

R. A. FISHER. THE RELEVANCE OF THE GENETICAL THEORY OF NATURAL SELECTION

Paola Monari

Department of Statistical Sciences, University of Bologna, Via delle Belle Arti 41, I-40126, Bologna, Italia

1. THE STARTING POINT

In his famous work, *The Genetical Theory of Natural Selection* (1930), R.A. Fisher built the foundation of the genetic theory of population starting from the main statement that "... natural selection is not evolution..." (Fisher, 1930a, p. vii). Fisher's statistical background and his method of reasoning in terms of population allowed him to conceptually distinguish the two fundamental components of the 'Theory of Evolution by means of Natural Selection' put forward by Charles Darwin: (a) genetic modifications and (b) environmental pressures.

It is true that the Theory of Evolution has assumed a dominant role in scientific debate and radically modified biological thought, but it is also true that without the process which supports it, that is natural selection, the theory would not hold. At the time, Darwin did not have the tools to supply rational answers to all the questions posed by his revolutionary theory as the mechanisms of heredity processes had not yet been discovered and some fundamental structure was lacking, even if the Malthus paradigm on population growth had justified many steps of Darwin's reasoning.

Darwin had clearly seen the role of the environment in selecting advantageous mutations, however the scientific reasons behind their dynamics had still to be discovered. On this point, Fisher wrote "*Yet, although this distinction has often been made clear, its logical cogency could never be fully developed in the absence of a separate investigation of the independently demonstrable modes of causation which are claimed as its basis. The present book ... is at least an attempt to consider the theory of Natural Selection on its own merits.*" (Fisher, 1930a, p. vii)

Fisher's statistical reinterpretation of the theory of evolution introduced a rigorous quantitative formulation of Darwin's genial intuition regarding the evolutionary advantage of groups provided with greater genetic variability. Fisher was ready to take up the challenge coming from variability.

Fisher's main objective was "... to combine certain ideas derivable from a consideration of the rates of death and reproduction of a population of organisms, with the concept of the factorial scheme of inheritance, so as to state the principle of Natural Selection in the form of a rigorous mathematical theorem, by which the rate of improvement of any species

of organisms in relation to its environment is determined by its present condition.” (...) “That theorem is well known as “Fundamental Theorem of Natural Selection” which can be summarized by the following few words: “The rate of increase of fitness of any species is equal to the genetic variance in fitness. . . ” (Fisher, 1930a, pp. 22–46).

To give a coherent statistical formulation to his theory, Fisher needed to introduce the concept of chance (frequency) of death and chance (frequency) of reproduction within a factorial model of inheritance “...by which the rate of improvement of any species of organism in relation to its environment is determined by its present condition.” (Fisher, 1930a, p. 22).

With reference to a given population, Fisher suggests measuring its fitness in terms of the growth rate of the group over time as a result of the synthesis of death and birth specific rates (a Malthusian parameter). This criterion is standard in demographic analysis, but its novelty lay in using it to interpret the dynamics of population evolution.

As had Darwin, in the Malthusian paradox Fisher found the genial spark which gives a sense to the entirely fortuitous process of genetic inheritance.

Malthusian competition, together with environmental conditioning, allowed Fisher to identify the ‘experimental’ factors that break the accidental variability of genetic recombination through the generations to trace out a new genetic course.

Since 1908 the Hardy-Weinberg theorem had shown the stability of the genetic structure of a population by means of a little formula — the Newton binomial — which is a fundamental mainstay in statistics

$$(p_A + q_a)^n \quad (1)$$

where p_A e q_a are respectively the frequencies of the two alleles A and a of a given genic locus (with $p_A + q_a = 1$).

However, as it is known, the genetic structure only holds under precise conditions:

- the absence of erraticness in population structure: which means that all mating takes place at random with respect to genetic differences (this condition requires a sufficiently large population);
- the absence of constraints at mating: this assures independence in the recombination of genes and alleles (no Mendelian segregation);
- the absence of the differential rates of migration and mutation: this assures equilibrium in the exchange in genes and alleles;
- the absence of mating selection in reference to the fertility rates of genotypes.

