

Advances in genetics and plant breeding 160 years after the publication of Mendel's research

Magno Antonio Patto Ramalho¹, Flavia Maria Avelar Gonçalves¹, Vânia Helena Techio¹ and Lucimara Cruz de Souza^{1*}

Crop Breeding and Applied Biotechnology
25(2): e521725212, 2025
Brazilian Society of Plant Breeding.
Printed in Brazil
<http://dx.doi.org/10.1590/1984-70332025v25n2a27>

Abstract: Mendel's research marked the start of a young science: Genetics. Among its various branches, genetic plant breeding merges ancestral domestication practices with scientific advances. This article briefly reports the main contributions of scientists who have marked history with countless contributions over the last 160 years, shedding light on what a gene is and how it works. Thus, it is possible to understand how information is passed between cells of the same individual and from parents to their offspring, as well as how the variability of different traits of individuals occurs. The advances were unimaginable, but it was also found that the production of the phenotype is much more complex than one could imagine at first. As in the past, improving the plant selection process remains the main strategy to continue producing new cultivars to meet humanity's needs regarding grains, fruits, fibers, and renewable energy.

Keywords: Genome, biotechnology, molecular genetics, quantitative genetics

INTRODUCTION

The similarity between related individuals, both in plants and animals, probably caught the attention of our ancestors thousands of years ago. However, it was only after the publication of the results of Mendel's research in 1865, 160 years ago, that the process of understanding what is now called heredity began, that is, the passing of information from parents to their descendants.

The Augustinian monk Gregor Johann Mendel (1822–1884) carried out all of his research in the gardens of his monastery, located in a city that was then part of Austria as is now Brno, in the Czech Republic. He hybridized several plant species, but focused his attention on one vegetable, the pea. Initially, he had 54 cultivars from Germany, of which he selected 42 for hybridization (Van Dijk et al. 2022). His research was carried out from 1854 to 1865 until its publication. He studied seven traits: two of the cotyledons and the rest of the plant. Everything indicates that, unlike his predecessors when explaining heredity, he noted the number of all phenotypes obtained in the offspring of the F₁ and F₂ generations. From this accumulated data, he established what is now recognized as Mendel's Laws.

The most surprising thing is that Mendel formulated all his assumptions without referencing what we know to be fundamental to biologically explain what occurred. For example, mitosis was only discovered in 1878 by the German biologist Walter Flemming, and meiosis, in 1885 by Friedrich Weissmann. Thus,



***Corresponding author:**

E-mail: lucimaracruz@ufla.br

 ORCID: 0000-0001-5795-7960

Received: 03 February 2025

Accepted: 09 April 2025

Published: 02 June 2025

¹ Universidade Federal de Lavras, Trevo Rotatório Professor Edmir Sá Santos, Universidade Federal, 37203-202, Lavras - MG, Brazil

he had no idea how gametes were produced, much less how fertilization was used to produce seeds. He also had no information that the expression of the characteristics of the cotyledons occurred at different times than those of the plant. In other words, he had no idea about the Xenia phenomenon, which Focke explained only a few years later, in 1881.

Over time, some scientists criticized the veracity of the phenotypic segregation ratios obtained by Mendel. One of them was Fisher in an article published in 1936. However, there is no reason to doubt the values reported by Mendel almost 60 years after their publication, especially considering that, at that time, the knowledge and applications of Mendel's Law were already widely consolidated. Even assuming that Fisher's argument was correct and that Mendel had adjusted the phenotypic ratios observed in the F_2 generations, he would have had to propose an a priori hypothesis that could explain the 3:1 segregation. Evidently, proposing segregation, which would have been obtained at that time to explain the genetic control of characteristics, would only have been possible if he had a much more privileged mind than we imagine today.

Thus, Mendel sought to explain what had occurred based on his results. He discovered that the proportions for the phenotypic segregations of the different traits were similar, and based on this, he established a hypothesis of what could have occurred. This hypothesis was later proven, giving rise to what came to be called Genetics. This science has accumulated a wealth of information in a short space of time, with repercussions in various areas of human society. The objective of this publication is to present a summary of what has occurred in the last 160 years in plant breeding with the evolution of knowledge in Genetics. Additionally, to comment on how this knowledge has contributed to meeting all the needs of grains, fruits, fibers, and energy of a population on planet Earth that has increased far beyond what could be imagined.

THE BEGINNING OF GENETICS AND PLANT BREEDING AS A SCIENCE

The domestication of plants was the most important event in human civilization when it ceased to be nomadic and began to live in communities. The process occurred in several very diverse regions of the planet and probably in the last 10 thousand years. It is estimated that, over this time, approximately 3,000 species were domesticated, although a small number of them received greater attention. It is important to mention that this domestication was conducted based exclusively on the visual capacity of our ancestors, who selected only a few individuals with the desirable phenotypes for the characteristics of interest to participate in the next planting.

Domestication continued, but since the 20th century, it has been called Genetic Breeding. In this context, there was a combination of the art practiced previously and the emerging science driven by the rediscovery of Mendel's Laws. Surprisingly, this rediscovery was made independently by three researchers of different nationalities: the German Carl Correns, the Dutch Hugo De Vries and the Austrian Erich von Tschermak. Their findings were published in three articles in 1900.

However, there was initially no unanimity in accepting Mendel's proposal. Among the defenders of his ideas was the English zoologist William Bateson, who played a fundamental role in disseminating Mendel's work. Bateson was the first to translate Mendel's article into English, publishing the book *Mendel's Principles of Heredity: A Defence*. In addition, he coined the term "Genetics" to name this new science and introduced fundamental concepts such as homozygote, heterozygote, alleles, and epistasis. Among the many opponents was Karl Pearson, a brilliant mathematician and disciple of Galton, who used biometrics to estimate the association between the phenotypes of parents and their offspring. They imagined that, since the traits studied by Mendel had a discontinuous distribution, what he proposed would not apply to those with a continuous distribution, which was the focus of his work. The divergence was gradually resolved, mainly with the research of East and Nilsson Elhe, who showed that the genetic control of qualitative and quantitative traits differs basically in the number of genes involved in the process.

