REVIEW ARTICLE

The birth and development of the DNA theory of inheritance: sixty years since the discovery of the structure of DNA

PETTER PORTIN*

Laboratory of Genetics, Department of Biology, University of Turku, FI-20014 Turku, Finland

Abstract

The development of the DNA theory of inheritance culminated in the publication of the molecular structure of DNA 60 years ago. This paper describes this development, beginning with the discovery of DNA as a chemical substance by Friedrich Miescher in 1869, followed by its basic chemical analysis and demonstration of its participation in the structure of chromosomes. Subsequently it was discovered by Oswald Avery in 1944 that DNA was the genetic material, and then Erwin Chargaff showed that the proportions of the bases included in the structure of DNA followed a certain law. These findings, in association with the biophysical studies of Maurice Wilkins and Rosalind Franklin with Raymond Gosling, led James Watson and Francis Crick to the discovery of the double-helical structure of DNA in 1953. The paper ends with a short description of the development of the DNA theory of inheritance after the discovery of the double helix.

[Portin P. 2014 The birth and development of the DNA theory of inheritance: sixty years since the discovery of the structure of DNA. *J. Genet.* **93**, 293–302]

Introduction

'We have discovered the secret of life'
Francis Crick, 28 February 1953

There are many ways to organize the course of the history of genetics. One, and a very illuminating one, is to describe the development and gradual emergence of our knowledge of the nature of the genetic material, which is what is attempted here. This method is very topical since during the present year we celebrate the 60th anniversary of the discovery of the molecular structure of DNA, the substance genes are made of, by American James D. Watson and Briton Francis H.C. Crick (Watson and Crick 1953a).

The model published by Watson and Crick has since been confirmed by several physicochemical and electron microscopical studies, and is still completely valid. In fact, the famous double-helix model has become the symbol of all of biology, and even a sort of cultural icon for the modern age.

After one and a half year's hard work, Watson and Crick perceived the model in Cambridge, England, on 28 February 1953, at that time Watson was only 24 and Crick, 36 years old. The birth of hardly any other scientific discovery can be dated so precisely—certainly not for any as important as this. The journey leading to the discovery, and the incidents following it, has been excitingly illuminated by James D. Watson in his famous book *The double helix* (Watson 1968). The article in which the model was published for the first time on April 25th 1953 is very short, consisting of only ca. 800 words and one figure. This short form was probably due to the competitive situation the authors were in, according to the American historian of molecular biology Horace F. Judson (Judson 1996, pp. 167). At least Watson felt that it was a question of a race, the rivals being the American chemist Linus C. Pauling and the British biophysicist and x-ray crystallographer Rosalind E. Franklin (Watson 1968).

Very soon after the publication of the model, however, Watson and Crick published a more detailed description of the structure of DNA, including an explanation of the genetic implications of the model (Watson and Crick 1953b). An even more extensive commentary on the matter followed in the following year's Cold Spring Harbor Symposia on Quantitative Biology (Watson and Crick 1954).

With its strong explanatory power, the Watson–Crick Model is ingenious; it accounts for all the four general properties required for genetic material. Firstly, the model explains the replication of the genetic material which for its part is the basis for reproduction, a definitive characteristic for life. Secondly, the model explains the specificity of genetic material, i.e., the quality of the genes, and how this

*E-mail: petter.portin@utu.fi.

Keywords. heredity; chromosome; nucleus; genetic material; history of genetics.

specificity is preserved in the duplication process. Thirdly, the model explains the information content of the genetic material: DNA is an informative macromolecule. Fourthly, the model explains the ability of the genetic material to change, i.e., the ability of the genes to mutate. These four important and necessary properties of the genetic material will be dealt with more closely later.

In the pages of *Nature* following the first paper of Waterson and Crick (1953a), the x-ray crystallographical works of Briton Maurice H. F. Wilkins and his coworkers as well as of his compatriots Rosalind E. Franklin and Raymond G. Gosling were published (Wilkins *et al.* 1953; Franklin and Gosling 1953). These works confirmed the model of DNA presented by Watson and Crick (1953a). Subsequently several biophysical and electron microscopical studies supporting the model have appeared.

Watson, Crick and Wilkins were awarded the Nobel Prize in physiology or medicine in 1962 'for their discoveries concerning the molecular structure of nucleic acids and its significance for information transfer in living material'. Rosalind Franklin would have absolutely been worthy of the prize also, but she died in 1958 at the age of only 37 years due to ovarian cancer. She was left without the prize because, according to the rules, it cannot be awarded posthumously nor be divided between more than three persons. What could have been done if she had been alive? Could it have been organized in such a way that the Nobel Prize in physiology or medicine would have been awarded to Watson and Crick and that in chemistry to Wilkins and Franklin?

The early stages of DNA research

DNA as a chemical substance as such was found by the Swiss physician and biochemist Friedrich Miescher in human leucocytes as early as 1869, which is noteworthy as this is at the same time as the German-Silesian scientist and Augustinian friar Gregor Mendel, the father of modern genetics, discovered his laws of inheritance (Mendel 1866). Miescher worked in the laboratory of Professor Felix Hoppe-Seyler at The University of Tübingen, Germany and was interested in the content of the cell nucleus in the chemical sense. The material for the studies was obtained from the pus of the surgical bandages from the local hospital. Pus contains large amounts of leucocytes, the nuclei of which were carefully purified. From these, Miescher was able to isolate a completely novel organic substance which he named, according to its origin, nuclein (Miescher 1871). Today, we know that the substance was DNA.

Nuclein was different from all other organic substances isolated from cells because, for example, of its exceptionally high content of phosphorus. At that time, this caused both attention and disbelief. Even Miescher's professor, Felix Hoppe-Seyler, the most influential representative of organic chemistry of the time, was so skeptical that he wanted to repeat the experiments for himself before he gave Miescher permission to publish the results (Dahm 2008). Because of this, the publication of the discovery was delayed by two years.