As a departure from even one of these conditions alters the equilibrium between genic proportions, a search for these factors of disturbance becomes relevant to understanding the evolutionary processes that can be triggered. The competition between and within species is one of the factors translated by Malthus in quantitative terms for human population. As it is well known, Malthus said that human populations tend to grow exponentially until checked by resource limits: in other words they increase in geometric progression so as to double after a certain period.

The general extension of Malthus' Theory to all forms of life and to all types of available resources means it can be translated into the well known expression "natural selection". It gave Fisher the methodological instruments to interpret Darwin's theory in terms of "statistical analysis of variance". It was statistics which endowed Fisher with the formal apparatus to interpret the fundamental natural processes, just as Probability Calculus gave him the formal language to interpret the Hardy-Weinberg Theorem and its deviations, that Fisher formalized in the stochastic processes reported in Chapter IV of his *Genetical Theory of Natural Selection* under the title of "Variation as determined by mutation and selection".

2. THE FUNDAMENTAL THEOREM OF NATURAL SELECTION

Fisher opens this fundamental chapter of his work (p. 22) with two enlightening quotations, the first by Darwin: "*One has, however, no business to feel so much surprise at one's ignorance, when one knows impossible it is without statistics to conjecture the duration of life and percentage of deaths to births in mankind*" (Darwin, 1845). The second one by the genial Darwin spokesman, Thomas H. Huxley: "*In the first place it is said — and I take this point first, because the imputation is too frequently admitted by Physiologists themselves — that Biology differs from the Physico-chemical and Mathematical sciences in being 'inexact.'*" (Huxley, 1854).

Darwin's reference to statistics could not escape the greatest statistician of all time, Fisher, who could not even get away from the provocation of Huxley. Fisher's Fundamental Theorem of Natural Selection definitively transformed biology into a science that could be rigorously represented by way of a theory, all written in the language of statistics, as Darwin had already intuited.

The phenomenal structure which permits biology to be represented in mathematical and statistical terms came from the rediscovery of the laws of Mendel (1866), which at the beginning of the Twentieth century constituted a new science, the genetics. Those discoveries brought in foreground the populations viewed as statistical aggregates and reassessed variability as a factor of innovation and thus of evolution.

Fisher was able to grasp these overlooked methodological aspects and, in fact, his Fundamental Theorem is founded essentially on the genetic variance of characters of a population. To give a coherent statistical formulation to the Theorem it is necessary to introduce the concepts of death and reproduction rates into the hereditary processes.

The analogy between the effects of natural selection in all stages of life of an organism and the formal premises on which the actuarial tables of human mortality are founded offered a first criterion to evaluate the adaptation of a species measured as the relative contribution of the components of a biological population to the reproduction of the successive population.

Fisher takes up many concepts of the demography and proposes the growth rate of a stable population as a measure of fitness, calculated as syntheses of the specific death rates and of the birth rates for age classes.

The integral

$$\int_0^{\infty} e^{-mx} l_x b_x dx \quad (2)$$

describes the overall contribution of a population to its actual growth, where l_x is the number of living individuals at age x , b_x is the rate of reproduction at age x and m is known — thanks to Fisher — as a Malthusian parameter: “*In view of the emphasis laid by Malthus upon the ‘law of geometric increase’ m may appropriately be termed the Malthusian parameter of population increase.*” And he adds, almost with foresight: “*It evidently supplies in its negative values an equally good measure of population decrease, and so covers cases to which, in respect of mankind, Malthus paid too little attention*” (Fisher, 1930a, p. 26).

Through these simple formal solutions Fisher was able to describe both the growth and the extinction of a population and, as an analogy, both the affirming and the extinguishing of a genetic characteristic within a population. Let v_x be the value of future offspring of individuals aged x

$$v_x = \frac{e^{mx}}{l_x} \int_0^{\infty} e^{-mt} b_t l_t dt \quad (3)$$

v_x expresses the relationship between the expected number of descendents of an age not less to x and of the number of subjects of age x . The value v_x assumes then the significance of the average number of offspring which an individual of x age can still generate and it is strictly correlated to a Malthusian parameter of the population.