Special mention should be made of the Danish scientist Wilhelm Johannsen (1857–1927), who, in 1909, published a book with relevant information on genetics. He coined the term "gene," replacing the "factors" or "elements" used by Mendel, and commented that the phenotype depends on the genotype and the environment. Besides, he encouraged the use of biometrics in the study of genetics. However, according to Rédei (2002), Johannsen was emphasizing that "We must pursue genetics with mathematics, not as mathematics."

This period of controversy in the new science was full of achievements still used today. One of them was the concept of populations in equilibrium, which, although there are still disagreements about the origin of the information, was independently attributed to a British mathematician, Hardy, and a German physician, Weinberg, and is therefore known as Hardy-Weinberg equilibrium.

Regarding genetic breeding, the phenomenon of hybrid vigor, observed in experiments with corn carried out from 1907 onwards, is also worth mentioning. Its discovery and the name for the phenomenon of heterosis are attributed to George H. Shull. Shull (1948) published an article "What is heterosis?", providing details of how it all happened and the reasons for choosing the name. However, Berlan (2021) cast doubt on whether it is correct to attribute the discovery above to Shull and presented the reasons for this. Leaving aside the controversy surrounding the origin of the name, heterosis is considered the superiority of the F_1 generation in relation to the average of the parents. Knowledge of heterosis made the commercial production of hybrid corn possible in the United States in the second decade of the 20th century. This is a striking example of how a theoretical concept in genetics had enormous practical repercussions. Its use was fundamental to the beginning of the entire seed industry, with enormous economic and social repercussions.

THE EMERGENCE AND CONTRIBUTIONS OF QUANTITATIVE GENETICS

Quantitative genetics is the part of genetics involved in studying traits with continuous distribution. In other words, those controlled by many genes and/or influenced by the environment. Most economically important traits in plants fall into this category. Although there are reports of research on quantitative traits carried out by other biometricians, Sir Ronald Fisher (1890–1962) must be credited as the main person responsible for generating information that has guided practically all knowledge of Quantitative Genetics over time. He was an excellent mathematician associated with an enormous practical vision. Most of his contributions occurred after he was 29 years old, when he started working at the Rothamsted Agricultural Experiment Station. He made contributions in several areas, one of which was Quantitative Genetics (QG) after publishing *The correlation between relatives on the assumption of Mendelian inheritance*, in 1918. He assumed that QG results from the effect of many genes acting cumulatively. Based on this condition, he established a model that considered the population of a diploid species in Hardy-Weinberg equilibrium, in which the effects of each locus are added together. In summary, considering this article and others published later, the genetic variance (σ_G^2) of populations can be decomposed into three parts: a) an additive part (σ_A^2), which is the part associated with the average effect of the alleles of each locus; b) dominance (σ_D^2), attributed to allelic interactions; and c) epistatic effects (σ_I^2), the part due to non-allelic or genic interactions. Thus, it can be inferred that: $\sigma_G^2 = \sigma_A^2 + \sigma_D^2 + \sigma_I^2$.

Notably, the decomposition of genetic variance in the presence of inbreeding occurred many years after Fisher's work mentioned above. This decomposition began in 1964 with Harris. Later, in 1983, Cockerham and Weir expanded the model to species with mixed reproduction, that is, with a certain rate of autogamy and the rest of allogamy. Souza Júnior (1989) demonstrated that, when considering an inbreeding rate F in the population, genetic variance can be analyzed without considering epistasis and can be broken down.

$$\sigma_G^2 = (1 + F)\sigma_A^2 + (1 - F)\sigma_D^2 + 4FD1 + FD2 + F(1 - F)\bar{H}$$

With inbreeding, three new components occur that were not included by Fischer. That is: $D1$, which is the genetic covariance between the average (additive) effects of the alleles and the dominance effects of homozygotes; $D2$, which is the genetic variance of the dominance effects of homozygotes; and $\bar{H}$, which is inbreeding depression squared. Note that without inbreeding ($F=0$), these three components will not occur. Considering each locus individually, Souza Júnior (1989) demonstrated that they also depend on the allele frequency; if it is equal to half or in the absence of allelic dominance interaction, they will be null.

Another area of Fisher's work was associated with agricultural experimentation. Evidently, it is also of great importance in QG. As mentioned, he worked at an agricultural experimental station where he learned the difficulty of comparing the evaluated treatments. This experience resulted in important contributions to hypothesis testing and the establishment of the basic principles of experimentation, that is, the existence of repetition, randomization, and local control. It is not without merit that Rothamsted Station is considered the place where **science in agriculture began**. According to some individuals who knew him personally, he was a fanatic for conducting experiments with maximum accuracy, and always commented, "Numbers first, mathematics later."

A few years later, in 1937, the American Jay L. Lush (1896–1982) published the book *Animal Breeding Plans*. This publication was very useful because it introduced the concepts of heritability (h^2) and expected gain from selection (SG). It showed that $SG = dsh^2$, where ds is the differential selection, that is, the difference between the means of the selected individuals and that of the population, and h^2 is the proportion of phenotypic variance transmitted to the offspring. Thus, heritability depends on the selective unit and the propagation mode of the selected individuals.

Another scientist, also an animal breeder, whose work impacted other areas of knowledge, was Charles Roy Henderson (1911–1989). Among his contributions is the mixed model methodology proposed in 1949. Due to the greater availability of computational facilities, it only became widely used in the last two decades of the 20th century.

Two large groups of researchers made significant advances in QG between 1950 and 1980. The first in England, led by K. Mather, researched *Drosophila* and some autogamous plants. The second in the U.S., at a few universities, especially Iowa State University and North Carolina State College, obtained information mainly from corn crops. One focus of the research was to estimate the variance components to verify whether the dominance variance was significant, aiming to guide plant breeding programs.