A little later, Miescher realized that fish sperm, milt, would be an ideal material for his purposes. Milt consists of very large cells containing, in addition to the nucleus, practically no cytoplasm at all, and, further, milt was also easy to obtain in large quantities. Accordingly, Miesher, now working in Basel, Switzerland, isolated nuclein from the milt of the salmon from the Rhine river, and the preparation was even more pure than that obtained from human leucocytes. Using this material, he was able to confirm that nuclein contained no sulfur, which had occurred as an impurity derived from proteins in the leucocyte preparations. Likewise, he could confirm the high content of phosphorus in nuclein, and measured its value almost correctly. Of importance was also his observation that all phosphorus in nuclein was present in the form of phosphoric acid (Miescher 1874a, b). At the same time Miescher also extended his studies to carp, frogs, chicken and bulls, and found nuclein in the sperm samples of all these species (Miescher 1874a, b).

During Miescher's lifetime, already several researchers started to solve questions raised by his work. Better methods for the purification of nucleic acids were developed. The German pathologist and histologist Richard Altmann, a pupil of Miescher, believed that he had isolated a completely new substance, which he named nucleic acid because it behaved like an acid in chemical reactions (Altmann 1889). However, he did not realize that it was precisely the same substance which Miescher had named nuclein (Dahm 2008).

A few years later several other biologists were able to show that nucleic acid was a part of the chromosomes, the first, however, was the botanist Edward Zacharias in 1884 (Mirsky 1968). In 1893, the German biochemists Albrecht Kossel and Albert Neumann demonstrated that the nucleic acid contained four different bases (Kossel and Neumann 1893). Further, Kossel observed that nuclein was a part of chromatin, the substance which composes chromosomes together with proteins such as histones, which he also discovered (Olby 1994; Portugal and Cohen 1977). On the basis of his studies, Kossel also made the important conclusion that nucleic acids are substantially involved in the synthesis of new cytoplasm during growth and replacement (Kossel 1913).

In spite of these great advances, the significance of nucleic acid remained obscure for decades, and interest in studying them gradually diminished until a renaissance in the 1930s.

Proof that DNA is the genetic material

The chromosome theory of inheritance

As mentioned earlier, at the beginning of the 20th century it was already known that DNA was a part of chromosomes. It is therefore informative to begin this section with a short

history of the theory of the chromosome as the material basis of inheritance.

The chromosome theory of inheritance, i.e., the theory that genes are located in the chromosomes of the cell nucleus, was created in 1902–1904 by the German biologist Theodor H. Boveri and the American geneticist and physician Walter S. Sutton, soon after the so called rediscovery of Mendel's laws of inheritance (Boveri 1902, 1904; Sutton 1903).

Boveri (1902) found evidence that the individual chromosomes in the sea urchin *Paracentrotus lividus* possess different qualities. Moreover, he was able to show that for normal development, a particular combination of chromosomes was more important than a particular number. In his studies of the divisions of the eggs of *Ascaris megalocephala*, a parasitic nematode worm, he also demonstrated an essential property of the chromosomes, namely their continuity (Boveri 1903). By continuity is meant that the chromosomes preserve their identity from one cell generation to the next which is a necessary feature of the genetic material. This observation was made by following the cycle of disappearance and appearance of the longitudinal morphology of the chromosomes during mitosis, particularly that of the chromomeres.

Sutton (1903) for his part, on the basis of his studies on the spermatogenesis of *Brachystola magna*, a large grasshopper drew attention to the resemblance between the separation of homologous chromosomes at the meiotic divisions and Mendel's postulated separation of character differences at gamete formation. He also pointed out that the independent orientation of each homologous chromosome pair, or bivalent, at the first meiotic metaphase parallels the independent inheritance of different character differences observed by Mendel. In other words, he explained Mendel's laws of inheritance, the law of segregation and the law of independent assortment on the basis of the behaviour of chromosomes during the course of meiosis.

The chromosome theory of inheritance was proven during the course of the 1910s by the American embryologist and geneticist Thomas Hunt Morgan and his school, who worked using the fruit fly *Drosophila melanogaster* as the experimental organism (Morgan *et al.* 1915; Morgan 1919, 1926).

First Morgan himself explained the sex linkage of the white eye colour character in *Drosophila*, a phenomenon that he had identified (Morgan 1910), by assuming that the respective gene was located on the sex-determining X chromosome (Morgan 1911). The chromosomal basis of sex determination had earlier been discovered by C. E. McClung (1902) and the American zoologist and geneticist Edmund B. Wilson (1905), findings which as such supported the chromosome theory of inheritance.

Following this, Alfred H. Sturtevant, a pupil of Morgan, laid the foundations of the mapping of genes by showing that the linkage map of genes was linear, corresponding to the linear structure of the chromosome, thus supporting the chromosome theory of inheritance (Sturtevant 1913). The first direct proof of the theory was somewhat later obtained by Calvin B. Bridges, another of Morgan's pupils, who observed that the nondisjunction, i.e., lack of segregation in gamete formation, of sex-linked genes in *D. melanogaster* was accompanied by an analogous nondisjunction of the X chromosomes in meiosis (Bridges 1914, 1916).

Additional direct evidence for the chromosome theory of inheritance was obtained in 1929 by the American geneticist Hermann J. Muller, again a pupil of Morgan, in collaboration with Theophilus S. Painter, likewise an American geneticist, and by the Ukrainian-American geneticist and evolutionary biologist Theodosius G. Dobzhansky, working at that time in Morgan's laboratory at the Columbia University, New York. They observed that structural changes in *D. melanogaster* linkage groups, following x-ray irradiation, were associated with corresponding changes in the chromosomes (Muller and Painter 1929; Dobzhansky 1929). Their work finally established that the order of the genes on the linkage map is the same as their physical order on the chromosome. Further support for this hypothesis was obtained when it was demonstrated by Harriet B. Creighton and Barbara McClintock in maize (Creighton and McClintock 1931) and Curt Stern in *D. melanogaster* (Stern 1931) that genetic recombination between linked markers was accompanied by physical exchange of cytologically marked chromosome segments.

The power of cytogenetical analysis was greatly increased as a consequence of the discovery of giant chromosomes in the cell nuclei of the salivary glands of *Drosophila* and other dipteran larvae. These chromosomes were first used by Painter (1933, 1934) to provide a detailed cytological map of the X chromosome of *D. melanogaster*. He showed that the succession of salivary chromosome bands corresponded to the linear order of blocks of genes in the linkage map of this chromosome. The physical mapping of genes reached its culmination when Bridges (1935) presented detailed cytological maps of all the chromosomes of *D. melanogaster* based on the study of these giant salivary gland chromosomes. Subsequently, he was able to narrow down the location of a given gene, in the best cases to within one salivary chromosome band (Bridges 1938). These discoveries by Painter and Bridges constituted the final proof of the chromosome theory of inheritance.