On the basis of these simple formal propositions, Fisher saw the first hints of a future progressive reduction of the numeric consistency of the English population, even if, at that time, it was still growing.

But Fisher’s methodological contribution is even greater when his formulas come to be applied to the capacity of genetic variants to reproduce and affirm themselves over successive generations under environmental pressure, which can modify the Malthusian parameter.

3. AVERAGE EXCESS AND AVERAGE EFFECT

In the measurement of natural selection it is important to define the relationships between the variability of a quantitative character in the natural population and the genetic constitutions of the organisms.

To this argument Fisher dedicates numerous pages in his *Genetical Theory of Natural Selection* and takes up the argument again in an essay from 1941 entitled *Average excess and average effects of a gene substitution*. It leads from the assumption, which was also in the thoughts of Darwin, that variability is a fundamental component of all natural and social phenomena and it is not an element of disturbance to annihilate in arithmetic mean, as all the sciences from the eighteenth and the nineteenth centuries had tried to do, orienting research to absolute certainty. Variability is a primary font of knowledge of phenomena and can become a factor of innovation. In any case, it is the primary

symptom of a dynamic reality that produces change in a population.

Due to his sensitivity as a great statistical methodologist, Fisher understood the cognitive value of variability in the explanation of phenomena and he transcribed it into his own statistical language that was able to plumb the depths of all the components of variability in the search for explanatory factors.

Let us go back to his words: *“We may imagine, in respect of any pair of alternative genes, the population divided into two portions, each comprising one homozygous type together with half the heterozygotes, which must be divided equally between the two proportions. The difference in average statures between these two groups may then be termed the average excess (in stature) associated with the gene substitution in question.”* (Fisher, 1930a, p. 30). After having made amply clear the empirical and methodological requisites necessary to measure the average excess between two alternative genes (modern genetics calls them alleles), Fisher specified that this measurement is always an estimate of true value, as it is carried out over a limited number of observations. Consequently the average excess has to be compared with the statistical randomness intrinsic to all sampling estimation. In this sense, the difference between the averages can be correctly evaluated by imagining a theoretic population including all the possible genetic combinations weighted by their respective probability of survival.

The average values of the character must be calculated on the basis of the values which these genotypes assume by making the environmental situation vary in all possible modes. That is, to annul the effect of irrelevant circumstances, thus to correctly evaluate the relationship between ‘genetic factors’ and ‘phenotypical results’. In these methodological recommendations, Fisher achieves the happy fusion of two essential logical moments: the design of experiments and the analysis of variance. Also his revolutionary reassessment of statistical inference was entirely based on the analysis of variability, from the decomposition of the components of correlation to the analysis of variance-covariance and discriminant analysis. When, however, the problem is to evaluate the effect on a character produced by the substitution of one allele with its alternative within a data population, it is necessary to introduce a second quantity that must not be confused with the former, that is *“... the average effect produced in the population as genetically constituted, by the substitution of the one type of gene for the other”* (Fisher, 1930a, p.31).

The genetic variance of a character, in the Fisherian formulation, is therefore a synthesis of average excess and average effect and corresponds to a Darwinian measure of fitness of a population. In the hypothesis of a polyfactorial genetic constitution, if the effect of each factor is additive, the quantity $pqhk$ represents the contribution, in terms of fitness, of each factor at the genetic variance of the character. In the quantity $pqhk$, p and q are the probabilities associated with two allelic states of a character, k and h are the respective average effects. The total genetic variance W is therefore given by the sum of the individual contributions:

$$\sum pqhk = W \quad (4)$$

This result is the significant synthesis of a plurality of effects ascribable to the combination of the two great components of variability:

1. The first, V_a , considers the additive variations in the individual outcome of the

character and takes into account the difference between the homozygotes. The basis of the calculation of V_a comes from the difference between the measurement of the characters in the homozygotic genotypes. In the hypothesis of random mating without dominance, the genes behave additively and therefore the measurement is obtained by the sum of those differences.