Estimates obtained in several situations showed that the additive variance predominated in autogamous plants, but dominance also occurred. This served as a stimulus for the commercial production of hybrids of autogamous species a few years later. In this context, among the numerous contributions of American researchers, three genetic-statistical designs used to obtain estimates of variance components, published in 1948 by Comstock and Robinson (1948), stand out. Several articles were published using these models in particular. Hallauer et al. (2010) compiled many estimates in the literature for various traits in corn; they found that the dominance variance was significant for grain yield but not greater than the additive variance. This information helped them demonstrate that overdominance does not occur; that is, the loci in heterozygotes in isolation do not perform better than both homozygotes. This fact allowed them to infer that heterosis, although its probable causes are not yet definitively known, should not be attributed to the occurrence of overdominance as previously imagined. This information, at least in principle, encouraged the implementation of recurrent selection programs. This is because increasing the average of plant populations requires raising the frequency of favorable alleles.

During the mentioned period, emphasis was also placed on obtaining information about the general and specific combining capacity of lines or populations. For this purpose, the most common procedure involved diallel crosses. Much theoretical information has been obtained, including from Brazil, which continues to be practically useful for many cultivated species. The same has occurred with other lines of research, such as the simultaneous selection of multiple traits.

With the availability of computational resources, comparison between selection methods and strategies has received great emphasis, whether through resampling or simulation based on pre-established genetic-statistical models. Although there are limitations in the models used, the inferences have been useful because, based on results obtained in the field, it is impossible to include all variables imagined by breeders.

The interaction of genotypes by environments (GE) has been extensively researched. GE is crucial in all situations, but its effects are particularly pronounced in tropical conditions. Over the years, a vast amount of research has been conducted to quantify the impact of GE. Subsequently, methodologies have been proposed to identify more adapted and stable cultivars. More recently, with the advent of new data analysis tools, greater attention has turned to what is referred to as ambientomics. In principle, the goal is to utilize prediction models that simultaneously incorporate climate, management, and genome data to aid in recommending new cultivars. The challenge is immense, and the expectations for what these models will deliver are enormous.

Since the first molecular markers emerged, there have been enormous expectations of their use in plant breeding. Since then, much of the effort of geneticists involved in QG has been directed towards this line of research. More details on the subject will be discussed later.

ADVANCES IN KNOWLEDGE OF CYTOGENETICS

The rediscovery of Mendel's Laws stirred the scientific community around new hypotheses and theories, and many scientists began to make important contributions that popularized Mendelian ideas. Walter Sutton and Theodor Boveri

carried out one of the most important works in 1902–1903. Both independently proposed the basis for what was later called the Chromosomal Theory of Inheritance (CHT). The CHT related the principles of Mendelian Laws to the behavior of chromosomes in cell divisions and considered that the factors designated by Mendel (elements responsible for heredity) were physical entities located in the chromosomes.

Although it represented an advance in understanding of Mendelian principles, the CHT also encountered important critics such as Wilhelm Johansen in 1909 and, initially, Thomas Hunt Morgan. Morgan gained recognition for his studies in embryology and for being skeptical about the ideas of Mendel and Darwin. Attracted by the genetic variants he observed in *Drosophila*, Morgan gathered several students and researchers in his laboratory. In a peculiar environment where the smell of bananas that fed the flies was dispersed amid piles of scientific articles and glassware, many discussions and stories took place. The most curious situation was discovering a white-eyed male among the predominantly red-eyed flies. However, the white-eyed mutant escaped when Morgan and his wife, Lilian, showed the students how the flies flew towards them. Luckily, the next day, it was recaptured, allowing the history and future of genetics to change (Rédei 2002).

So much so that, based on the results obtained from the study of these mutant flies, Morgan published the first relevant article titled *Sex-limited inheritance*. He documented many basic principles of genetics with other scientists, discovering that factors (now known as genes) form linkage groups that exhibit the same inheritance pattern as chromosomes. Conceptually, alleles could then be considered alternative forms of the gene located at the same locus on homologous chromosomes.

Experimental data mapping sex chromosomes in *Drosophila* mutants supported the hypothesis that genes are independent physical entities arranged in a linear fashion on chromosomes, adhering to Mendel's Law of Independent Assortment. From this, Morgan and his colleagues theorized that: "Although Mendel's Law does not explain the phenomena of development (quantitative trait) and does not claim to explain them, it stands out as a scientific explanation of heredity." Despite criticism of Mendel's work, which he ignored, or failed to report data that did not support his hypothesis, Morgan's statements and publications began to credit Mendel for discovering the principles of heredity (Bellen and Yamamoto 2015). From 1926 onwards, with the publication of the book *The Theory of the Gene*, Morgan definitively incorporated the term "gene" into his manuscripts. From that moment on, it was widely adopted by the scientific community.

While his studies with *Drosophila* continued, Morgan was also busy teaching and developing other projects. Gradually, he became interested in physics and chemistry, as he understood that both were essential for genetics. In 1928, Morgan, aged 62, went to work at Caltech (California Institute of Technology). By then, he had already gained a worldwide reputation as a remarkable teacher, writer, and impressive researcher. In the genetics laboratory at Caltech, the atmosphere of the fly room was recreated, and the studies continued. The culmination of his brilliant career was recognized in 1933 when Thomas Hunt Morgan was awarded the Nobel Prize in Physiology or Medicine for discovering the chromosomal mechanism by which the phenotypes of different traits are passed from father to son.

During a visit to Caltech, a particularly picturesque situation involved a conversation between Morgan and Albert Einstein. When asked by Einstein, "What the hell are you doing in a place like this?", Morgan replied, "The future of biology lies in the application of the methods and ideas of physics, chemistry and mathematics." However, Einstein persisted, "Do you think you will ever be able to explain in terms of chemistry or physics a biological phenomenon as important as first love?". Then Morgan attempted to explain to him something about the connection between the sense organs, the brain, and hormones. He did not truly believe in it either, but he felt he had to tell Einstein something (Rédei 2002).

