

Evolution from hypothetical hereditary ‘factors’ to synthetic genomes: celebrating Mendel’s birth bicentenary*

H. A. Ranganath and Subhash C. Lakhotia

Genetics, emerging as a branch of biology at the beginning of the 20th century to study biological inheritance, has evolved into an important discipline that encompasses all aspects of structure and function of the hereditary genetic information. Consequently, it has impacted every area of biology leading to the integration of its different fields, and of biology with other disciplines of natural sciences and humanities. Although the seed of genetics was sown in 1865 by Gregor Johann Mendel, it germinated only in 1900 when the significance of his forgotten results was understood.

We pay tribute to Mendel on his birth bicentenary through this narration that briefly recapitulates the remarkably far-reaching consequences of his experiments that have placed genetics not only at the centre stage of biology, but have also deeply influenced researchers in nearly all areas of natural sciences and other disciplines like anthropology, history, judiciary, economics, etc. This has made genetics and its derivative branches truly multidisciplinary.

Mendel, considered as the father of genetics, was born 200 years ago on 20 July 1822. A Catholic Abbott, he conducted experimental studies between 1856 and 1863 on patterns of inheritance of a few characters in the pea plant. Using meticulously designed experiments involving hybridization of plants showing contrasting characters, Mendel examined their transmission in successive generations of the hybrid progeny. Based on the results, he arrived at a few generalizations, now known as Mendel’s laws of inheritance. Mendel’s interdisciplinary knowledge in biology, physics and mathematics, particularly combinatorial mathematics and statistics, was indeed pivotal for the success of his experiments and analysis of the results. Mendel presented his findings in 1865 to the Natural Science Society in Brünn, Austria, and published

his remarkable findings in the paper ‘Versuche über Pflanzen-Hybriden’ (Experiments on plant hybrids) in the Society’s journal, *Verhandlungen des naturforschenden Vereines in Brünn* in 1865.

Mendel’s conceptualization of the behaviour of ‘factors’ that he presumed to ‘control’ a given character was far ahead of the times, since in the 1860s phenomena like mitosis and meiosis were unknown. Consequently, and ironically for Mendel, his meticulous observations and unusually precise visionary interpretations remained unnoticed and unrecognized in his lifetime.

Independent investigations by Carl Correns in Germany, Hugo de Vries in Holland and Erich von Tschermak in Austria facilitated rediscovery and validation of Mendel’s results in 1900, 35 years after the publication of his results. The year 1900, therefore, marks the birth of the science of genetics. Some criticisms and disagreement with the statistical ratios of Mendel’s experiments by Ronald Fischer and others followed the rediscovery, since the data in many cases appeared too good to be true. Notwithstanding the criticisms, cytologists and geneticists soon found parallels between the behaviour of chromosomes during gamete formation and the expected behaviour of Mendel’s factors. Within the next 2–3 years of re-discovery of Mendel’s laws in 1900, Walter S. Sutton, Theodor Boveri and others independently noted these parallels, leading to the chromosomal theory of inheritance, which implied that the Mendelian factors were located on chromosomes and the behaviour of chromosomes during meiosis resulted in the formation of gametes according to the laws of inheritance postulated by Mendel. A few years later in 1909, Wilhelm Johannsen introduced three words to dictionary of biology, namely ‘phenotype’, ‘genotype’, and ‘gene’. Phenotype describes the appearance of an organism, genotype refers to the genetic constitution of an individual as revealed by breeding experiments and gene is used to replace Mendel’s ‘factor’. It may be noted that although the naming of the science of heredity as ‘genetics’ is generally ascribed to W. Bateson (1905), the term ‘genetic’ was first used by the Hungarian sheep-breeder Count Imre (Emmerich) Festetics

in his 1819 paper describing the results of his experiments in sheep-breeding, and published in *Oekonomische Neuigkeiten und Verhandlungen (Economic News and Announcements)*, wherein he made a reference to *Die genetischen Gesetze der Natur* (‘the genetic laws of nature’).

Soon after the demonstration of chromosomal basis of inheritance, T. H. Morgan and his group introduced the fruit fly, *Drosophila melanogaster* as a convenient model for studies in genetics. These studies established the chromosome theory of inheritance on firm footing. Despite the fact that the gene was an abstract entity during much of the first half of the 20th century, statistical ratios of phenotypes among the experimental progeny in diverse organisms were used to uncover new facets of inheritance, like linkage, sex linkage, interaction of genes, etc. which were not encountered by Mendel. It is interesting that Morgan was initially critical of the Mendelian laws, but soon changed his views and became one of the main architects of ‘modern genetics’ with Mendel’s laws being the basis.