2. The second, V_d , produces variability when, in the presence of dominance, the heterozygote systematically modifies the phenotype and the measure of the character is not intermediate in respect to the homozygotes.

In short, if we set some essential “statistical” conditions which are easy to verify in the large numbers of statistical populations, such as independence and non dominance, the genetic variation W will always be additive and it can be decomposed additively:

$$W = V_g = V_a + V_m + V_i + V_d \quad (5)$$

The components V_a and V_d are already defined, while V_m measures the variability around the average level of the character when the hypothesis of random mating falls and a preferential choice at mating is realised. V_i , on the other hand, expresses the effect of the interaction between genes when the hypothesis of independence falls.

“The appropriateness of the term genetic variance lies in the fact that quantity X (expected value) is determined solely by the genes present in the individual, and is built up of average effects of these genes. (...) Without obtaining individual values, the genetic variance of the population may be derived from the correlations between relatives, provided these correlations are accurately obtained. For this purpose the square of the parental correlation divided by the grandparental correlation supplies a good estimate of the fraction, of the total observable variance of the measurement, which may be regarded as genetic variance.” (Fisher, 1930a, p. 33)

It is evident that even little variations in the values of p or q in respect to the involved alleles modify the variance W . The equation

$$\sum pqhkd t = W dt \quad (6)$$

measures the total increase of fitness.

From Eq. (6) the main proposition of Fisher’s fundamental theorem it follows: *“The rate of increase in fitness of any organism at any time is equal to its genetic variance in fitness at that time”* (Fisher, 1930a, p. 35).

Though the statement appears concise, its phenomenal contents have a huge impact. To measure the rate of increase in fitness statistically, one has to identify an estimate of the expected rate of the variation of fitness. The difference between real values and expected values of p in the presence of chance fluctuations of genetic frequency is $pq/2n$ (the denominator is equal to the number of alleles of character considered in a population of n individuals).

The variance of the increase in fitness becomes

$$V(k \delta p) = k^2 pq/2n \quad (7)$$

where k is still the average effect consequent to the measurement of p . In a stable population, in the sense of Hardy Weinberg (Monari and Scardovi, 1989) it is $h = k$, therefore the standard error of the increase of fitness in a generation is

$$< \sqrt{\frac{W}{2n}} \quad (8)$$

When n increases, the effects of random fluctuation on W tend to become insignificant even within the same generation.

These variations can be determined by diverse factors and they feel the effects of a differential aptitude of the phenomena to affirm themselves through the generations, that is of the diverse survival probability of the genes through the generations. A gene carried by an individual in reproductive age could be present in the successive generation in a variable number of individuals. The probabilities that individuals will also reach the age to hereditarily transmit the gene can be described by a stochastic process (Kimura, 1964). For this the strength of numbers in the large species enters the core of the concept of fitness stated in the *Fundamental Theorem of Natural Selection*. Obviously, the concept of fitness preserves its presupposition even in the differential variations between the genotypic frequencies that alter the genetic variance of a population.

Only after a careful statistical analysis of impressive collections of data, Fisher maintained to have shown on empirical basis that the most numerous species have the greatest evolutionary capacity and therefore a higher probability of substituting the species in competition. Concerning this, Fisher's continuing preoccupation regarding the quality and quantity of data, which never falls away from the cognitive objective, is extraordinary.

4. THE ROLE OF VARIABILITY IN FISHER'S THOUGHT

In the history of the genetics of populations, Fisher (1930a), Haldane (1924–27) and Wright (1931), contended for a temporal supremacy, but it was later judgement which delivered the absolute verdicts. Without taking anything from the contribution of his competitors, Fisher is distinguished by his methodological broadness and for the rigour of his statistical language.