The research of Barbara McClintock marked another important period in post-Mendelian genetics. Like Mendel, she worked largely alone and devoted most of her scientific life to studies on corn. Barbara McClintock began her scientific career at Cornell University, where she pioneered cytogenetic studies. The marriage of cytology and genetics became official in 1931 when McClintock and Harriet Creighton provided the first experimental proof that genes were physically positioned on chromosomes, describing the crossing-over phenomenon. McClintock's observations allowed her to prove Thomas Morgan's ideas about the link between genetic traits and the exchange of segments between homologous chromosomes (Pray and Zhaurava 2008).

A few years later, a discovery by McClintock impacted genetics. By analyzing corn cobs with grains of varying colors, she discovered the presence of jumping genes, initially identified as Ac/Ds, which could change their position within the chromosome and reversibly alter the expression of other genes. The results showed that corn was the perfect model organism for studying the first transposable elements. The pioneering information was published in 1950 in the Proceedings of the National Academy of Sciences (PNAS), *The origin and behavior of mutable loci in maize*. McClintock's work was revolutionary because it suggested that an organism's genome is not a stationary entity but is subject to changes and rearrangements - a concept criticized by the scientific community at the time. However, the role of transposons later became widely appreciated, and in 1983, McClintock received the Nobel Prize in Physiology or Medicine in recognition of this and her other contributions to genetics (Pray and Zhaurava 2008).

During the award ceremony, the Royal Swedish Academy of Sciences compared Barbara McClintock to Gregor Mendel, publicly demonstrating the relevance of her research to genetics. With the advent of the genomic era in the late 1990s, McClintock's insight became even more evident, as the availability of genome sequences accelerated the study of transposable elements (TEs) and demonstrated that they have a profound impact on evolution and genome rearrangement, driving genetic diversity and speciation. The molecular mechanisms of "cut and paste" (transposons) or "copy and paste" (retrotransposons) were elucidated, and it was found that each new copy of a retrotransposon is potentially as active as the original sequence, leading to the possibility of exponential increases in the number of copies (SanMiguel et al. 1996). To understand the importance of this discovery, it is worth mentioning that transposable elements play a fundamental role in studies on the origin of grasses such as barley, rice, sorghum and corn. Different levels of TE accumulation have been observed since the divergence from a common ancestor over a few tens of millions of years.

Another important aspect is that TEs can influence genetic regulation through epigenetic modifications and play a fundamental role in shaping the expression patterns of neighboring genes, demonstrating their potential to produce various phenotypic results and adaptation to abiotic stress. More recently, it has been found that TEs have had and should continue to have a fundamental role in the phenomenon currently called Pangenomics, which will be reported in more detail later.

FROM THE DISCOVERY OF NUCLEIC ACIDS TO GENOME SEQUENCING

Coincidentally, between 1869 and 1871, after the discovery of Mendel's Laws, a young Swiss biochemist named Friedrich Miescher isolated a phosphorus-rich substance from the cell nucleus, which was not found in known proteins or lipids. He was very surprised and called this substance nuclein without fully understanding its purpose. Later, in 1889, Richard Altmann renamed nuclein to nucleic acid. Research on this chemical substance conducted during the first decade of the 20th century led biochemist Phoebus Aaron Levene to discover two types of nucleic acids. One contained a sugar called ribose, giving rise to what we know as ribonucleic acid (RNA). The other contained a sugar missing an oxygen atom on one of its carbons, deoxyribose (deoxyribonucleic acid – DNA). Furthermore, he discovered that DNA contained four nitrogenous bases: Adenine (A), Guanine (G), Thymine (T), and Cytosine (C). At that time, no hypothesis suggested that nucleic acids were involved in heredity.

However, some reflections on the genetic control of traits began to emerge. In 1909, the English physician A. E. Garrod wrote the book *Inborn Errors of Metabolism*, in which he was the first to associate human anomalies with the action of genes and enzymes. A few years later, these observations started to get confirmed through the research of American scientists George W. Beadle (1903–1989) and Edward L. Tatum (1909–1975), who proposed the "one gene, one enzyme" hypothesis.

The indication that nucleic acids were involved in the process began with studies on bacterial transformation in *pneumococcus*. These studies were conducted by the English physician Frederick Griffith in the early 20th century. He had already identified that pneumococci could be distinguished by the presence of a polysaccharide capsule, which gave colonies a smooth appearance when present. In contrast, its absence resulted in colonies with a rough appearance. He also knew smooth bacteria were pathogenic when inoculated into mice, while rough ones caused no harm. Griffith conducted a simple experiment, relying on a bit of luck. He exposed smooth bacteria to heat, which killed them, rendering them harmless when injected into mice. He then combined live rough (non-pathogenic) bacteria with heat-killed smooth bacteria and injected this mixture into mice. Surprisingly, the mice died. When analyzing the dead mice, he recovered

smooth bacteria, indicating that something from the heat-killed smooth bacteria had passed to the live rough bacteria, resulting in what became known as bacterial transformation. Griffith appeared to have been quite fortunate in his discovery. It is now known that both proteins and nucleic acids are inactivated when exposed to high temperatures. However, proteins denature at lower temperatures than DNA. It can be inferred that if he had heated the bacteria to a higher temperature than he did, he would not have discovered anything about the mechanism of bacterial transformation.

The debate over whether protein or DNA was the hereditary molecule persisted for several years until three researchers - Oswald Avery, Colin MacLeod, and Maclyn McCarty - built upon Griffith's work and determined that the nature of the transforming principle was due to deoxyribonucleic acid (DNA). They demonstrated that DNA was the fundamental unit capable of transforming non-pathogenic bacteria into pathogenic ones, rather than proteins. From 1944 onward, it became established that DNA was the molecule involved in heredity.