Thanks largely to the studies in Morgan’s ‘Fly-room’ at Columbia University, USA, beginning in the decade of 1910, the abstract genetic factors postulated by Mendel were finally mapped on chromosomes in cells. This association of genes with chromosomes was more definitively established when the ‘giant’ chromosomes seen in the salivary glands of larvae of *Drosophila* and some other dipteran insects were correctly interpreted as ‘polytene chromosomes’ in 1934 by T. S. Painter, a member of Morgan’s laboratory. Realizing the importance of these chromosomes for studies on the gene, the Editor of the *Journal of Heredity* introduced Painter’s classic 1934 paper as follows: ‘Ever since the gene hypothesis was generally accepted, geneticists and cytologists have dreamed of the day when it would be possible to see the actual genes, instead of having to be satisfied with studying their “shadows”, which were “reflected” in the morphological development of generations of organisms. Giant salivary gland chromosomes of the fruit fly (*Drosophila melanogaster*), whose study is the subject of the accompanying article, are proving to be a new genetic tool of the

*We dedicate this article to the memory of late Prof. S. P. Raychaudhuri (University of Calcutta, Kolkata and Banaras Hindu University, Varanasi) and late Prof. M. R. Rajasekarasetty (University of Mysore, Mysuru) who initiated teaching and research in Genetics in India around the middle of the last century.

utmost importance'. This prophecy indeed came true in subsequent decades. It is not surprising that during the heydays of classical genetics, investigations using *Drosophila*, the Cinderella of genetics, uncovered many of the basics of inheritance.

During the formative years, genetics forged linkage with several other disciplines. Within a year of the rediscovery of Mendel's laws, a connection between genetics and medicine was established in 1901 by the physician, Archibald Garrod, who described some human diseases to be due to 'inborn errors of metabolism'. A few years later in 1914, T. H. Boveri suggested a connection between chromosome instability and cancer.

Studies during the 1930s suggested a relationship between genes and biochemical reactions in metabolic pathways, leading to the 'one gene one enzyme' hypothesis of G. W. Beadle and E. L. Tatum. Some years later, C. H. Waddington, an embryologist and evolutionary biologist, linked genetics and embryology in 1940s, and this laid the foundation of developmental biology. Waddington also introduced the concept of 'epigenetics' while explaining how the single-celled zygote produces different tissues with diversified structures and functions. Epigenetics, literally meaning 'beyond genetics', has now become a major discipline in contemporary biology since epigenetic inheritance relies on the diverse non-genetic modifications of bases in the DNA and of specific amino acids of chromatin-associated histones.

Julian Huxley connected Charles Darwin's theory of evolution and Mendel's ideas on heredity in a mathematical framework to develop the modern synthetic theory of evolution. This had its roots in and was further developed by the studies of mathematicians and biologists, including G. H. Hardy, W. Weinberg, T. K. Pearson, Th. Dobzhansky, J. B. S. Haldane, R. A. Fisher, Sewall Wright, G. L. Stebbins and Ernst Mayr. Power of the synthetic theory of evolution is aptly reflected in the famous statement of Dobzhansky, 'nothing in biology makes sense except in the light of evolution'.

In the 1940s, Barbara McClintock, a maize geneticist, encountered some genes in maize that were unstable with reference to their function and position in the genetic map. As had happened with Mendel's laws, acceptance of McClintock's suggestion of some genes being mobile had to wait for nearly two decades till microbial geneticists discovered and characterized trans-

posable elements in bacteria in the 1960s. The mobile genetic elements or transposons are now realized as significant components of the genetic systems of all organisms. Studies on transposon biology have contributed significantly to our understanding of biological evolution besides catalysing the generation of genetically modified transgenic organisms.