The underlying thread of Fisher's theory is the study of variability of one or more characters with reference to one or more groups and of the factors that can explain them. The taxonomy of Linneus had enabled Darwin to recognise within the interweaving of similarities and differences between the species the symptoms of a nature that evolved, the same taxonomy which suggested to Fisher to observe variability in a new way and to discover the part of variability (heterogeneity) between the individuals within a group and of variability (divergence) that distinguishes the groups and renders them systematically diverse among themselves.

New biology — like new physics — definitely left the epistemological models of the nineteenth century in its wake. After the normalisation of Galilean science, the whole nineteenth century had tried to bring the human sciences back to rules of a deterministic description devoted to annulling the variability of natural and social phenomena,

interpreted as an element of disturbance, rather than as a symptom of evolution and innovation. The work of Adolphe Quetelet (1794–1874) is emblematic in that he based his research on the anthropometrical characteristics of man and his social behaviour, so to describe that which will be defined as the Queteletian ‘average man’. Quetelet based his arguments exclusively on the theory of the instrumental errors of measurement and on an interpretation of variability modelled by the Gaussian curve. His way of reasoning contained a logical fallacy, since he attributed the observed phenomena the same causal premises that in other contexts had determined that curve.

With Darwin, the study of variability thus became an element of innovation in statistic methodology, that fifty years later Fisher would develop in many aspects, from the analysis of variance to the design of experiments and discriminant analysis. In 1919, when confronted with a choice of either a post at University College offered by Karl Pearson or a post at the Rothamsted Centre offered by John Russell, he chose the latter. Kendall writes, “*There was never a happier appointment. Over the next fourteen years Fisher established Rothamsted as one of the Holy Places of the statistical world*” (Kendall, 1963, p. 439)

He was fully aware that the study of variability is exclusively statistical and that at every stage it is necessary to evaluate if the differences between two individuals, two groups or between an individual and a group are significant or not. It was on this understanding that Fisher designed his methods.

These are methods which did not come from theory, but from the experimental research which Fisher has continuously got ahead throughout his career and reposed in all his principle publications as if it were a permanent theme. And this does not mean laboratory research, but research conducted on observed data where the separation of the relevant factors from those irrelevant was not easy, nor automatic. In consequence, the recognition of ‘accidental’ variability from ‘systemic’ variability generated by experimental or discriminant factors became just as important.

All the analysis of variance is founded on this capacity to distinguish or separate the two types of variability and the methodological key rationalized by Fisher is the design of experiments, which suggested to redistribute accidental variability, thus to render it non influential in respect to the recognisability of the systematic components. Fisher therefore re-established the randomization of experimental design with the same probabilistic logic of random sampling which enabled the justification and construction of his most known sample distributions (Fisher, 1935, 1966). The single sample is nothing more than an empirical determination of a random variable that in statistical research performs the role of a catalyst that permits the settlement of accidental variability rendering it recognisable and separating it from that which is systematic.

The design of the experiments in all its variants, and the developments of the analysis of variance that followed, have allowed statistical research to be given the same persuasive force as more traditional experimental research. In *The Design of Experiments* (1935), Fisher writes: “... a secure basis of new knowledge” and again, recalling Darwin “... the method of comparison adopted by Darwin is that of pitting each self-fertilised plant against a cross-fertilised one, in condition made as equal as possible. The pairs so chosen for comparison had germinated at the same time, and the soil conditions in which they grew were

largely equalized by planting in the same pot. . . . The evident object of these precautions is to increase the sensitiveness of the experiment, by making such differences in growth rate as were to be observed as little as possible dependent from environmental circumstances, and as much as possible, therefore, from intrinsic differences due to their mode of origin". The best rule suggested by Fisher for selecting the pairs to be submitted under different experiments is to "assign at random, as by tossing a coin. . . . Since each particular error has thus an equal and independent chance of being positive or negative, the error of our average will necessarily be distributed in a sampling distribution, centred at zero, which will be symmetrical. . . ." (Fisher, 1935, 1966, pp. 32, 43).

