275-11-16 (cited on page 11).

Maindonald, J. and W. J. Braun (2010). *Data Analysis and Graphics Using R - an Example-Based Approach*. 3rd Ed. Cambridge University Press (cited on page 6).

Malthus, T. (1798). *An Essay on the Principle of Population*. J. Johnson. 125 pp. (cited on pages 16, 17).

Marouli, E. *et al.* (2017). Rare and low-frequency coding variants alter human adult height. *Nature* **542**: 186–190. DOI: 10.1038/nature21039 (cited on page 12).

Mayr, E. (1977). Darwin and Natural Selection: How Darwin may have discovered his highly unconventional theory. *American Scientist* **65**: 321–327 (cited on pages 15, 16).

Mayr, E. (1982). *The growth of biological thought: diversity, evolution, and inheritance*. Cambridge, Mass.: Harvard Univ. Pr. 974 pp. (cited on page 21).

Mayr, E. (1988). *Toward a new philosophy of biology: Observations of an evolutionist*. Harvard University Press. 564 pp. (cited on page 15).

Meaburn, K. J. and T. Misteli (2019). Assessment of the Utility of Gene Positioning Biomarkers in the Stratification of Prostate Cancers. *Frontiers in Genetics* **10**: 1029. DOI: 10.3389/fgene.2019.01029 (cited on page 5).

Mendel, G. (1866). Versuche über Plflanzenhybriden. *Verhandlungen des naturforschenden Vereines in Brünn*, 3–47 (cited on pages 4, 19).

Moose, S. P., J. W. Dudley, and T. R. Rocheford (2004). Maize selection passes the century mark: a unique resource for 21st century genomics. *Trends in Plant Science* **9**: 358–364. DOI: 10.1016/j.tplants.2004.05.005 (cited on page 14).

Mrode, R. A. (2013). *Linear models for the prediction of animal breeding values*. 3rd ed. Boston, MA: CABI (cited on pages 6, 26).

Nielsen, R. and M. Slatkin (2013). *An Introduction to Population Genetics: Theory and Applications*. Oxford, New York: Oxford University Press. 287 pp. (cited on page 7).

O'Farrell, P. H. (1975). High Resolution Two-Dimensional Electrophoresis of Proteins. *The Journal of biological chemistry* **250**: 4007–4021 (cited on page 11).

Pearson, K. (1904a). Report on Certain Enteric Fever Inoculation Statistics. *BMJ* **2**: 1243–1246. DOI: 10.1136/bmj.2.2288.1243 (cited on pages 19, 26).

Pearson, K. (1904b). III . Mathematical contributions to the theory of evolution. XII. On a generalised Theory of alternative Inheritance, with special reference to Mendel's laws. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **203**: 53–86. DOI: 10.1098/rsta.1904.0015 (cited on page 19).

Perason, K. and A. Lee (1903). On the Laws of Inheritance in Man: I. Inheritance of Physical Characters. *Biometrika* **2**: 357–462 (cited on page 19).

Pierce, B. A. (2012). *Genetics: A conceptual approach*. 4th ed. New York: W.H. Freeman. 1 p. (cited on page 5).

Portin, P. and A. Wilkins (2017). The Evolving Definition of the Term "Gene". *Genetics* **205**: 1353–1364. DOI: 10.1534/genetics.116.196956 (cited on page 12).

Price, G. R. (1970). Selection and covariance. *Nature* **227**: 520–521 (cited on page 7).

Robertson, A. (1966). A mathematical model of the culling process in dairy cattle. *Animal Science* **8**: 95–108. DOI: 10.1017/S0003356100037752 (cited on page 7).

Rogers, S. O. (2017). *Integrated Molecular Evolution* (cited on page 12).

Rowley, M. J. and V. G. Corces (2018). Organizational principles of 3D genome architecture. *Nature Reviews Genetics* **19**: 789–800. DOI: 10.1038/s41576-018-0060-8 (cited on page 5).

Ruse, M. (2014). 'Gould, Stephen Jay'. *eLS. John Wiley & Sons, Ltd: Chichester.* American Cancer Society, pp. 1–10. DOI: 10.1002/9780470015902.a0025067 (cited on page 19).

Sanders, M. F. and J. L. Bowman (2015). *Genetic Analysis - An Integrated Approach*. OCLC: 915123587 (cited on pages 5, 15).

Santiago, E., J. Albornoz, A. Dominguez, M. A. Torot, and C. Lopez-Fanjul (1992). The Distribution of Spontaneous Mutationson Quantitative Traits and Fitness in Drosophila melanogaster. *Genetics* **132**: 771–781 (cited on page 13).

