A historical perspective on QG

3

Authors: Author 1, Author 2, ..., H. Jorjani
Last update: 2020 01 05

The light of evolution

Nothing in biology makes sense except in the light of evolution.

Theodosius Dobzhansky (1973)

This chapter provides a brief description of Darwin's theories of evolution, and argues that it was the study evolution that provided the intellectual milieu in which the foundations of QG started to take shape. This also provides the link between evolution and QG.

3.1 Why do we study QG?

There are many ways of summarizing Darwin's ideas on evolution (1859) in modern terminology. The majority of introductory texts on genetics refer to Darwin's *"theory"*, and mention three principles of variation, heredity, and selection (see *e.g.* Sanders and Bowman (2015, p 51), Griffiths *et al.* (2015, p 766). In a most lucid interpretation, Mayr (1977, see also Mayr 1988, Ch 13), using the information available to him at the time of writing, postulated that Darwin had actually presented five *"theories"* in his book "On the Origin of Species" (Darwin, 1859)[1] , namely:

- Evolution as such;
- Evolution by common descent;
- Origin of diversity;
- Gradualness; and
- Natural selection.

Mayr (1977) also suggested how Darwin's ideas can be put into *"one long [chain of] argument"*[2] .

Accordingly, Darwin's theories of evolution comprise five *"facts"*, and three *"inferences"*. In the original figure by Mayr (1977) the facts and the inferences are enclosed in text-boxes. However, Mayr (1988) also believes that Darwin treated these, especially the facts, as black boxes, *i.e.* irrespective of the mechanisms explaining them, Darwin's theories of evolution are valid. As an example, the exact mechanisms of heredity for Fact 5 (inheritance of much of the individual[3] variation) were unknown at Darwin's time. Nonetheless, if the inheritance of much of variation of the individuals can be assumed (as Mendel and re-discoverers of Mendel's rules showed many years later), then Darwin's theories would withstand.

1: **Origin of Species:** Full name of Darwin's book is: "On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life"

2: **Darwin's theories:** Mayr (1988) admitted that, in the light of the discovery of Darwin's notebooks, and from a philosophical/historical point of view, this figure is not an entirely accurate representation of Darwin's chronology. However, this will not affect the conclusions that are drawn from this figure in this book.

3: **What is an individual?** Some evolutionary biologist and philosophers of biology (see *e.g.* Hull (1988)) argue that a species is an *"individual"*, and consequently another word, for example *"specimen"*, should be used for a person, an animal, or a plant. In this book, the word *"individual"* is used in its colloquial meaning.