One of biology's greatest challenges lies in the close relationship between structure and function. Thus, what is the molecular structure of DNA? This question captured the attention of numerous researchers, but the answer was proposed in 1953 by two scientists: physicist James Watson and chemist Francis Crick. They, along with Maurice Wilkins, were later awarded the Nobel Prize in Medicine or Physiology in 1962. Notably, the paper *Molecular structure of nucleic acids*, published in *Nature*, which earned them the prize and revolutionized biology, was only one page long. Their proposal for DNA's structure was made possible thanks to the remarkable imagination of the two scientists, combined with key prior discoveries. One of these was made by Austrian Erwin Chargaff in 1949, who showed that in DNA molecules, adenine (A) equals thymine (T), and guanine (G) equals cytosine (C). However, the proportion of A+T relative to G+C varied between species while remaining consistent across different tissues of the same species. The findings obtained through X-ray diffraction by Rosalind Franklin and Maurice Wilkins in 1952 were even more crucial. The stunning image known as "Photograph 51", produced by Franklin, revealed essential clues about DNA's molecular structure. Remarkably, more than 70 years later, during an era of intense molecular biology research with far greater financial and technological resources, Watson and Crick's proposed structure has remained virtually unchanged.

The next step was to demonstrate how the DNA molecule replicates to transmit hereditary information between cells during cell division and between parents and their offspring during gamete formation. To achieve this, Meselson and Stahl conducted one of the most ingenious experiments in biology in 1958. They used the ability to label nucleic acids with different nitrogen isotopes. They knew that when bacteria were grown in a medium containing ^{15}N , their DNA became denser than those grown in ^{14}N . Initially, they cultured *Escherichia coli* in a ^{15}N medium for several generations to ensure that all DNA became "heavy." Then, they transferred the bacteria to a ^{14}N medium and allowed them to multiply once (one generation). DNA samples from the daughter bacteria were then centrifuged, revealing an intermediate density compared to DNA made exclusively of ^{14}N or ^{15}N . Repeating the process for additional generations demonstrated that DNA replication is semiconservative. During replication, the two strands of DNA separate, each serving as a template for synthesizing a new complementary strand. The enzyme responsible for this duplication process, DNA polymerase, was discovered by two scientists, Arthur Kornberg and Severo Ochoa, who were awarded the Nobel Prize in Physiology or Medicine in 1959.

With the unveiling of the DNA molecule model, many questions arose. One of the most intriguing was that the four DNA bases encode the 20 amino acids (AAs) involved in synthesizing individuals' vast array of proteins. In other words, the challenge was to decipher the genetic code. Some scientists considered this period, which unfolded over a short span in the 1960s, the most thrilling era in science. Initially, a mathematical inference was made: How could the four DNA bases encode the 20 AAs found in proteins? For this, the unit of the code had to consist of at least three bases. This reasoning was based on the calculation $4^3=64$, providing more than enough combinations to encode the 20 AAs. The actual deciphering of the code only became feasible when scientists succeeded in synthesizing proteins *in vitro* in a cell-free system. As a result, the 64 triplets, or codons, were rapidly identified, along with the properties of this "code of life." At this point, it is worth recalling the definition of a gene at the time: "a segment of the DNA molecule, located at a specific position on a particular chromosome, that contributes to the phenotypic expression of a given trait."

The concepts grew more complex when scientists Richard Roberts and Phil Sharp (awarded the Nobel Prize in Physiology or Medicine in 1983) demonstrated that eukaryotic genes consist of coding regions (exons) and non-coding regions (introns). Thus, the process of transcribing DNA into RNA involves additional steps to remove introns and join

exons - a phenomenon known as RNA processing or splicing. The existence of introns was initially met with much skepticism. However, it was later discovered that some genes undergo alternative RNA processing. This means that the same sequence of a given gene can be used to encode entirely different proteins. While many questions remain, it is now evident that introns play a role in regulating gene expression.

The transfer of information to synthesize RNA molecules occurs through the complementary copying of one of the DNA strands at specific sites. This process is called transcription. Specific enzymes, such as RNA polymerases, are involved in this mechanism. Different types of RNAs are produced, whose functions have been described over time. Among them are the various messenger RNAs (mRNAs), which are involved in producing the phenotypes of different traits of individuals throughout their lifespan. Another type is ribosomal RNAs (rRNAs), which assemble ribosomes with other proteins, where polypeptide chain synthesis occurs. Additionally, there are transfer RNAs or transporter RNAs (tRNAs), whose function is to transport amino acids (AAs) to the ribosomes during protein synthesis. Other types of RNAs are also produced, such as miRNAs, snRNAs, and piRNAs, whose functions are still being described. Transcription factors, which will be discussed later, are transcribed to produce mRNAs that form proteins. These proteins act as enzymes in regulating the transcription of other mRNAs, which, in turn, will be responsible for the phenotypic expression of a given trait. Details about the transcription process and the roles of the numerous RNAs can be found in various literature sources (Griffiths et al. 2016, Ramalho et al. 2021).

Translation is the term for synthesizing a polypeptide chain from the message encoded in mRNAs. Although there are differences in translation between prokaryotes and eukaryotes, the principle remains the same. However, many questions remain unanswered despite the vast amount of knowledge accumulated about the translation process since the discovery of the DNA molecule's structure. This is particularly true in the case of regulatory processes during and after translation.

Some research that greatly impacted biotechnology, seemingly unrelated to what happened later, came from studies on bacterial defense mechanisms. This is because bacteria are frequently exposed to exogenous DNA molecules due to their reproductive systems, which could threaten their existence as a species. Werner Arber, Daniel Nathans, and Hamilton Smith made the first such breakthrough. They discovered that some bacteria produced restriction enzymes capable of cutting exogenous DNA at specific points. Since each bacterial species had a unique restriction enzyme, this discovery allowed various bacteria to be identified and used to cut DNA into segments of different sizes. Sometime after the discovery, these enzymes started to be commercialized and used to cut DNA from any source.