The period of 1940s and 1950s attracted researchers with expertise in diverse fields like microbiology, physics, chemistry, biochemistry, crystallography, etc. to study the gene. Following Frederick Griffith's discovery of 'transforming principle' in bacteria in 1929, Oswald Avery, Colin MacLeod, Maclyn McCarty, Alfred Hershey and Martha Chase conclusively established DNA as the hereditary material in the 1940s. Thus the gene, till then an abstract entity, was identified as a distinct chemical substance, i.e. DNA. James Watson and Francis Crick, taking cognizance of Erwin Chargaff's observation of equality of purines and pyrimidines in DNA, and the X-ray diffraction patterns of DNA observed by Rosalind Franklin and Maurice Wilkins, proposed the famous double-helical structure of DNA. Quickly following this, the astrophysicist George Gamow proposed the concept of genetic code to understand the language of inherited genetic information. Rapid and exciting progress between 1950 and 1970 in molecular biology by researchers with training and expertise in diverse branches of natural sciences unravelled the basics of vital transactions of the gene, including how the DNA copies itself and how the hereditary information coded in its base sequence is transferred to the intermediary ribonucleic acid (RNA) for translation into a defined sequence of amino acids in the protein. The molecular bases of mutation, recombination and repair of DNA, regulation of gene expression, etc. were also beginning to be understood as transactions of nucleic acids. Development of biophysical techniques helped understand molecular structures of various proteins and sub-cellular organelles, some examples being the deciphering of collagen structure by G. N. Ramachandran and of ribosomes by V. Ramakrishnan and others.

Virologists and microbiologists contributed significantly to the birth and progress of molecular biology, and applications of molecular biological approaches transformed these fields. Physicists like Max Delbrück, Seymour Benzer and others, using the characteristic reductionist approach, exploited 'simplicity' of viruses infecting

bacteria (bacteriophages) to expand the scope of genetics and establish the unity of life processes across living beings. The first mechanistic details of regulation of gene expression learnt in bacteria by F. Jacob and J. Monod in the early 1960s, paved the way for understanding the greater complexities of gene regulation in eukaryotes. Questions were initially raised if the lessons learnt from microorganisms would be applicable to higher organisms. Subsequent studies indeed largely confirmed the prophetic statement of Jacques Monod: 'What is true for *E. coli* is true for an elephant', since the basic principles of gene activity turned out to be generally similar in bacteria and eukaryotes. With increasingly deeper studies and better understanding, the gene is now variously described through context-dependent terms like cistron, protein-coding gene, non-coding gene, regulatory gene, split gene, overlapping gene, nested gene, poly-protein gene, processed gene, pseudo-gene, transposable gene (element) and so on.

Parallel to the astounding progress in molecular genetics/biology, the cytology of chromosomes and chromatin, the physical bearers of genes, was also transforming into newer fields like cell biology, molecular cytogenetics, chromatin biology, etc. Herbert Taylor's demonstration of semi-conservative replication of chromosomes in the late 1950s allowed the possibility of directly visualizing gene activity, especially using polytene chromosomes. Studies by W. Beer-mann's group, F. M. Ritossa, A. S. Mukherjee, M. Ashburner and others in the 1960s established a direct correlation between opening and closing of the chromatin with activity and inactivity of genes respectively. The need to examine chromatin and chromosomes at increasingly better spatial resolutions prompted the development of advanced microscope systems like electron microscope, various tunnelling microscopes, confocal microscope, etc. Development of such tools obviously required close interactions between biologists and experts in diverse disciplines ranging from engineering/technology, physics, chemistry, materials science and computer science. Visualization of 'beads on string' (nucleosomes) by Donald Olins and Ada Olins in the early 1970s through electron microscopy transformed our understanding of how the enormous length of DNA in cells is physically packaged within the small nuclear volume. The dynamic nature of nucleosomes and its impact on gene expression are now being understood in the light of the epigenetic changes suggested by Waddington in the

1940s. Heterochromatin-associated repetitive DNA and euchromatin-associated non-coding DNA sequences discovered since the 1960s were once considered as 'junk' or 'selfish', but are now the major players in genetic and epigenetic regulation.

The strong foundations of molecular biology/genetics laid by many inter-disciplinary researchers catalysed unprecedented expansion in basic and applied research, especially after 1970. Success with the earlier discovered DNA ligase in recombining separate DNA molecules in a test tube by Paul Berg, and the use of bacterial restriction enzymes by Stanley Cohen and Herbert Boyer heralded the recombinant DNA technology. Development and large-scale exploitation of biophysical techniques like DNA electrophoresis using agarose gel by Joseph Sambrook, Southern hybridization by Edwin Southern, Frederick Sanger's enzymatic chain termination method for DNA sequencing, Walter Gilbert and Allan Maxam's chemical method for DNA sequencing, Marvin Caruthers and Leroy Hood's invention of automated DNA sequencing, and development of polymerase chain reaction (PCR) using the thermophilic bacterial DNA polymerase by Kary Mullis' group made it easier for numerous laboratories spread across the globe to address deeper aspects of gene structure and function, and also exploit the new knowledge and methods for biotechnological and other applied outputs. Some examples of the biotechnological applications emerging from basic studies on the gene include production of important therapeutic biomolecules using bacteria or other cell systems as 'factories', and improving the quality and quantity of agricultural produce.