Path to the proof of the DNA theory of inheritance

The first hints that the nucleic acid component of the chromosomes rather than the protein component would constitute the genetic material were received from mutation studies done using the *Sphaerocarpus donnellii* liverwort, certain microbial fungi and maize as experimental organisms. These studies were performed by the German Edgar Knapp, the Americans Alexander Hollaender and Lewis J. Stadler with their coworkers (Knapp and Schreiber 1939; Hollaender and Emmons 1941; Stadler and Uber 1942). They observed that the wavelength of ultraviolet light that gave the maximum frequency of mutations following irradiation

corresponded to the maximum absorption by DNA. None of these groups, however, was yet ready to draw from this result the conclusion that DNA was the genetic material.

Conclusive evidence for the hypothesis that DNA was the carrier of hereditary characters was obtained by the Americans Oswald T. Avery, Collin M. MacLeod and Maclyn McCarty (Avery *et al.* 1944) with *Diplococcus pneumoniae* bacterium (the pneumococcus). In 1928, the British bacteriologist Frederick Griffith had discovered the phenomenon of genetic transformation in pneumococcus (Griffith 1928). His experiments leading to this discovery are well known, being discussed in almost all text books of genetics, and thus they are not described in detail here.

Later, the American Martin H. Dawson and the Chinese Richard H. P. Sia (Dawson and Sia 1931) along with the American J. Lionel Alloway (Alloway 1932) demonstrated transformation *in vitro*, and subsequently Alloway (1933) obtained cell-free extracts that could cause the specific genetic transformation observed by Griffith.

After many years of work, Avery and his associates isolated in a highly purified form the substance responsible for this genetic change, and showed that it was DNA (Avery *et al.* 1944). Subsequently, many other characters were shown to be capable of being transferred by DNA, both in pneumococcus and a number of other species of bacteria.

This discovery by Oswald T. Avery and his coworkers did not, however, convince all members of the scientific community. It was commonly thought that their preparation contained proteins as an impurity; at that time it was generally believed that only proteins could have the specificity required for genetic material. In addition, concerning the structure of DNA, the so-called tetranucleotide hypothesis, first proposed by Hermann Steudel in 1906 and developed by the Lithuanian-American biochemist Phoebus Levene in 1931, prevailed (Olby 1994; Deichmann 2004, quoted by Falk 2009, pp. 191). According to this hypothesis, DNA was composed of identical units of tetranucleotides, which were thought to contain one of the four bases each (Dahm 2008). Such a structure would be too monotonous for a molecule carrying specific information, a characteristic required for genetic material.

On the other hand, true knowledge, too, about the structure of DNA increased during the 1930s, when the English physicist and molecular biologist William T. Astbury proposed that DNA was a long and helical linear molecule. This he determined on the basis of the x-ray crystallographical material (which was provided by Torbjörn Caspersson, a Swedish cytologist and geneticist) in collaboration with Florence O. Bell (Astbury and Bell 1938; Rheinberger 1998: 645).

More convincing for the entire scientific community were the experiments of the American geneticists Alfred D. Hershey and Martha C. Chase (Hershey and Chase 1952), in which they showed that DNA was responsible for bacteriophage multiplication. Hershey and Chase's experiment, described in most text books of genetics, made it clear that the genetic material of the virus is its DNA. This agreed with Avery, MacLeod and McCarty's discovery with pneumococcus, showing that hereditary characters may be carried by DNA alone, since little or no protein appears to be associated with the process of hereditary transmission. These discoveries suggested that in higher forms of life, too, it might be the DNA alone that was the chemical basis of heredity, with the protein part of the chromosome having some other function. This had been suggested in 1948 on the basis of the DNA content of nuclei and the base composition of DNA by the French microbiologist André F. Boivin and his coworkers Roger and Colette Vendrely (Boivin *et al.* 1948; Vendrely and Vendrely 1948). Today, it is fully established that, apart from the important exception of the RNA viruses, DNA is the universal genetic material of life on Earth.

The Austrian biochemist Erwin Chargaff was among the few scientists who understood the significance of the work of Oswald Avery and his coworkers, considered its results true and worked accordingly. In the late 1940s Chargaff and his collaborators showed that the tetranucleotide hypothesis must be wrong, and at the same time he showed the specificity of the structure of DNA (Chargaff 1950, 1951; Chargaff *et al.* 1949). Chargaff discovered an important rule, today termed Chargaff's rule, concerning the proportions of the bases of DNA. It was found that DNA invariably contained equal amounts of adenine (A) and thymine (T) on the one hand and guanine (G) and cytosine (C) on the other, which rule then suggested to Watson and Crick the base-pairing rule in the structure of DNA.

The double-helix model of DNA, and its genetic implications

Genetic implications of the model

Leaning on the Chargaff group's chemical studies, the physicochemical studies described above, and the x-ray crystallographic studies of Wilkins *et al.* (1953) and Franklin and Gosling (1953), James D. Watson and Francis H. C. Crick were finally able to propose the structure for DNA known as the double-helix model (Watson and Crick 1953a).

As pointed out by Watson and Crick (1953b, 1954), their model of the structure of DNA fulfilled the characteristics necessary for the genetic material, that is, autoreplication, specificity, and information content. In addition, the model accounts for the ability of the genetic material to change, i.e., the ability of the genes to mutate.

The autoreplication of DNA is based on the base-pairing rule of its structure, suggested by Chargaff's (1950) work, that adenine is always opposed to thymine, and cytosine is always opposed to guanine. The specificity and information content of DNA is based on its linear structure; the sequence of nucleotides in DNA constitutes the genetic information. The ability of the genes to mutate, which is a necessary condition for biological evolution, is based on the built-in

property of the model that on rare occasions the nucleotides of DNA can be changed. Nevertheless, the structure of DNA is stable enough that organisms can rely on its encoded information during development. These characteristics of DNA are the necessary conditions of the genetic material which life and its evolution is based on. If there is life elsewhere in the universe, and if the life there is based on some other macromolecule than DNA, even that molecule must have these properties.