Another restriction enzyme was discovered in 1987 by a research team led by Y. Ishino in Japan, who identified repetitive DNA sequences now known as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). This work was further complemented in 2002 when genes adjacent to the CRISPR sequence, known as Cas (CRISPR-associated proteins), were identified. In 2008, an RNA sequence was also discovered that guides the Cas protein to its target. For example, CRISPR-Cas9 can break both strands of DNA at specific points. This marked the starting point for gene-editing technology, which will be discussed later.

However, it was soon discovered that having various DNA segments obtained through restriction enzymes was insufficient without a mechanism to amplify them. In other words, there was a need to produce multiple copies of a DNA sample in the shortest time possible to provide more detailed studies. The solution to this problem began in 1983 when Kary Mullis (awarded the Nobel Prize in Chemistry in 1993) described the process. Later, with the involvement of other researchers, the PCR (Polymerase Chain Reaction) technique was developed and became widely known by the end of that decade. PCR, combined with restriction enzymes, was fundamental for molecular biology to be effectively recognized as a science capable of generating technology of immense importance for humanity.

With all these advancements, the need to understand the base sequence of the genomes of different species became increasingly urgent. It all began years earlier with the creation of the methodology for DNA sequencing by Frederick Sanger and his collaborators. A few years later, organisms with small genomes were sequenced. However, the greatest emphasis was placed in the 1990s with the beginning of the Human Genome Project, which aimed to sequence the human genome.

The excitement was immense regarding the possibility of better understanding the biological processes of the human

being and the pharmaceutical industry's anticipated opportunity to develop new drugs. This was a very particular period in science, as the magnitude of the challenge was so vast that it required the participation of multiple groups and institutions worldwide, coordinated by American researchers. The initial expectation was that the project would take several years to complete. However, humanity's ability to create new equipment accelerated the process and drastically reduced its cost. As a result, by early 2000, the first draft of the human genome was released. The scientific community's response was monumental. However, the definitive conclusion (the complete sequencing of the genome) was only achieved in 2022.

Among other remarkable findings, the scientific community was surprised to discover that the number of genes in the human species was much lower than expected. The initial expectation was to identify over 100,000 genes, coinciding with the estimated number of proteins humans need to perform their functions. However, approximately 25,000 genes were identified in the human genome, highlighting greater complexity in regulating gene expression. It was also found that the number of genes varies among individuals of the same species, giving rise to a new field of study, Pangenomics, which will be discussed later. Numerous genes have been identified, but much work remains to elucidate all genes' functions fully. The study of the model plant *Arabidopsis thaliana* genome has served as a reference for identifying the occurrence of orthologous genes. Additionally, new terms have been incorporated into the genetics vocabulary, such as haplotypes, contigs, and reads.

With the feasibility of complete genome sequencing, from one telomere to the other (symbolized as T-2-T) and involving multiple individuals of the same species, it was found that the genes of individuals are not identical. The first report of this fact occurred in the early 21st century (Tettelin et al. 2005). They sequenced the genomes of six strains of *Streptococcus agalactiae* and found that the six genomes were not identical. Since then, with the advancements and lower cost of complete genome sequencing, pangenomics has been observed in many organisms, including humans. This means that different individuals in a population do not share the same genes. Commenting on this, Yang et al. (2019) emphasize that a single genome from an individual is not enough to represent the diversity of a species. In the past decade, there has been a surge in publications on this subject, including several cultivated plant species, such as maize (Cannon et al. 2024), beans (Cortinovis et al. 2024), eucalyptus (Lötter et al. 2023), soybean (Ni et al. 2023), and many others.

Numerous studies have demonstrated the existing diversity within the species' genome in maize cultivation. One example is the work of Fu and Dooner (2002), who, even without using the term pangenomics, compared two widely known lines, B73 and Mo17. They observed a significant difference in the amount of DNA between these lines. They associated this difference with one of the likely causes of the high vigor observed in various traits in the hybrid resulting from their cross. They even suggested that heterosis, in this case, could be due to the complementation of the different genes between them.

Commenting on the occurrence of pangenomics in maize, Andorf et al. (2019) cite the research conducted by Hirsch and other authors in 2014. They mention that the genome of B73 was compared with 503 other lines with broad diversity. It was found that only 38% of the genes were common to all. That is, 62% of the genes were part of the pangenome. Cannon et al. (2024) provide information about the maize genome using the maize genetic and genomic database (MaizeGDB, Maize Genetics and Genomics Database - <https://www.maizegdb.org>). This community-based database provides data curation and computational resources to support maize genetics, genomics, and breeding research.

Although this classification is not widely used in many articles, the genome has been grouped into Core, Shell, and Cloud. The Core includes genes that are present in all the analyzed genomes. The Shell genome consists of genes most genome shared (10-95% occurrence). Gene families in only one genome or with less than 10% occurrence are described as the Cloud genome (Bezuidt et al. 2016).

The consensus is that pangenomes are associated with genes conferring tolerance to biotic and abiotic stresses. They likely originated from transposons and retrotransposons, highly dynamic genome components. They may also have arisen through mechanisms not yet fully understood, involving recombination events between DNA sequences from non-homologous chromosomes - a previously unimaginable phenomenon (Chen et al. 2018).

An important question arises: What is the utility of Pangenomics knowledge? Much of the understanding of this

phenomenon is directed toward evolutionary studies. For instance, Cortinovis et al. (2024) mention that domestication contributed to the shrinking of the pangenome in the common bean (*Phaseolus vulgaris*). In other words, human selection during domestication reduced the number of genes in the process. They emphasize that the partial or complete loss of genes represented an adaptive genetic change. However, it is also plausible that with the expansion of cultivation into new regions significantly different from the species' area of origin and the occurrence of various biological and ecological stresses, natural selection may have contributed to the emergence of previously nonexistent genes. It is worth noting that Mendel's Laws remain applicable even in the context of Pangenomics, this is because when crossing a line containing a specific gene absent in the other, the segregation in the F_2 generation will still follow a 3:1 ratio. The absence of the gene acts as if it were a recessive allele.