Molecular genetics had a pervasive influence on diverse fields in biology. Developments in molecular biology/genetics greatly benefitted medical genetics, initiated by Garrod in 1901, by not only permitting precise identification of the genetic changes and factors underlying inherited disorders, cancer and other diseases, but also facilitated development of precise diagnostic tools and machines using genetic information. The rapid progress in human genetics is making it feasible to diagnose and predict possibilities of genetic diseases based on the genomic sequence data of an individual, and thus provide personalized medical treatment.

Molecular genetic studies initiated with a bacteriophage by Seymour Benzer, originally trained in physics, unravelled the fine structure of the gene in the 1950s. A few years later, Benzer switched to *Drosophila*

and laid the foundations of chronobiology, behavioural genetics and neurogenetics, each of which has grown into independent disciplines with wide-ranging impact in basic as well as clinical studies.

Molecular biology also transformed taxonomy, an important branch in classical biology reflecting the evolutionary divergence of organisms or their phylogeny. Gel electrophoresis to separate proteins led to the discovery of enzyme polymorphism in the 1970s by R. C. Lewontin and others. Subsequent studies on enzyme polymorphism and more recently on DNA sequence variations have strongly linked taxonomy with DNA science in a new branch, namely molecular taxonomy. DNA barcoding has now become popular in taxonomy as well as biomedical industry. DNA-fingerprinting, initiated by Alec Jeffreys and later popularized by Lalji Singh in India, has found significant applications in forensic science and judiciary. Analysis of DNA sequences in fossils of diverse organisms has impacted palaeontology and earth sciences. Extension of these methods to human fossils and living populations is improving the understanding of our own origins and providing new insights in socio-biology.

Morgan's original interest in understanding the mechanism of early embryonic development led him to start using *Drosophila melanogaster* as the model organism. During the 1980s, almost 70 years after Morgan's early studies, with combinations of techniques used in genetics, developmental biology and recombinant DNA-based molecular biology began to reveal the genetic blueprint that shapes *Drosophila* development. Lessons learnt from the fruit fly development were soon applied to diverse multicellular organisms leading to the discovery of common 'homeotic' and other 'developmental' genes, which explain how the genetic information inherited from parents provides the temporally and spatially regulated 'instructions' in different cells of an early embryo to produce the organisms that we see as the end result. This knowledge is now being exploited through inter-disciplinary researches to generate synthetic organs and even organisms.

Birth of the Human Genome Organization (HUGO) and the launching of Human Genome and other genome projects depended on remarkable cooperation among geneticists, molecular biologists, computer (hardware and software) experts, materials scientists, engineers and industry. Rapid advances in the new field of genomics already includes several sub-divisions like

structural genomics, functional genomics, comparative genomics, evolutionary genomics, etc. While publishing the genome sequence of Mendel's pea plant, the editorial in *Nature Genetics* (<https://www.nature.com/articles/s41588-019-0501-0>) stated: 'The genome of the model genetic organism *Pisum sativum*, or pea plant, links nineteenth-century genetics to twenty-first-century genomics, serving as a symbol of how far the genetics field has developed and how greatly technologies have advanced. Almost every student's introduction to genetics currently involves learning Mendel's laws; we envision that genomics and genome sequencing will become just as foundational in the education of future geneticists'. Continuing further, the editorial concludes 'we are excited to publish the genome sequence of Mendel's pea plant. Although the individual genes and sequences of Mendel's seven original traits have been known for a while, we believe that the genome sequencing of the pea plant represents a symbolic milestone for genetics, bringing the foundational experimental studies in basic models into the modern sequencing era. We hope that Gregor Mendel would approve.'

Following the success of genomics, other 'omics' sciences like RNomics, proteomics, metabolomics, microbiomics, etc. have also come into being and are contributing to solving the ever-deepening mystery of the gene and its actions. Such 'big data sciences', requiring powerful computational and statistical tools, have been possible because of large-scale international collaborations, including in virtual mode. The expanding need for DNA-related data to be shared across the world in real time has been one of the catalytic factors in the exponential spread of internet since the 1990s. The powerful new discipline of bioinformatics, which interprets the biological implications of the enormous digital data, requires active collaboration among computer scientists, mathematicians, molecular biologists, geneticists etc. The emerging field of artificial intelligence (AI) has much to learn from neurobiology and neurogenetics.