Savage, J. E. *et al.* (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics* **50**: 912–919. DOI: 10.1038/s41588-018-0152-6 (cited on page 12).
Schaeffer, L. R. (2019). *Animal models*. L. R. Schaeffer. 381 pp. (cited on pages 6, 26).
Shannon, W. D., M. A. Watson, A. Perry, and K. Rich (2002). Mantel statistics to correlate gene expression levels from microarrays with clinical covariates. *Genetic Epidemiology* **23**: 87–96. DOI: 10.1002/gepi.1115 (cited on page 12).
Sorensen, D. and D. Gianola (2002). *Likelihood, Bayesian and MCMC methods in quantitative genetics*. Statistics for biology and health. New York: Springer-Verlag. 740 pp. (cited on pages 6, 26).
Strachan, T., J. Goodship, and P. F. Chinnery (2015). *Genetics and genomics in medicine*. New York: Garland Science/Taylor & Francis Group. 526 pp. (cited on page 5).
Terman, L. M. (1917). The Intelligence Quotient of Francis Galton in Childhood. *The American Journal of Psychology* **28**: 209–215 (cited on page 18).
Tymoczko, J. L., J. M. Berg, and L. Stryer (2015). *Biochemistry, a short course*. Third edition. New York: W.H. Freeman & Company, a Macmillan Education imprint. 900 pp. (cited on page 5).
Uhlén, M., L. Fagerberg, B. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-K. Szigyarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, . Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. v. Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. v. Heijne, J. Nielsen, and F. Pontén (2015). Tissue-based map of the human proteome. *Science* **347**: DOI: 10.1126/science.1260419 (cited on page 11).
Visscher, P. M., N. R. Wray, Q. Zhang, P. Sklar, M. I. McCarthy, M. A. Brown, and J. Yang (2017). 10 Years of GWAS Discovery: Biology, Function, and Translation. *The American Journal of Human Genetics* **101**: 5–22. DOI: 10.1016/j.ajhg.2017.06.005 (cited on page 12).
Wackerly, D. D., W. Mendenhall III, and R. L. Scheaffer (2008). *Mathematical Statistics*. 8th Ed. Brooks/Cole, Cengage Learning. 946 pp. (cited on page 6).
Walsh, B. and M. Lynch (2018). *Evolution and selection of quantitative traits*. New York, NY: Oxford University Press. 1459 pp. (cited on pages vi, 7, 25).
Watson, M. A., A. Perry, V. Budhjara, C. Hicks, W. D. Shannon, and K. M. Rich (2001). Gene Expression Profiling with Oligonucleotide Microarrays Distinguishes World Health Organization Grade of Oligodendrogliomas. *Cancer Research* **61**: 1825–1829 (cited on page 12).
Wilhelm, M., J. Schlegl, H. Hahne, A. M. Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.-H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber, and B. Kuster (2014). Mass-spectrometry-based draft of the human proteome. *Nature* **509**: 582–587. DOI: 10.1038/nature13319 (cited on page 11).
Wood, A. R. *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: 10.1038/ng.3097 (cited on page 12).
Wright, S. (1921a). Systems of mating: I-V. *Genetics* **6**: 111–178 (cited on pages 10, 23, 24).
Wright, S. (1921b). Systems of mating. I. The biometric relations between parent and offspring. *Genetics* **6**: 111–123 (cited on page 10).
Wright, S. (1921c). Systems of mating. II. The effects of inbreeding on the genetic composition of a population. *Genetics* **6**: 124–143 (cited on page 10).
Wright, S. (1921d). Systems of mating. III. Assortative mating based on somatic resemblance. *Genetics* **6**: 144–161 (cited on page 10).
Wright, S. (1921e). Systems of mating. IV. The effects of selection. *Genetics* **6**: 162–166 (cited on page 10).
Wright, S. (1921f). Systems of mating. V. General considerations. *Genetics* **6**: 167–178 (cited on page 10).
Yoo, B. H. (1980). Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster* : 1. Response to selection. *Genetical Research* **35**: 1–17. DOI: 10.1017/S0016672300013896 (cited on page 13).
Yule, G. U. (1902). MENDEL'S LAWS AND THEIR PROBABLE RELATIONS TO INTRA-RACIAL HEREDITY. *New Phytologist* **1**: 222–238. DOI: 10.1111/j.1469-8137.1902.tb07336.x (cited on pages 19, 26).

Yule, G. U. (1906). 'On the theory of inheritance of quantitative compound characters on the basis of Mendels laws – a preliminary note.' Report of the Third International Conference 1906 on Genetics : hybridisation (the cross-breeding of genera or species), the cross-breeding of varieties, and general plant-breeding. London: Royal Horticultural Society (cited on page 26).

Yule, G. U. (1907). On the Theory of Inheritance of Quantitatively Compound Characters on the Basis of Mendel's Laws. *Biometrika* **5**: 481–482. DOI: 10.2307/2331701 (cited on page 19).

Zhang, Y., D. Li, and B. Sun (2015). Do Housekeeping Genes Exist? *PLOS ONE* **10**: e0123691. DOI: 10.1371/journal.pone.0123691 (cited on page 11).

Zheng, H. and W. Xie (2019). The role of 3D genome organization in development and cell differentiation. *Nature Reviews Molecular Cell Biology* **20**: 535–550. DOI: 10.1038/s41580-019-0132-4 (cited on page 5).

Zhu, J., F. He, S. Song, J. Wang, and J. Yu (2008). How many human genes can be defined as housekeeping with current expression data? *BMC Genomics* **9**: 172. DOI: 10.1186/1471-2164-9-172 (cited on page 11).