All this permitted Fisher to write the genetics of populations in terms of genetic variance — which phenomenally defines the fitness capability of an individual or group — and to decompose it into specific genetic components (average excess and average affect) and into environmental components (selective effect).

Fisher was therefore not limited to constructing methods to explain phenomena through variability thus to formulate scientific laws. Taking advantage of the innumerable information contained in variability, Fisher suggested absolutely new solutions for the attribution of single individuals to the respective groups they belong. We are referring to the discriminant analysis. The capacity to correctly assign an observed individual to a class is simultaneously both analysis and synthesis, and it is also one of the most significant moments of any process of knowledge since every correct attribution of an individual to its own class is a renewed confirmation of the statistical laws which include that class.

Obviously the question is posed when the attribution occurs on the basis of quantitative characters that, singularly taken, can place the same individual in more than one class. Discriminant analysis is one of the most suggestive examples of the meeting between statistical method and scientific research and is based on the best combination of observed characters that maximise the separation between the classes: the so called discriminant function. Once again Fisher's trick consists of the comparison between the diverse components of variability (within and between the classes) which constitutes the cognitive basis of this method.

At the basis of the methodological achievements of Fisher and the multitude of techniques which, at the time, also revitalised statistics, as well as the other sciences which have borrowed his language, there is a brilliant idea that Fisher developed in all the possible exceptions: that of sample distribution as a probabilized random variable. It was a milestone in statistical methods which has had extraordinary effects, not only in statistical inference, but also in the interpretation of natural phenomena.

Not only this, but by rewriting the rules of combinatorial analysis, Fisher was able to transform a fact, an event or an experimental result into an outcome of an exhaustive space of possibilities, each of which can be associated with a measure of probability.

5. A DARING ANALOGY WITH THE SECOND LAW OF THERMODYNAMICS

To clarify the stabilizing effect of dimension n of a population with respect to the accidental variations of the individual gene composition, Fisher introduced an audacious

analogy between his *Fundamental Theorem of Natural Selection* and the greatest proposition of physics of his time, the *Second Law of Thermodynamics*. Fisher (1930a) wrote that there were many points in common: “Both are properties of populations, . . . both are statistical laws; each require the constant increase of a measurable quantity, in the one case the entropy of a physical system and in the other the fitness. . . of a biological population.” Moreover, this analogy was advanced at the beginning of the Twentieth century by Boltzmann, who also recognised a visible snare (and at the same time a confirmation) in his Laws, in terms of potential negative entropy, when he wrote: “The general struggle for existence of animate beings is therefore not a struggle for raw materials. . . but a struggle for entropy” (Boltzmann, 1905, p. 40).

The attraction for the new *weltanschauung* tinged with Darwinian theory certainly intrigued the complex intellectual and scientific personality of the great physicist, even if the statistical interpretation of the evolutionary processes of the living population would emerge fully only a few years later with the rediscovery of the Mendel’s Laws and affirmation of genetics. Molecular genetics, like particle physics, took combinatorial calculus out of the game, transformed the microstates represented by the potential random mating between the elementary genetic components in probabilized macrostates, expressed his laws in an intrinsically statistical language and therefore gave demonstrative strength to the Theory of evolution for natural selection.

Boltzmann quickly understood how much was occurring in science, also because he himself was a protagonist of this revolution, and as early as 1905 he announced his famous claims: “If you ask me for my innermost conviction whether it will one day be called the century of iron, or steam, or electricity, I answer without qualms that it will be named the century of the mechanical view of nature, the century of Darwin” (Boltzmann, 1905, p. 28). The search for a common language between physics and biology suitable for interpreting the new discoveries of these two *avant-garde* sciences was completed in the first forty years of the twentieth century.

The first statistician who realised he had the privilege of possessing this language and was able to transform it into the most powerful method of scientific research was Ronald A. Fisher. He brought together the immanency of this new way of managing the natural sciences (physics and biology) founded on an essentially statistical language and he was fascinated by the force in which, at the same time, such a physicist as Eddington was asserting the Second Law of Thermodynamics.