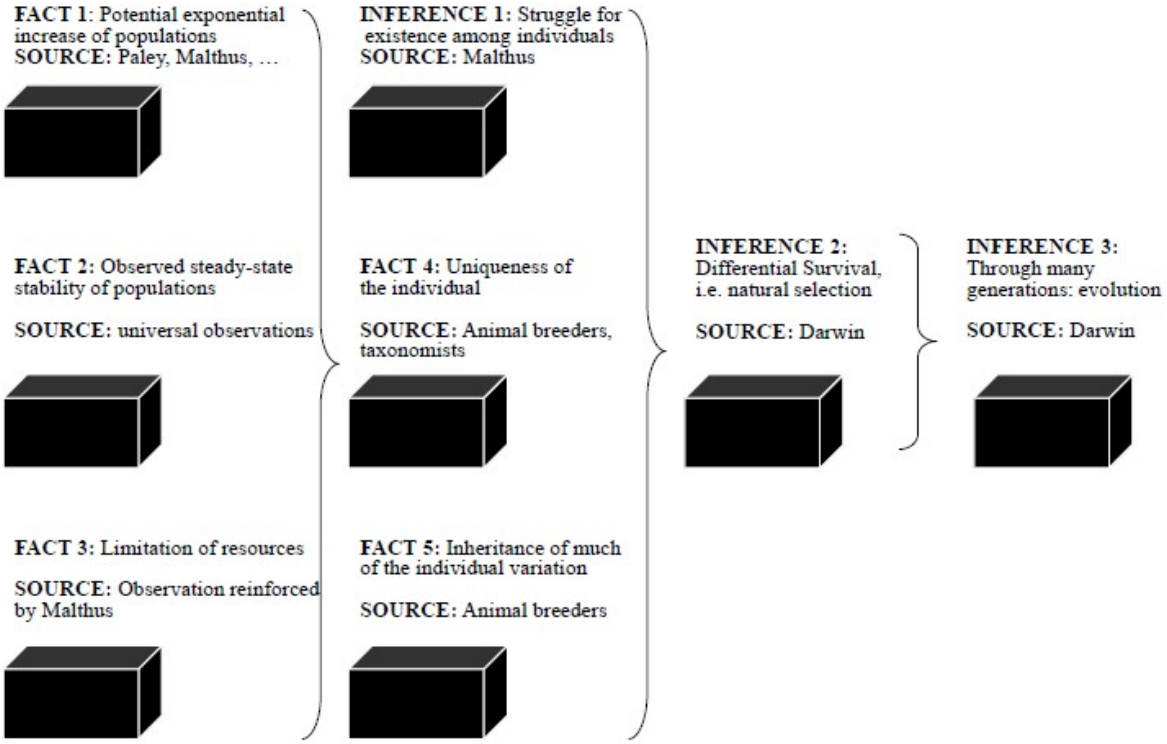

Figure 3.1: Darwin's theories - An adaptation of a figure from Mayr (1977)

From Darwin's theories (Figure 3.1) it can be seen that the three principles of variation, heredity, and selection are shown as Fact 4, Fact 5, and Inference 2, respectively.

Fact 1: The chain of argument starts with the fact established by many scientists, especially Malthus (1798), that all populations, irrespective of the rate of reproduction of the species, are capable of exponential growth. As an example of a species with very low rate of reproduction, Darwin (1859, P 64) mentions elephants. He asserts that starting with a pair of elephants, if the population growth is unchecked, the number of elephants in 500 years would grow to 15,000,000 elephants.

Fact 2: Despite the potential exponential growth in the population size, the general observation is that natural populations, within bounds of certain amount of fluctuations, exist is a steady-state stability of population size (Darwin (1859, P 67-68).

Fact 3: However, there is always a limitation of resources, especially food. Every unit of land can feed a limited number of herbivores, and a limited number of herbivores for carnivores to feed on.

Given the above three facts, it is understandable that some textbooks on population genetics start with models of population growth (e.g. Crow and Kimura (2009)).

Inference 1: The sum of the first 3 facts leads to the first inference that because of limitation of resources the potential exponential growth cannot be realized, and the relative constancy of population size is an

indication of competition among members of a population for access to the resources.

This inference had already been made by Malthus (1798). Indeed, the Malthusian inference might have been the factor that was used by Darwin to tie together his ideas. There is evidence to suggest that Darwin started to read the book by Malthus on September 28 of 1838, and he finished it within 5 days (by October 2 of 1838). It was sometime during these five days that Darwin could piece together a general sketch of his theories.

Fact 4: Darwin's own observations, as well as observations by taxonomist, and especially animal breeders, could establish the fact that each individual is unique. Darwin himself had worked with pigeons and had done experiments on them (Darwin was also member of two organized groups of pigeon breeders around London).

Fact 5: It was also obvious to Darwin, again based on the experiences of animal breeders, that offspring could inherit the individual variation expressed in the uniqueness of their parents.

There were many theories of inheritance in the middle of 1800's, including Darwin's own *pangenesis* theory[4] . Of course, none of these theories were correct. However, the nature of the inheritance is inconsequential to structure of evolutionary theory that Darwin was conceiving.

4: **Darwin's pangenesis theory:** There is no doubt that pangenesis was wrong in the sense that it could not explain many observations, and also Galton could disprove it easily by blood transfer between rabbits (for a description of Galton's work on this subject, and Darwin's reaction to it, see Bulmer (2003)). However, and ironically, there are elements of pangenesis that might be correct (see *e.g.* Liu and Li (2014)).

Inference 2: Sum of Inference 1, and Facts 4 and 5, can lead to the inference (hypothesis) that the heritable individual differences could contribute to the success (or failure) of some individuals, compared to their contemporaries, in their struggle for existence (adaptation to the prevailing conditions of life in that group of individuals), and consequently differential survival / reproduction of individuals, i.e. natural selection.

Inference 3: The long term effect of natural selection in consecutive generations, and under ever changing conditions of life will eventually lead to divergence among populations.

It is now obvious that studying quantitative genetics can contribute to the elucidation of mechanisms underlying Fact 4 and Fact 5, as well as Inference 2 and Inference 3. Further, study of animal and plant breeding can be considered as the evolution in high speed, and provide a better understanding of evolution. Studying quantitative genetics in the context of animal and plant breeding, therefore, is of mutual interest to the study of evolution, and Darwin's facts and inferences.