Difficulties with the model and their solutions

Notwithstanding the elegance of the double-helix model, it posed some serious difficulties at the physico-chemical level. As already mentioned by Watson and Crick (1953a), the unwinding of the helix during replication was a major one. A prominent skeptic in this instance was the German–American biophysicist Max Delbrück (Falk 2009: 196). Specifically, this problem involved long DNA molecules in the chromosomes of the eukaryotic organisms.

The unwinding problem was resolved when Huberman and Riggs (1968) showed that replications initiated at numerous sites along the chromosome, each unraveling, followed by synthesis extending in both directions in the form of a 'replication bubble' that eventually coalesced with its neighbour bubbles.

A number of papers (quoted by Crick *et al.* 1979), presented at the end of the 1970s, suggested that the two strands of DNA do not coil round one another but lie side-by-side. At least some of these models had an even better fit with the x-ray crystallographic data than the double-helix model. Moreover, these side-by-side models involved no difficulties in unwinding, were tape-like rather than rope-like, thus having fewer problems in packing than the double-helix model. Therefore, these models constituted a noteworthy alternative for the Watson–Crick Model. However, Crick *et al.* (1979) were able to disprove the side-by-side models on the basis of the mobility of circular DNA molecules in electrophoresis, and, moreover, they presented a huge amount of evidence existing at that time in favour of the double-helix model.

The highlights of DNA research after the discovery of the structure of DNA

The discovery of the structure of DNA signified the starting point of a new branch of science, viz., molecular genetics, which has flourished since. The history and present status of molecular genetics will be briefly summarized here.

The discovery of the mechanism of DNA replication

In principle, there are three alternative models for the replication of DNA, the semi-conservative, the conservative and the dispersive. In the semi-conservative replication, the double-helix of each daughter DNA molecule contains one strand from the original DNA molecule and one newly synthesized strand. In the conservative replication instead, the parent molecule is conserved, and a single daughter double helix is produced, consisting of two newly synthesized strands. In the dispersive replication, daughter molecules consist of strands, each containing segments of both parental DNA and newly synthesized DNA.

The Americans Matthew Meselson and Franklin W. Stahl succeeded in performing an elegant experiment in 1958 to make a distinction between these alternatives, and observed that the semi-conservative model was the correct one (Meselson and Stahl 1958b). They noted that the models for the mechanisms of DNA replication 'differ in the predictions they make concerning the distribution among progeny molecules of atoms derived from parental molecules.' It was only needed to 'label the DNA of an organism, allow it to reproduce in a nonlabelling medium, and then determine the distribution of parental label among progeny DNA molecules' (Meselson and Stahl 1958a).

In essence, *Escherichia coli* bacteria were fed for several generations with nutrients that contained a heavy isotope of nitrogen (^{15}N). When all the nitrogen of the bacterial DNA was of the heavy isotope, the bacteria were allowed to replicate synchronously once, twice and thrice, in a medium with only light DNA precursors. Samples of DNA from given cycles of replication were run in a cesium chloride gradient in an ultracentrifuge. One generation of replication produced hybrid molecules that gave a band located halfway between that of the heavy and light DNA. Another cycle of replication produced equal amounts of two buoyancies, the hybrid and the light-only buoyancy. Subsequent cycles further diluted the band of the hybrid buoyancy (Meselson and Stahl 1958b; explained in Falk 2009, pp. 197). The results permitted the conclusion that DNA replication in *E. coli* evidently followed the semi-conservative pattern, which, in fact, Watson and Crick (1953b, 1954) had postulated.

A little earlier, the American J. Herbert Taylor with his coworkers had already demonstrated, via the use of an experiment analogous to that of Meselson and Stahl, that the replication of plant chromosomes is semi-conservative (Taylor *et al.* 1957). As explained by Whitehouse (1973: 188), it does not, however, follow from this that the DNA molecule or molecules that form part of the chromosomes necessarily also show this pattern of replication, because at that time there was no evidence to indicate that each chromatid was one strand of complementary DNA strands. In contrast to this, the experiment by Meselson and Stahl (1958b) indicated that the DNA of bacteria replicated as a unit, semi-conservatively, rather than in the dispersive manner, strongly indicated that all the DNA of these cells acts as one continuous molecule (Falk 2009, pp. 198). Therefore, their experiment deserves its high repute as 'the most beautiful experiment in biology' (Judson 1996, pp. 163).

The birth and development of the theory of the genetic code

Earlier, before it was found by Avery *et al.* (1944) that DNA is the genetic material, American geneticists George W. Beadle and Edward L. Tatum as well as Adrian M. Srb and Norman H. Horowitz had demonstrated that genes direct the biosynthesis of proteins in the cells using the *Neurospora crassa* microbial fungus as an experimental organism (Beadle and Tatum 1941; Srb and Horowitz 1944). After this, the American biochemist Alexander L. Dounce and the Russian-American theoretical physicist and cosmologist George Gamow created, independently of each other, a theory according to which the sequence of nucleotides in the DNA of the genes determines the sequence of amino acids in the primary structure of the proteins (Dounce 1952; Gamow 1954). This theory has variously been called the sequence hypothesis, co-linearity hypothesis and the theory of the genetic code. James D. Watson and Francis H. C. Crick adduced the same idea in their papers discussing the genetic implications of the double-helix model of the structure of DNA (Watson and Crick 1953b, 1954). During the first part of the 1960s, several scientists presented evidence for the co-linearity of given genes and the respective proteins by comparing the genetic fine structure maps of the genes and the primary structures of corresponding proteins (see Portin 1993).

In a 1958 theoretical paper concerning the mechanism of protein synthesis, Francis H. C. Crick presented a hypothesis according to which the biosynthesis of proteins is a biphasic process. In the first phase, called genetic transcription, the genetic information residing in the genes is copied to RNA in the nucleus. This is then transported into the cytoplasm, where, in the second phase of protein synthesis, called genetic translation, the genetic information, now carried by the intermediating RNA, is converted into an amino acid sequence (Crick 1958). According to the Central Dogma of Molecular Biology, presented by Crick for the first time in the same paper, genetic information in the cells can flow in one direction only: first from DNA to RNA, and then from RNA to proteins, but never from proteins to nucleic acids. Nor can information be transferred from protein to protein. This fundamental theory is, in its updated form presented by Crick in 1970 (Crick 1970), still in force.