In Boltzmann’s thermodynamics the direction of physical phenomena is irreversible and goes on towards an entropic degradation of the system from order to disorder. It is the same for life when it evolves from one state to another which is not necessary but only probable. “And even if it were be possible to know all the initial conditions of a system at a certain point in time, it would not be possible to predict its state at a different point: the only thing one can derive from it is the distribution of probabilities of the possible states” (Scardovi, 1983, p. 253).

As a refined statistician, Fisher understood the language of statistical physics, but he knew the language of genetics even better and did not therefore refute the analogy between the two sciences. Nevertheless he was prudent and with great acumen underlined the difference between the two fundamental laws that have revolutionized modern sci-

ence: ... we should note that the laws as they stand present profound differences. (1) The systems considered in thermodynamics are permanent; species on the contrary are liable to extinction, although biological improvement must be expected to occur up to the end of their existence. (2) Fitness, although measured by a uniform method, is qualitatively different for every different organism, whereas entropy, like temperature, is taken to have the same meaning for all physical systems. (3) Fitness may be increased or decreased by changes in the environment, without reacting quantitatively upon that environment. (4) Entropy changes are exceptional in the physical world in being irreversible, while irreversible evolutionary changes form no exception among biological phenomena. Finally, (5) entropy changes lead to a progressive disorganization of the physical world, ... while evolutionary changes are generally recognized as producing progressively higher organization in the organic world." (Fisher, 1930a, p. 37).

In these extremely advanced remarks, Fisher clearly distinguished between statistical determinism, that guides the passage of heat towards its most probable entropic state in a practically irreversible way, and statistical indeterminism of the evolutionary becoming of populations of organisms, which is also an irreversible and non predictable process, but which is not entropic, because it always converges towards a newly organised system. This message was taken from the most shrewd physicists when confronted with the similarities and differences between the respective paradigms. Among the greatest physicists who have involved themselves in this foundational problem is Erwin Schrödinger. In 1944 he wrote in *What is Life?*: "What is the characteristic feature of life? When is a piece of matter said to be alive? When it goes on 'doing something', moving, exchanging material with its environment, and so forth, and that for a much longer period than we should expect an inanimate piece of matter to 'keep going' under similar circumstances. When a system that is not alive is isolated or placed in a uniform environment, all motion usually comes to a standstill very soon as a result of various kinds of friction; differences of electric or chemical potential are equalized, substances which tend to form a chemical compound do so, temperature becomes uniform by heat conduction. After that the whole system fades away into a dead, inert lump of matter. A permanent state is reached, in which no observable events occur. The physicist calls this state of thermodynamical equilibrium, or of 'maximum entropy'." (Schrödinger, 1967, p. 74).

Well then, what makes the phenomena of life different from the phenomena of physics governed by the Second Law of Thermodynamics? A response can be found by looking at micro processes. Any time an error of replication (mutation) occurs at the subatomic level of DNA and of its components, whatever the internal or external causes to the system of the 'living organism', the thermodynamic equilibrium is altered. This is how a new process is set in motion towards a different equilibrium that nevertheless can alter itself at any time owing to a new 'error', which is a chance event, that is not predictable and not programmable. If we consider all the combinatorial alternatives present in the structure of a genome, which in the process of hereditary transmission are decomposed in four possible matings for every site, a single genome constituted of n couples of nucleotides gives 4^n possible alternatives. If one considers that a single chromosome can contain thousands of genes and that each gene can contain thousands of pairs of nucleotides, one can understand how the number of possible combinations

is phenomenally equivalent to mathematical infinity, such to make it absolutely unforeseeable whether they are single microstates or the consequent macrostates.

This essentially indeterministic character of the phenomena of life, which is non-predictable in results, finds a winning strategy in statistical methodology and in the language of probability to interpret and describe the great natural (biology and physics) and social phenomena (economics, finance, sociology and communications).