Our understanding of the evolution, and the evolutionary processes, has changed enormously since Darwin's times. Interestingly, even our understanding of Darwin's ideas has changed since Figure 3.1 was conceived around mid-1970's. Today, we are cognizant of all evolutionary processes[5] (forces), and how their role differs at different levels of biological complexity. For example, at the level of biological complexity that we may loosely call *"form"* and *"function"*, while we acknowledge the driving force of natural selection, we also have an abundance of evidence to accept the role of other evolutionary processes. At another level of biological complexity, *e.g.* at the DNA level, another evolutionary process, *i.e.* random drift, might be the driving force (Kimura, 1968, see

5: **Apologetic definitions of evolution:** There are five evolutionary processes: random drift, mating, mutation, migration, and selection.

also the debate between Kern and Hahn, 2018 and Jensen *et al.*, 2019 on the status of Neutral Mutation Theory).

There are many definitions of evolution[6] , most of which are more suitable for a specific branch of evolution, *i.e.* they lack generality.

Evolution

Evolution is the study of biological variation.

3.2 A brief history of QG

Publication of the *Origin of species* (Darwin, 1859) forced scientists, from all branches of science[7] , to take a stance for or against Darwin's theories of evolution. The most notable among all these scientists was Francis Galton. Insistence of Darwin on the importance of continuous variation (*i.e.* small differences among individuals) persuaded Galton to study many such traits. As early as 1865, Galton (1865) published a study entitled *"hereditary talent and character"*. In this study he used data from several extensive biographical listings, and simple arithmetic (*i.e.* proportions and percentages) of the *"counts"* of number of individuals, to show the resemblance among relatives, *i.e.* a possible role for hereditary factors. He admitted that the traits that he had studied were complex, and many factors contributed to what he called *"talent"*, or *"character"*.

Galton soon realised that *"counting"* might not be enough, and he perhaps should also use *"measuring"*, and also more sophisticated methods of numerical analysis. By 1874 (Galton, 1877), he had constructed a device (known as the Galton Board, or Bean Machine*), which he demonstrated in the Royal Society of London, to show how random processes can lead to different distributions.

In his 1877 study, Galton used published data on human stature among Americans, French, and Belgians men, and also lifting power from Belgian men. He also, with the help of his colleagues and friends, including Darwin, had performed replicated experiment with sweet-peas planted in different environments. In explaining the results of the sweet-pea experiment he used data simulated by a new version of his Bean Machine to examine possible effects of different evolutionary forces on the outcome of the experiment. Galton (1877) argued that most of the traits have distributions that follow the *"law of deviation"* or the *"law of frequency of error"*[8] . Galton's understanding of these laws can be best seen in the following lines:

> The essence of the law is that differences should be wholly due to the collective actions of a host of independent petty influences in various combinations.

For the modern reader the words "collective actions of a host of independent petty influences" reminds of the concept of *"iid"* (independent and identically distributed), which is used in invoking the Central Limit Theorem.

6: **Apologetic definitions of evolution:** Compare the definitions of a few branches of biology (such as anatomy, physiology, . . . , evolution) in a few dictionaries. Most of the definitions for *"evolution"* are unnecessarily long and include reference to evolutionary mechanisms, as if one needs to justify working with this branch of biology.

7: **Reaction to Darwin's theories:** For example, the physicist Lord Kelvin was among the non-biologists who participated in discussions. Kelvin believed that evolution could not have happened, because he wrongly assumed that the Earth had not existed long enough to allow the gradual evolution advocated by Darwin.

A little bit of history

Reading some of the old literature is excruciating painful. For example, in Galton (1865), he expresses many opinions that are horrendous. By today's moral norms he is a pure and simple racist of the worst kind. For a person whose IQ is estimated to be about 200 (Terman, 1917), he is very uncritical of many so-called *"facts"*, and he holds many contradictory opinions.

In such cases, distinguishing between the *"person"* and the *"scientist"*, and acknowledging that each of these is multifaceted, may be of some value. It is important to see how much the person and the scientist have risen above their background beliefs and knowledge.

8: **Law of frequency of error:** For all contents and purposes, these terms are equivalent to invoking the Central Limit Theorem and the resulting normal distribution. Central Limit Theorem and normal distribution will be discussed in more details in Section 5.1

* For a description of the Bean Machine see: https://en.wikipedia.org/wiki/Bean_machine. Accessed 2019-12-27.