The existence of the intermediating RNA molecule, being of a metabolically unstable nature, and whose function would be to carry information from DNA to protein, was postulated by Riley *et al.* (1960) and Jacob and Monod (1961), as a result of their studies of inducible enzymes in *E. coli*, and the name 'messenger RNA' (mRNA) was coined by Jacob and Monod (1961).

RNA was already being seriously considered as the intermediary between DNA and proteins in the late 1950s, but the ideas had not yet crystallized (see e.g. Portin 1993 for a review). According to an early version of the theory of information transfer from genes to cytoplasm, a gene was imagined to give rise to the formation of one specialized kind of ribosome, which in turn would direct the synthesis of one and only one kind of protein—a scheme that Brenner *et al.* (1961) had epitomized as the 'one gene / one ribosome / one protein' hypothesis. This theory had, however, already been shown to be incorrect in 1959 by Pardee *et al.* (1959) on the basis of their analyses concerning the regulation of the lactose operon of *E. coli*. Rather quickly thereafter, several studies conducted by different scientists during the first years of the 1960s confirmed the hypothesis concerning the role of messenger RNA in protein synthesis beyond doubt (see Portin 1993).

Purely genetic mutation studies conducted again by Francis Crick and his coworkers (Crick *et al.* 1961) had a great impact on the solution of the problem of the genetic code. These studies indicated for the first time that the genetic code, the rule according to which the correspondence of the genes and proteins is determined, is a comma-less triplet code where the code words do not overlap but occur consecutively. 'Triplet code' means that code words, or codons, in DNA consist of groups of three nucleotides, and the word 'comma-less' means that there are no nucleotides between the codons as though punctuation marks.

The biochemical deciphering of the genetic code *in vitro* succeeded soon after this during 1961–1965. In other words, it was discovered which code triplets in DNA corresponded to each amino acid in the primary structure of proteins. In this endeavor, the biochemists Marshall W. Nirenberg, J. Heinrich Matthaei, H. Gobind Khorana and Severo Ochoa, all working at that time in the United States of America, distinguished themselves. The chapters of this great achievement of molecular genetics have been reported in the best textbooks of genetics (e.g. Bresch and Hausmann 1972; Griffiths *et al.* 2008; Janning and Knust 2004; Whitehouse 1973), and they will therefore not be repeated here.

Partly, simultaneously with these stages, the American geneticist Charles Yanofsky and his coworkers confirmed biologically that the code holds true also *in vivo* (Yanofsky 1963; Yanofsky *et al.* 1966). They came to this conclusion after investigating mutations in the gene encoding for tryptophan synthetase in the *E. coli* bacterium. The order of mutations on the fine structure map of the gene was compared with the order of changes in the primary structure of the enzyme protein caused by these mutations. It was observed that these orders corresponded with each other (Yanofsky 1963). Further, it was observed that the changes in the amino acid sequence caused by these mutations could be explained assuming single-nucleotide alterations in the corresponding gene, and by assuming that the genetic code holds true also *in vivo*. Also, the phenotypic effects of intragenic recombination on the amino acid sequence could be explained in the same way by assuming that recombination occurred between two nucleotides and that the genetic code held true *in vivo* (Yanofsky *et al.* 1966).

Further, mutation studies performed by the German biochemists Heinz-Günter Wittmann and Brigitte Wittmann-Liebold using the tobacco mosaic virus (TMV) as the

experimental organism confirmed the genetic code *in vivo*. TMV belongs to the RNA viruses, and mutations in its genome were induced with nitrous acid, the mutagenic effects of which are fully specific. Namely, it causes mutations in RNA by altering cytosine to uracil, and adenine to hypoxanthine, which is then expected to pair like guanine (in RNA, instead of thymine, uracil occurs as one of the bases). The effects of a total of 24 mutations on the amino acid sequence of the coat protein of the virus were studied. Of these, 23 cases could be explained on the basis of single-nucleotide changes mentioned, and assuming that the genetic code held true *in vivo*. The one case left over was interpreted to be a spontaneous mutation (Wittmann and Wittmann-Liebold 1966).

The genetic code was shown to be universal. This means that practically speaking all organisms on earth, be they a virus, bacterium, archae, fungus, plant or animal, use the same code, a fact which is one of the strongest pieces of evidence in favour of the theory of evolution. Only a very few exceptions of this rule exist. For example, in the mitochondrial genome some few code triplets have a different meaning than in the nuclear genome.

The birth and victories of gene technology and biochemical gene analysis

At the turn of the 1960s and 1970s, after the discovery of the bacterial restriction enzymes, and the description of their ability to cleave DNA at specific sites, it became possible to isolate, amplify by cloning, manipulate, and transfer genes artificially, even from one species to another, as well as to sequence them (see Portin 1993). The sequencing of DNA made it possible to analyse the biochemical fine structure of individual genes, and consequently also that of entire genomes.

The first artificial gene transfer from one species to another was carried out by a team led by the American biochemist Paul Berg in 1972 (Jackson *et al.* 1972). They simultaneously transferred genetic material from a virus and a bacterium into the genome of another virus species. With the method invented by Berg's team, it is possible to deliberately construct different recombinant DNA molecules consisting of DNA from different sources. By introducing a desired DNA segment into such a recombinant DNA molecule, and by using specific gene vectors, it is possible to transfer, for instance, human DNA into a bacterial cell and let it amplify there. Alternatively, the polymerase chain reaction (PCR), invented by the American biochemist Kary Mullis in 1983, can be applied in amplification (Bartlett and Stirling 2003). In both ways, a sizeable amount of the DNA in question, sufficient for many different chemical and biochemical analyses and purposes, can be obtained. For instance, by applying restriction mapping, invented by the British molecular biologist Edwin Southern (Southern 1975), a physical map of a given gene can be constructed, or the DNA segment in question can be sequenced.