The knowledge of Pangenomics can significantly impact plant breeders' activities, especially regarding intrapopulation recurrent selection (IRS). While reciprocal recurrent selection (RRS), which works with two distinct heterotic populations, does not directly benefit from this phenomenon, since it is limited to two genomes of the species and does not introduce new germplasm, IRS can take advantage of the genetic diversity provided by Pangenomics. In SRI, which aims to increase the frequency of favorable alleles and improve the average trait in a population, the Pangenomic phenomenon can enrich the genetic pool with each selection cycle. By incorporating individuals with genes not present in the original population, recombination can generate new genetic material, increasing variability and potentially improving plant performance against biotic and abiotic stresses, disease resistance, or other environmental challenges. Using germplasm banks, combined with whole genome sequencing, will allow plant breeders to identify and select useful genes that were not previously recognized. This will enable the inclusion of these genes in IRS populations, enhancing breeding efforts and, more precisely, adapting plants to cultivation conditions. Furthermore, selecting new genes specific to resistance to pathogens or environmental stresses could transform breeding programs, increasing their efficiency and effectiveness. In summary, Pangenomics offers a new dimension for plant genetic improvement, including greater genetic diversity, which can accelerate the development of more resistant and productive cultivars.

As previously emphasized, the biggest challenge that remains to be elucidated is the issue of gene expression regulation, although many aspects have already been uncovered. Among the various unanswered questions is why there are more proteins than genes. Probably, exon shuffling and post-translational protein rearrangements, addressed in the field of molecular biology known as proteomics or, even more advanced, metabolomics, may contribute to the explanation. Furthermore, epigenetic mechanisms that modulate chromatin should also be considered. In fact, in the context of gene expression regulation, the 2024 Nobel Prize in Physiology or Medicine was awarded to scientists Ambros and Ruvkun, who demonstrated in research conducted a few years ago, the role of microRNAs (miRNAs), which contain 21 to 25 nucleotides, as silencers of some genes post-transcription. To perform their functions, miRNAs pair with specific mRNA and regulate their stability and translation. In the October 2024 issue of Revista FAPESP, biomedical scientist Marcelo Mori commented on the work of these two scientists, expressing the following: "MicroRNAs function like the conductor of an orchestra. They define which musician [gene] will play, at what time, and with what rhythm and intensity. They fine-tune the amount of proteins produced by cells, something important for both development and the functioning of organisms."

With the advancement of the knowledge already mentioned, it becomes difficult to have a concept of genes that can encompass all the existing information. For this reason, one of the editors of *Nature* magazine, Pearson (2006), mentioned that: "Gene is not a typical four-letter word. It is not offensive. It is never bleeped out of TV shows. Moreover, where the meaning of most four-letter words is all too clear, that of gene is not. The more expert scientists become in molecular genetics, the less easy it is to be sure about what, if anything, a gene is." In the same text, it is mentioned that Professor Karen Eilbeck, from the University of California, Berkeley, invited 25 scientists to discuss what a gene is. After extensive debate, they proposed the following definition: "A locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions and/or other functional sequence regions." Nearly ten years later, the geneticists from ENCODE, a group involved in sequencing the human genome, commented, with the information now available, on the difficulties of having a single concept of what a gene is (Gerstein et al. 2007). However, they offered a suggestion that is not much different from the one previously presented: "A gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products."

Additionally, everything that has been discussed so far, focusing on the DNA molecule and the effect of the environment

on the regulation and expression of genes, cannot be disregarded. To illustrate, let's consider the timing for the onset of flowering in a plant. It has been well established that environmental factors such as temperature, humidity, and day length are relevant in the manifestation of this trait. All indications are that they act on transcription factors, producing RNAs and proteins that will subsequently regulate the expression of the flowering genes themselves. These genes will then be transcribed and translated, producing the proteins/enzymes initiating the flowering process.

APPLICATION OF DNA MOLECULE KNOWLEDGE IN PLANT BREEDING

As previously mentioned, the information obtained from the DNA molecule has unveiled a vast range of opportunities for its application in plant breeding. The approach to developing new cultivars has undergone profound changes. The interest among young professionals in this field has been immense, and graduate programs (both pre-existing and newly established) have emphasized training in understanding the DNA molecule and, more importantly, its applications. This section will briefly discuss what has happened and what is currently happening in this new phase of plant breeding.

Molecular markers in plant genetic breeding

It is essential to emphasize that every marker used in plant genetic improvement relies on two fundamental principles of Mendelian genetics: pleiotropy and gene linkage. In this context, some markers have been employed for a considerable time, including phenotypic markers and certain enzymes. However, these markers have a very limited scope. With the advent of DNA molecule-based information, some of which has been previously discussed, its use as a powerful tool for identifying individuals with desirable phenotypes indirectly, through DNA information, became evident. Unsurprisingly, the scientific community experienced great enthusiasm regarding this possibility.

The expectation surrounding RFLP (Restriction Fragment Length Polymorphism), one of the first molecular markers developed, was immense. In this context, it is important to highlight Edwin Southern's contribution in 1975 when he developed the Southern Blot technique, which was essential for visualizing DNA fragments obtained by exposing the molecule to various restriction enzymes. RFLP markers were widely used throughout the 1980s and 1990s, enabling, for the first time, the differentiation of genotypes based on DNA fragment patterns. It was as if scientists had discovered a genetic "fingerprint", marking the beginning of a new era in molecular biology. However, despite its innovation, RFLP had some limitations: it was expensive, labor-intensive, and required DNA samples of exceptionally high quality.

The search for new types of DNA markers began intensely, leading to the development of RAPDs (Random Amplified Polymorphic DNA), AFLPs (Amplified Fragment Length Polymorphisms), and microsatellites (or SSRs, Simple Sequence Repeats), among others. These techniques accelerated genetic analyses and simplified processes, significantly broadening the range of practical applications in many other fields.