In analyzing the sweet-pea data, Galton also laid down the foundations of the statistical methods regression and correlation. He explicitly used regression in a publication entitled *"Regression towards mediocrity*[9] *in hereditary staure"* (Galton, 1886) on the data on human stature that he himself had collected. He also explicitly spelled out the conditions for invoking the Central Limit Theorem in the following lines (Galton, 1886):

> One of [the advantages of using stature as a trait] lies in the fact that stature is not a simple element, but a sum of the accumulated lengths or thicknesses of more than a hundred bodily parts, each so distinct from the rest as to have earned a name by which it can be specified.

Even though the length (or thickness) of different bodily parts are not independently distributed, the above quote makes it clear that Galton was cognizant of the *iid* conditions as we understand them today. By 1889 Galton had collected so many empirical evidence and theoretical arguments for justification of using normal distribution to handle continuously distributed traits that a research program to be pursued by others had been established. Impact of Galton's statistical methods to study the evolution (of continuously distributed traits) was so deep and wide that many talented young scientists, with more formal mathematical/statistical training than Galton himself, directly or indirectly, became associated with his methods. Pearson and Yule are just two examples of such scientists.

9: **Old terminology:** In Galton's terminology mediocrity meant average.

3.3 From Mendelian genetics to QG

Some of Darwin's five theories, *e.g.* "evolution as such", and "evolution by common descent", were almost universally accepted immediately after the publication of the "On the origin of species" (Darwin, 1859). However, the other theories, especially "gradualness" [10] , were subjected to doubt. When the Mendelian rules became known to the scientific community in 1900, the proponents of gradualness of evolution needed to unite Galton's ideas about the continuous variation (Galton, 1889) and the discrete Mendelian factors (Mendel, 1866). Many attempts were made at uniting these two seemingly opposite scientific tenets (see *e.g.* Yule (1902), Perason and Lee (1903), Pearson (1904), Pearson (1904), Castle (1905), Yule (1907), East (1910)). However, all of these efforts lacked generality, mainly because of somewhat erroneous assumptions, *e.g.* about allele frequencies, or the degree of dominance.

In the studies published in the first two decades of the 20th century, the arguments for justification of using normal distribution was the same as suggested by Galton. For example, Yule (1902) writes the following lines:

> Surely it would be a very moderate estimate that the number of units could not be less than 50? Yet this would suffice to give, on the simplest Mendelian assumption that each unit can only exhibit two types, not some mere ten thousand different values of stature, the run of which would be quite

10: **Gradualness of evolution:** The main competing theory to gradualness of evolution is *"punctuated equilibrium"* due to Eldredge and Gould (1972). Stephen Jay Gould, who must be considered a brilliant scientist, gradually became an ardent champion of a full-fledged alternative evolutionary theory based on punctuated equilibrium. Towards the end of his life, Gould wrote a massive, but overly blown out of proportion, book (Gould, 2002) about his ideas. I, for one, believe that Gould has damaged his theories, and his legacy, by writing his last book (see also Ruse (2014)).

> indistinguishable from strictly continuous variation, but over a thousand-million million different types!

Finally, the reconciliation came in 1918 with the paper by Fisher (1918). The title of Fisher's paper *"The correlation between relatives on the supposition of Mendelian inheritance"* makes it clear that the subject is still that of Galton's efforts (*i.e.* to explain resemblance among relatives) assuming involvement of Mendelian factors. In order to do this, at the outset, Fisher (1918) embraces Galton's ideas:

> The great body of available statistics show us that the deviations of a human measurement from its mean follow very closely the Normal Law of Errors.

A little bit of history

Fisher submitted a version of his famous paper to the Royal Society of London in 1916, but it was rejected based on some obscure reasons. Thereafter, he submitted his paper to the Royal Society of Edinburgh.

This is similar to what happened to Mendel, who sent a copy of his paper to the famous German botanist Nägeli. The response from Nägeli was so indifferent, or even worse, discouraging, that Mendel did not send his paper to a journal with wider readership.

The lesson to be learned by the young scientists is "don't let your good ideas to wither away.