In the sequencing of DNA, the sequence of nucleotides or what is known as the base sequence is determined. The first methods for the sequencing of DNA were the enzymatic method developed by the British biochemist Frederick Sanger and his coworkers (Sanger *et al.* 1977) and method based on chemical degradation invented by the American molecular geneticist Allan M. Maxam and the American physicist and biochemist Walter Gilbert (Maxam and Gilbert 1977). Later, far quicker and cheaper methods for sequencing of DNA have been developed, most of which, however, have the same basic principle as these first methods: DNA is cleaved into fragments at particular sites, and then the fragments are separated using electrophoresis, or some comparable method, according to their sizes. The most modern sequencing methods allow us to see the progression of sequencing in real time. Of these methods, the most new rely on nanotechnology.

The sequencing of genes and genomes is the most accurate physical gene mapping imaginable. Today, entire genomes of countless numbers of eukaryotic species have been analysed in detail, not to mention the genomes of bacteria and other prokaryotes. This has led to a completely new orientation of research in all areas of biology, biomedical research included.

The discovery of the structure of DNA 60 years ago has been a prerequisite for the triumphant march of molecular genetics described above. The inventions and findings reported here have brought to their composers numerous Nobel prizes. The publication of the entire sequence of the human genome, first as a draft in February, 2001 (International Human Genome Sequencing Consortium 2001; Venter *et al.* 2001), and then in its final form on 24 October 2004 (International Human Genome Sequencing Consortium 2004), can be considered the climax of the area so far. This sequence will form the basis for biomedical research for decades to come.

Notably, one splendid example of the significance of the Human Genome project and the respective projects concerning our nearest relative species—both living and extinct—is the deep understanding that followed regarding the relationships between human populations and their ancestors. The chimpanzee genome sequence and its comparison with the human genome was published in 2005, and that of gorilla in 2012 (The Chimpanzee Sequencing and Analysis Consortium 2005; Scally *et al.* 2012). In 2006, two international teams were able to present an initial sequence of the chromosomal genome of Neanderthal man (Green *et al.* 2006; Noonan *et al.* 2006). In 2008, the distal manual phalanx of a juvenile hominin was excavated at Denisova Cave located in the Altai Mountains in southern Siberia. This finding represented a new extinct hominin species, subsequently called Denisova man. The sequence of the genome of the Denisovans was published in 2010 by Reich *et al.* (2010). Attempts to extract DNA from the remnants of the Flores man (*Homo floresiensis*), a third victim of recent extinction in the genus *Homo*, found in 2003

from the island of Flores in Indonesia, have so far been unsuccessful.

The significance of these results for our understanding of the evolution and polymorphism of our species has been reviewed by, among others, the present author (Portin 2007). I believe that the most important findings in this area, from the point of view of human natural history, are the following: firstly, we share a vast majority of our genes with other primates, about 80% with vertebrate animals in general, circa 60% with all animals, and ∼20% with all living organisms (International Human Genome Sequencing Consortium 2001). Secondly, mankind is, genetically-speaking, very homogenous, so homogenous that the concept of race in human biology can be abandoned (Pääbo 2003). These facts signify that we are brothers and sisters not only to our human companions but also to all living beings.

Another very significant consequence of the extensive sequencing projects is the profound change that followed in our understanding of the genome organization and function, also reviewed recently by the present author (Portin 2009). Thanks specifically to the Encyclopedia of DNA Elements (ENCODE) project, we now know that, contrary to earlier belief, the vast majority of the genome is biochemically functional (The ENCODE Project Consortium 2007, 2012), and consequently such terms as 'junk DNA' and 'selfish DNA' can in actual fact be abandoned. In particular, regulatory elements have been found physically associated with each other and with well-studied protein-coding regions. Overall, the project provides new insights into the organization and regulation of our genes and genome, and is a valuable and extensive resource for functional annotations for biomedical research.

Acknowledgements

My friend, Professor Harri Savilahti, Department of Biology, University of Turku, read the first version of the manuscript and made several suggestions for improvement, for which I express my sincere thanks. Special thanks to Maaria Tringham, M.Sc. and Damon Tringham, M.Phil., for checking the language.

References

Alloway J. L. 1932 The transformation in vitro of R pneumococci into S form of different specific types by the use of filtered pneumococcal extracts. *J. Exp. Med.* **55**, 91–99.

Alloway J. L. 1933 Further observations on the use of pneumococcus extracts in affecting transformation of type in vitro. *J. Exp. Med.* **57**, 265–278.

Altmann R. 1889 Über Nukleinsäuren. *Arch. f. Anatomie u. Physiol., Leipzig Physiol. Abt.* 524–536.

Astbury W. T. and Bell F. O. 1938 X-ray study of thymonucleic acid. *Nature* **141**, 747–748.

Avery O. T., MacLeod C. M. and MacCarty M. 1944 Studies on the chemical nature of the substance inducing transformation of Pneumococcal types. Induction of transformation by a deoxyribonucleic acid fraction isolated from Pneumococcus type III. *J. Exp. Med.* **79**, 137–159.

Bartlett J. M. and Stirling D. 2003 A short history of the polymerase chain reaction. *Methods Mol. Biol.* **226**, 3–6.

Beadle G. W. and Tatum E. L. 1941 Genetic control of biochemical reactions in *Neurospora*. *Proc. Natl. Acad. Sci. USA* **27**, 499–506.

Boivin A., Vendrely R. and Vendrely C. 1948 L'acide désoxyribonucléique du noyau cellulaire, dépositaire des caractéres héréditaires: arguments d'ordre analytique. *C. R. Hebd. Séanc. Acad. Sci., Paris* **226**, 1061–1063.

Boveri T. 1902 Über mehrpolige Mitosen als Mittel zur Analyse des Zellkerns. *Verh. Phys.-Med. Ges. Würzb. N.F.* **35**, 60–90.

Boveri T. 1903 Über die Konstitution der chromatischen Kernsubstanz. *Verh. Deutsch. Zool. Ges. Würzb.* **13**, 10–33.

Boveri T. H. 1904 *Ergebnisse über die Konstitution der chromatischen Substanz des Zellkerns*. Gustav Fischer, Jena.

Brenner S., Jacob F. and Meselson M. 1961 An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* **190**, 576–581.

Bresch C. and Hausmann R. 1972 *Klassische und molekulare Genetik*. Third expanded edition. Springer-Verlag, Berlin, Heidelberg, New York, USA.

Bridges C. B. 1914 Direct proof through non-disjunction that sex-linked genes of Drosophila are borne by the X-chromosome. *Science* **40**, 107–109.