ACKNOWLEDGEMENTS

My thanks go to Prof. Italo Scardovi, who introduced me to the fascination of these topics and taught me to read natural phenomena in the statistical language.

REFERENCES

- F. AYALA (1982). *Population and evolutionary genetics*. The Benjamin Cummings Publisher, Menlo Park.
- L. BOLTZMANN (1905). *Populäre Schriften*. J.A. Barth, Leipzig. Reprinted in 1919 and 1925.
- M. BORN (1959). *Natural philosophy of cause and chance*. Cambridge University Press.
- C. DARWIN (1845). *Life and Letters of Charles Darwin*. John Murray, London.
- C. DARWIN (1872). *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. John Murray, London.
- R. FISHER (1930a). *The genetical theory of natural selection*. Clarendon Press, Oxford.
- R. FISHER (1930b). *Statistical method for research workers*. Oliver and Boyd, Edinburgh.
- R. FISHER (1935, 1966). *The design of experiments*. Oliver and Boyd, Edinburgh.
- R. FISHER (1936). *The use of multiple measurements in taxonomic problems*. Ann. Eug., 7.
- R. FISHER (1941). *Average effect and average excess of a gene substitution*. Ann. Eug., 11.
- R. FISHER (1953). *Population genetics*. Proc. Roy. Soc. B, 141.
- F. GALTON (1889). *Natural inheritance*. MacMillan, London.
- J. HALDANE (1924–27). *A mathematical theory of natural and artificial selection*. Proc. Cambridge Phil. Soc., 23.
- G. HARDY (1908). *Mendelian proportion in a mixed population*. Science, 28.
- T. HUXLEY (1854). *On the educational value of the natural history sciences*. Collected Essays, III.
- M. KENDALL (1963). *Ronald aylmer fisher, 1890–1962*. Biometrika, 50.
- M. KENDALL (1970). *Ronald aylmer fisher, 1890–1962*. In E. PEARSON, M. KENDALL (eds.), *Studies in the History of Statistics and Probability*, Charles Griffin and Co., London.

- M. KIMURA (1964). *Diffusion models in population genetics*. Methuen, London.
- T. MATHUS (1798). *Essay on principle of population*. London.
- G. MENDEL (1866). *Versuche uber pflanzenhybriden*. Verhan. Naturf. Ver. Brunn, IV.
- P. MONARI, I. SCARDOVI (1989). *I fondamenti statistici dell'equilibrio genetico nelle popolazioni*. Ed. Martello, Milano.
- J. MONOD (1970). *Le hazard et la nécessité*. Le Seuil, Paris.
- K. PEARSON (1903–04). *Mathematical contribution to the theory of evolution*. Trans. Roy. Soc., 21–22.
- I. SCARDOVI (1982). *The idea of chance in statistical intuition of natural variability*. Genus, 3–4.
- I. SCARDOVI (1983). *Chance and order in a statistical picture of life*. Epistemologia, VI, 2.
- E. SCHRÖDINGER (1967). *What is life?: the physical aspect of the living cell; Mind and matter*. Cambridge University Press.
- L. C. SFORZA, A. BODMER (1971). *The genetic of human population*. Freeman, San Francisco.
- S. WRIGHT (1931). *Evolution in mendelian populations*. Genetics, 16.

SUMMARY

R. A. Fisher. The relevance of the genetical theory of natural selection

Starting from the main statement that "... natural selection is not evolution...", R.A. Fisher built the foundation of the genetic theory of population in his famous work *Genetical Theory of Natural Selection* (1930). He rewrote the scientific paradigm proposed by Darwin in statistical terms using the calculus of probability and, most importantly, statistics. The key to his formal transposition is in the analysis of variance in which Fisher interpreted as phenomenal variability by means of random variability: this completely original result would become a fundamental chapter of statistical method. It is not by chance that at the same time he published his statistical method for research workers in which the analysis of variance dominated his primary elements of the design of experiments.