Then, Fisher explains how the variation in the phenotype caused by different alleles of a single Mendelian factor (locus) can be modeled. Extending the one locus model to a multiple loci model (or as Fisher puts it to a *"great number of factors"*, as long as the effect of each locus is small in comparison to the total variance, is straight forward. Having established the hypothesis that the sum of the effects of many loci would lead to (at least approximate) normal distribution, he poses the question of how the causal effects can be obtained from the data on a continuously (normally) distributed trait. Starting from the best known equation in quantitative genetics, $P = G + E$, and given the theoretical impossibility of separating genetic and environmental causes, Fisher devised the statistical method of analysis of variance to partition the variances, $V_P = V_G + V_E$. His search for causal components of variance, especially variances due to genetic variation, can be sketched as follows:

- Divide the population under study into groups of relatives and non-relatives (*e.g.* parent-offspring, twins, full-sibs, half-sibs, cousins, *etc.*;
- Measure the resemblance (correlation) within the groups of relatives, which statistically is the same as the within group covariance;
- Partition the total variance into *"observational"* variances, *i.e.* within group (V_W), and between group (V_B) variances; and
- Attribute the *"observational"* variances to *"causal"* variances.

Fisher's model will be explained in details later in this book, especially in Part V.

Thus, Fisher's paper (1918) convinced the scientific community that although the continuously distributed traits can be explained by the involvement of many Mendelian factors, the analysis of such traits cannot be restricted to *counts* and *ratios*, and most definitely requires new statistical methods that rely on continuous distributions, and especially the normal distribution. From that point of time, quantitative genetics was born, and was set out to study evolution of continuously distributed traits, and shed some light on Darwin's Fact 4 (uniqueness of the individual), Fact 5 (inheritance of much of the individual variation), Inference 2 (differential survival, i.e. natural selection), and Inference 3 (change of traits through many generations).

With regard to Darwin's Inference 2 it is important to emphasize (as it was mentioned towards the end of Section 3.1) that nowadays all evolutionary processes are studied under evolution and quantitative

genetics, as well as other branches of genetics. In this book we do not intend to explicitly rank the role of different evolutionary processes. The reason is twofold. First reason is that different parts of the QG theory are at different developmental stages. Perhaps, some parts have been easier to develop, and have been more developed. Second reason is that the needs for different parts of QG theory have been at varying levels. Perhaps, the parts that were most needed have been more developed. Therefore, we try to create a balance between the level of theory development, and the level of needs.

3.4 Demarcation lines of QG

In the chapter on the prerequisites of QG theory (Chapter 1) five other branches of genetics were mentioned. These were:

- Mendelian genetics;
- Molecular genetics;
- Biochemical genetics;
- Mathematical and statistical genetics; and
- Population genetics.

It is not easy to draw definite lines between QG theory and these neighboring fields, because there are fuzzy borders among them. In order to keep the book as accessible as possible to a wide range of readership, it is our intention to cross the borders as little as possible.

This chapter started with a brief description of the components of the evolutionary theories, as Darwin might have understood them. Darwin's emphasis on traits with continuous variation persuaded many scientists, especially Galton, to concentrate on such traits. Galton (and many other scientists) resorted to the use of normal distribution to study evolution. Rediscovery of Mendelian rules of inheritance in 1900 forced many scientist to examine how the discrete Mendelian factors could lead to the normal distribution. The general solution was provided in a publication By Fisher in 1918. Fisher's paper can be considered as the birth of the QG theory as we know it today.

Reading recommendation

For a historical perspective on early development of theories of evolution and genetics, read Mayr (1982).

For the early history of quantitative genetics, including many historical papers from the first 60 years of the 20th century, and Hill's commentary on them, read Hill (1984a; 1984b).

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).

Notation

The next list describes several symbols that will be later used within the body of the document.

c Speed of light in a vacuum inertial frame

h Planck constant

Greek Letters with Pronounciation

Character	Name	Character	Name
α	alpha *AL-fuh*	ν	nu *NEW*
β	beta *BAY-tuh*	ξ, Ξ	xi *KSIGH*
γ, Γ	gamma *GAM-muh*	o	omicron *OM-uh-CRON*
δ, Δ	delta *DEL-tuh*	π, Π	pi *PIE*
ϵ	epsilon *EP-suh-lon*	ρ	rho *ROW*
ζ	zeta *ZAY-tuh*	σ, Σ	sigma *SIG-muh*
η	eta *AY-tuh*	τ	tau *TOW (as in cow)*
θ, Θ	theta *THAY-tuh*	υ, Υ	upsilon *OOP-suh-LON*
ι	iota *eye-OH-tuh*	ϕ, Φ	phi *FEE, or FI (as in hi)*
κ	kappa *KAP-uh*	χ	chi *KI (as in hi)*
λ, Λ	lambda *LAM-duh*	ψ, Ψ	psi *SIGH, or PSIGH*
μ	mu *MEW*	ω, Ω	omega *oh-MAY-guh*

Capitals shown are the ones that differ from Roman capitals.