Bridges C. B. 1916 Non-disjunction as proof of the chromosome theory of heredity. *Genetics* **1**, 1–52, 107–163.

Bridges C. B. 1935 Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila Melanogaster*. *J. Hered.* **26**, 60–64.

Bridges C. B. 1938 A revised map of the salivary gland X-chromosome of *Drosophila melanogaster*. *J. Hered.* **29**, 11–13.

Chargaff E. 1950 Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209.

Chargaff E. 1951 Structure and function of nucleic acid as cell constituent. *Fed. Proc.* **10**, 654–659.

Chargaff E., Vischer E., Doniger R., Green C. and Misani F. 1949 The composition of the desoxypentose nucleic acid of thymus and spleen. *J. Biol. Chem.* **177**, 405–416.

Creighton H. B. and McClintock B. 1931 A correlation of cytological and genetical crossing-over in *Zea mays*. *Proc. Natl. Acad. Sci. USA* **17**, 492–497.

Crick F. 1970 Central dogma of molecular biology. *Nature* **227**, 561–563.

Crick F. H. C. 1958 On protein synthesis. *Symp. Soc. Exp. Biol.* **12**, 138–167.

Crick F. H. C., Barnett L., Brenner S. and Watts-Tobin R. J. 1961 General nature of the genetic code for proteins. *Nature* **192**, 1227–1232.

Crick F. H. C., Wang J. C. and Bauer W. R. 1979 Is DNA really a double helix? *J. Mol. Biol.* **129**, 449–461.

Dahm R. 2008 Discovering DNA: Friedrich Miescher and the early years of nucleic acid research. *Hum. Genet.* **122**, 565–581.

Dawson M. H. and Sia R. H. P. 1931 In vitro transformation of pneumococcal types. I. A technique for inducing transformation of pneumococcal types in vitro. *J. Exp. Med.* **54**, 681–700.

Deichmann U. 2004 Early responses to Avery *et al*.'s paper on DNA as hereditary material. *Hist. Stud. Phys. Biol. Sci.* **34**, 207–232.

Dobzhansky Th. 1929 Genetical and cytological proof of translocations involving the third and fourth chromosome in Drosophila melanogaster. *Biol. Zentralbl.* **49**, 408–419.

Dounce A. L. 1952 Duplicating mechanism for peptide chain and nucleic acid synthesis. *Enzymologia* **15**, 503–507.

Falk R. 2009 *Genetic analysis: a history of genetic thinking*. Cambridge University Press, Cambridge, UK.

Franklin R. E. and Gosling R. G. 1953 Molecular structure of nucleic acids. Molecular configuration in sodium thymonucleate. *Nature* **171**, 740–741.

Gamow G. 1954 Possible relation between deoxyribonucleic acid and protein structures. *Nature* **173**, 318.

Green R. E., Krause J., Ptak S. E., Briggs A. W., Ronan M. T., Simons J. F. *et al.* 2006 Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**, 330–336.

Griffith F. 1928 Significance of pneumococcal types. *J. Hyg.* **27**, 113–159.

Griffiths A. J. F., Wessler S. R., Lewontin R. C. and Carroll S. B. 2008 *Introduction to Genetic Analysis*, 9th edition. W. H. Freeman, New York, USA.

Hershey A. D. and Chase M. 1952 Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **36**, 39–56.

Hollaender A. and Emmons C. W. 1941 Wavelength dependence of mutation production in ultraviolet with special emphasis on fungi. *Cold Spring Harb. Symp. Quant. Biol.* **9**, 179–186.

Huberman J. A. and Riggs D. A. 1968 On the mechanism of DNA replication in mammalian chromosomes. *J. Mol. Biol.* **32**, 327–341.

International Human Genome Sequencing Consortium 2001 Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.

International Human Genome Sequencing Consortium 2004 Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945.

Jackson D., Symons R. and Berg P. 1972 Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **69**, 2904–2909.

Jacob F. and Monod J. 1961 Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356.

Janning W. and Knust E. 2004 *Genetik: allgemeine genetik, molekulare genetik, entwicklungsgenetik*. Georg Thieme Verlag, Stuttgart, New York.

Judson H. F. 1996 *The eighth day of creation. Markers of the revolution in biology.* Expanded edition. Cold Spring Harbor Laboratory Press, New York.

Knapp E. and Schreiber H. 1939 Quantitative analyse der mutationsauslösende wirkung monochromatischen UV-lichtes in spermatozoiden von *Sphaerocarpus*. In *Proceedings of the 7th International Congress of Genetics, Edinburgh* (ed. R. C. Punnett), pp. 175–176. Cambridge University Press, Cambridge.

Kossel A. 1913 Beziehungen der Chemie zur Physiologie. In *Die Kultur der Gegenwart, ihre Entwicklung und ihre Ziele: Chemie* (ed. Ev. Meyer), pp. 376–412. Teubner, Leipzig.

Kossel A. and Neumann A. 1893 Über das thymin, ein spaltungsprodukt der nukleinsäure. *Ber. Deutsch. Chem. Ges.* **26**, 2753–2756.

Maxam A. M. and Gilbert W. 1977 A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**, 560–564.

McClung C. E. 1902 The accessory chromosome-sex determinant? *Biol. Bull. (Woods Hole)* **3**, 43–84.

Mendel G. 1866 Versuche über Pflanzenhybriden. *Verh. Naturf. Ver. Brünn* **4**, 3–47.

Meselson M. and Stahl F. W. 1958a The replication of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **23**, 9–12.

Meselson M. and Stahl F. W. 1958b The replication of DNA in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **44**, 671–682.

Miescher F. 1871 Über die chemische zusammensetzung der eiterzellen. *Hoppe-Seyler's Med.-chem. Unters.* **4**, 441–460.

Miescher F. 1874a Das protamin, eine neue organische basis aus den samenfäden des rheinlachses. *Ber. Deutsch. Chem. Ges.* **7**, 376–379.

Miescher F. 1874b Die spermatozoen einiger wilbertiere. Ein Beitrag zur histochemie. *Verh. Naturf. Ges. Basel* **6**, 138–208.

Mirsky A. E. 1968 The discovery of DNA. *Sci. Am.* **218**, 78–88.