Despite the misdirected efforts of Lyenko in Russia, genetics has played a significant role in the development of agriculture, for example, the Green Revolution in India in the 1960s which ensured food security. The more recent development of genetically modified organisms using biotechnological approaches is a direct outcome of synergy among genetics, biotechnology and the various 'omics'.

Genetic analysis of the biological ‘warfare’ between viruses and bacteria led to the discovery of restriction endonucleases in the 1960s, which catalysed the recombinant DNA-based biotechnology. Later other ‘nucleases’ like meganucleases, zinc-finger nucleases and transcription activator-like effector-based nucleases (TALEN) were engineered for editing DNA sequences and genomes. More recently, the CRISPR/Cas system and its variants have emerged as a powerful technology to edit the genome, even *in vivo* for therapeutic applications to address inherited disorders. These methods have also greatly impacted commercial activities through novel gene-based industries. With such powerful tools in hand, the new field of ‘synthetic genomics’ has emerged to manipulate and engineer the genome of an organism so that products and processes of value can be obtained on a large scale. Genesis of the modern-day high-efficiency genome engineering tools can be traced back to 1965 when Robert Holley’s group reported the 77 nucleotides long sequence of yeast alanyl tRNA, which was utilized by H. G. Khorana’s team for manual synthesis of the gene for this tRNA. The first protein-coding gene to be synthesized *in vitro* was the 514 bp long human leukocyte interferon gene. A landmark achievement reported in 2013 was the editing of the 4.6 Mb long *Escherichia coli* genome by replacing all UAG stop codons with synonymous UAA codons. Taking the challenge to produce a complete genetic system through chemical synthesis, starting with the digitized DNA sequence, Craig Venter’s team synthesized the complete genome of a bacterial species, *Mycoplasma*

genitalium in 2008. In 2010, this group produced the first novel biological species, with a computer as its parents, following synthesis of the entire genome of *M. mycoides* and transplanting in a recipient *Mycoplasma capricolum* cell.

Synthesis of more complex eukaryotic chromosomes also progressed in parallel. Following the successful synthesis of chromosome 25 (497 kb) and 26 (441 kb) of the diatom *Phaeodactylum tricorutum*, which has 33 chromosomes, synthesis of fully functional chromosome III of baker’s yeast (*Saccharomyces cerevisiae*) was achieved in 2014. Synthetic genomics and synthetic biology have now become active approaches to produce organisms with industrially high-value phenotypes.

Application of genetics and its diverse ramifications also made it possible to develop various vaccines, including the recent ones for the COVID-19 pandemic. These have indeed been saviours of humans.

A darker side of the great success stories that make it possible to manipulate genes, genomes and organisms, is the unfortunate selfish human desire to control and dominate over others. Misuse of eugenics in the past serves as a grim reminder of what can happen when powerful techniques are misdirected.

Concluding remarks

Mendel’s ‘factor’ (gene) was an imaginative entity. Understanding the biology and chemistry of nucleic acids not only enabled us to ‘see’ the gene, but to also ‘dissect’ and ‘direct’ it. Isolation and manipulation

of naturally occurring genes, synthesis of novel genes and genomes are no longer dreams. Mendel would have certainly marvelled at the blossoming of the seed that he had sown through the simple breeding experiments into an enormously intricate tree of knowledge of genetics that besides occupying the centre stage in biology has become the meeting ground for researchers with expertise in disciplines other than biology. This has indeed been possible through curiosity-driven basic and applied research, with attention to unusual findings and sometimes doing ‘crazy’ experiments, and above all through deep interdisciplinary interactions and collaborations. We may expand Dobzhansky’s famous statement ‘nothing in biology makes sense except in the light of evolution’, by adding ‘and nothing in evolution makes sense except in the light of genes’.

ACKNOWLEDGEMENT. We thank Prof. Rajiva Raman (Banaras Hindu University, Varanasi) and Prof. Pradeep Burma (University of Delhi, New Delhi) for their useful suggestions and comments on the manuscript.

H. A. Ranganath is in the Centre for Human Genetics, Biotech Park, Electronic City, Phase 1, Bengaluru 560 100, India; Department of Zoology, University of Mysore, Manasagangotri, Mysuru 570 006, India; Subhash C. Lakhotia is in the Cytogenetics Laboratory, Department of Zoology, Banaras Hindu University, Varanasi 221 005, India.*

**e-mail: lakhotia@bhu.ac.in*