Morgan T. H. 1910 Sex limited inheritance in *Drosophila*. *Science* **32**, 120–122.

Morgan T. H. 1911 The application of the conception of pure lines to sex-limited inheritance and to sexual dimorphism. *Am. Nat.* **45**, 65–78.

Morgan T. H. 1919 *The physical basis of heredity*. Yale University Press, New Haven, USA.

Morgan T. H. 1926 *The theory of the gene*. Yale University Press, New Haven, USA.

Morgan T. H., Sturtevant A. H., Muller H. J. and Bridges C. B. 1915 *The mechanism of mendelian heredity*. Henry Holt, New York, USA.

Muller H. J. and Painter T. S. 1929 The cytological expression of changes in gene alignment produced by X-rays in *Drosophila*. *Am. Nat.* **63**, 193–200.

Noonan J. P., Coop G., Kudaravelli S., Smith D., Krause J., Alessi J. *et al.* 2006 Sequencing and analysis of Neanderthal genomic DNA. *Science* **314**, 1113–1118.

Olby R. C. 1994 *The path to the double helix the discovery of DNA*. Dover Publications, New York, USA.

Painter T. S. 1933 A new method for the study of chromosome rearrangements and the plotting of chromosome maps. *Science* **78**, 585–586.

Painter T. S. 1934 A new method for the study of chromosome aberrations and the plotting of chromosome maps in *Drosophila melanogaster*. *Genetics* **19**, 175–188.

Pardee A. B., Jacob F. and Monod J. 1959 The genetic control and cytoplasmic expression of inducibility in the synthesis of β-galactosidase by *E. coli*. *J. Mol. Biol.* **1**, 165–178.

Portin P. 1993 The concept of the gene: Short history and present status. *Q. Rev. Biol.* **68**, 173–223.

Portin P. 2007 Evolution of man in the light of molecular genetics: A review. Part I. Our evolutionary history and genomics. *Hereditas* **144**, 80–95.

Portin P. 2009 The elusive concept of the gene. *Hereditas* **146**, 112–117.

Portugal F. H. and Cohen J. S. 1977 *A century of DNA: a history of the discovery of the structure and function of the genetic substance*. MIT Press, Cambridge, UK.

Pääbo S. 2003 The mosaic that is our genome. *Nature* **421**, 409–412.

Reich D., Green R. E., Kircher M., Krause J., Petterson N., Durand E. Y. *et al.* 2010 Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468**, 1053–1060.

Rheinberger H.-J. 1998 Kurze Gesichte der Molekularbiologie. In *Gesichte der biologie. Theorien, methoden, institutionen, kurzbiographien*, 3., neubearbeitete und erweiterte Auflage (ed. I. Jahn), pp. 642–663. Gustav Fisher, Jena, Germany.

Riley M., Pardee A. B., Jacob F. and Monod J. 1960 On the expression of a structural gene. *J. Mol. Biol.* **2**, 216–225.

Sanger F., Nicklen S. and Coulson A. R. 1977 DNA sequencing with chain-termination inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

Scally A., Dutheil J. Y., Hillier L. W., Jordan G. G., Goodhead I, Herrero J. *et al.* 2012 Insights into hominid evolution from the gorilla genome sequence. *Nature* **483**, 169–175.

Southern E. M. 1975 Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**, 503–517.

Srb A. M. and Horowitz N. H. 1944 The ornithine cycle in *Neurospora* and its genetic control. *J. Biol. Chem.* **154**, 129–139.

Stadler L. J. and Uber F. M. 1942 Genetic effects of ultraviolet radiation in maize. IV. Comparison of monochromatic radiation. *Genetics* **27**, 84–118.

Stern C. 1931 Zytologisch-genetische Untersuchungen als Beweise für die Morgansche Theorie des Faktorenaustauschs. *Biol. Zentralbl.* **51**, 547–587.

Sturtevant A. H. 1913 The linear arrangement of the six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool.* **14**, 43–59.

Sutton W. S. 1903 The chromosomes in heredity. *Biol. Bull. (Woods Hole)* **4**, 231–251.

Taylor J. H., Woods P. S. and Hughes W. L. 1957 The organization and duplication of chromosomes as revealed by autoradiographic studies using tritium-labeled thymidine. *Proc. Natl. Acad. Sci. USA* **43**, 122–128.

The Chimpanzee Sequencing and Analysis Consortium 2005 Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87.

The ENCODE Project Consortium 2007 Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816.

The ENCODE Project Consortium 2012 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74.

Vendrely R. and Vendrely C. 1948 La teneur du noyau cellulaire en acide désoxyribonucléique a travers les organes les individus et les espéces animales. *Experientia* **4**, 434–436.

Venter J. C. and 275 other authors 2001 The sequence of the human genome. *Science* **291**, 1304–1351.

Watson J. D. 1968 *The double helix: a personal account of the discovery of the structure of DNA*. Atheneum, New York, USA.

Watson J. D. and Crick F. H. C. 1953a Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738.

Watson J. D. and Crick F. H. C. 1953b Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964–967.

Watson J. D. and Crick F. H. C. 1954 The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **18**, 123–131.

Whitehouse H. L. K. 1973 *Towards an understanding of the mechanism of heredity*, 3rd edn. Edward Arnold, London, UK.

Wilkins M. H. F., Stokes A. R. and Wilson H. R. 1953 Molecular structure of nucleic acids. Molecular structure of deoxypentose nucleic acids. *Nature* **171**, 738–740.

Wilson E. B. 1905 The chromosomes in relation to the determination of sex in insects. *Science* **22**, 500–502.

Wittmann H. G. and Wittmann-Liebold B. 1966 Protein chemical studies of two RNA viruses and their mutants. *Cold Spring Harb. Symp. Quant. Biol.* **31**, 163–172.

Yanofsky C. 1963 Amino acid replacements associated with mutation and recombination in the *A* gene and their relationship to *in vitro* coding data. *Cold Spring Harb. Symp. Quant. Biol.* **28**, 581–588.

Yanofsky C., Ito J. and Horn V. 1966 Amino acid replacements and the genetic code. *Cold Spring Harb. Symp. Quant. Biol.* **31**, 151–162.

Received 28 August 2013, in revised form 15 October 2013; accepted 18 October 2013
Published on the Web: 15 April 2014