However, in genetic breeding, the gain achieved through marker-assisted selection (MAS) can be compared to the gain expected from phenotypic selection (PS). Considering the same selection intensity applied in both cases, the relationship is expressed as: $\frac{MAS}{PS} = \sqrt{fm/h}$ where fm is the proportion of additive genetic variance explained by the marker, and h is the square root of the heritability. It becomes evident that marker-assisted selection is advantageous when the trait's heritability is low, below 0.3. However, for fm to be reliable, it must be derived from well-conducted experiments where heritability is high. This requirement explains why the successful application of markers has predominantly occurred in traits controlled by a small number of genes, often with high heritability but challenging phenotypic evaluation. Among many notable examples, markers have been a highly successful application in selecting soybean plants resistant to soybean cyst nematode (*Heterodera glycines*).

Regarding quantitative traits such as grain yield, tremendous efforts have been made to identify QTLs (Quantitative Trait Loci). However, the success has been limited. In this context, Andorf et al. (2019) stated: "While the numbers of mapping studies and identified QTL and loci are overwhelming, the statement of Bernardo (2009) is certainly still valid: "the vast majority of the favorable alleles at these identified QTL reside in journals on library shelves, rather than in cultivars that have been improved through the introgression or selection of these favorable QTL alleles."

A new wave of optimism for using markers emerged with the next-generation sequencing (NGS), which enabled the

identification of millions of SNPs (single-nucleotide polymorphisms) across the genome. This advancement expanded possibilities and applications in genetic breeding and evolutionary studies, including genomic-wide selection (GWS) and genome-wide association studies (GWAS). The hope was that SNPs, combined with predictive models utilizing more powerful biometric techniques, would help the analysis of larger datasets and aid in predicting phenotypes for quantitative traits. This approach would enable early plant evaluation before the expression of the desired phenotype, resulting in significant time savings in the selection process. Simulation studies demonstrated the potential efficiency of this new selection method. However, for most cases, the accuracy estimates obtained, particularly for traits such as grain yield or wood volume, have not been of small magnitude, limiting the effectiveness of this approach.

Due to its low efficiency, the use of SNPs has been questioned. Commenting on this topic, Bhat et al. (2021) emphasized that SNP-based models are inherently biallelic, which leads to several limitations: a low level of polymorphism, a higher probability of false positives, loss of rare alleles, inability to detect epistatic interactions, reduced power in identifying marker-trait associations, and low accuracy in the genomic prediction of breeding values (GEBVs). Consequently, haplotypes have been proposed for genomic selection (Bevan et al. 2017, Varshney et al. 2020, Bhat et al. 2021, Liu et al. 2023). The fundamental difference is that genomic selection traditionally focuses on genes/loci to identify desired alleles. In the case of haplotypes, the focus is on small genomic regions where recombination is unlikely, involving a small number of genes with favorable alleles, which are selected together (Rai and Tyagi 2022).

The development of genetically modified organisms

The use of genetically modified organisms (GMOs) began with American biochemist Paul Berg, who demonstrated that it was possible to “reassemble” DNA, creating combinations that do not occur naturally. His research made genetic manipulation a reality. One of the first practical applications of this technique was conducted by Cohen et al. (1973). Using restriction enzymes, they inserted a gene for antibiotic resistance from the plasmid of one *E. coli* strain into another strain of the same species. Their work demonstrated that living organisms could artificially receive and express genes from other organisms.

Shortly afterwards, using the same principle, scientists successfully inserted two human genes responsible for producing human insulin into *E. coli*. Remarkably, both genes were expressed in the bacterium, and more importantly, the insulin they produced was excreted. This achievement was announced in September 1978. Subsequently, strategies were developed *in vitro* to combine the two insulin chains, A and B, through biochemical techniques. Simultaneously, the process of approving the commercialization of transgenic insulin began, which was achieved in 1982 in several countries. This breakthrough enabled enough insulin production to meet the global demand of millions of diabetic individuals. It is worth noting, however, that despite significant advancements in biotechnological techniques over the past four decades, and the increasing availability of human and financial resources for human genetics, no achievement comparable to the insulin production has been accomplished to date, as far as current evidence suggests.

The potential use of genetically modified organisms (GMOs) in plants was undoubtedly a significant achievement for genetics and plant breeding. One of the first articles published in *Science* clarifies this, specifically in Abelson’s *A Revolution in Biology* (Abelson 1980). In this article, for example, the author discusses one possibility that would have immense societal impact: “At the applied level, I believe that the most important and pressing applications will come in providing solutions to our energy problems. Photosynthesis is a solution to the need for a renewable energy source that has been perfecting itself for over three billion years. Almost certainly, an important component of the solution to the energy problem will come from biology. Engineering may be required to adapt to our needs for the biological conversion of sunlight to carbon skeletons. It should be evident that many of the tools needed for this task are already at hand.”

The early years were marked by intense euphoria in society as a whole. Some believed that the potential of GMOs could be fully realized. On the other hand, many people expressed significant concerns about the possible consequences of using GMOs. Forty-four years after the publication of the mentioned article, it became evident that the radical views of both groups did not come to fruition. There have been some notable successes using GMOs in plants, but not to the extent originally anticipated. Furthermore, no conclusive evidence has been found regarding the effects of GMOs on public health or any significant negative environmental impacts.

There is no doubt that the development of certain GMOs has had a significant impact on agricultural management,

especially in large-scale crops such as corn and soybeans. Among these achievements is the creation of GM plants that assist in the application of post-emergence herbicides for weed control. This feat was unimaginable through conventional genetic improvement methods. The same has occurred with pest insects, where genes from microorganisms previously used in biological control were transferred to cultivated species. In both cases, weed control and pest management, the success of GMOs has had and continues to have a tremendous agronomic, economic, and environmental impact.

However, despite the enormous efforts of the global community of thousands of scientists involved in the use of GMOs, more complex problems, such as improving photosynthetic efficiency and drought tolerance, have not yet been solved. It should be emphasized, however, that articles are frequently published on developing GM plants with improved photosynthetic efficiency (Souza et al. 2020) and drought tolerance (Khan et al. 2019). The hope is that these promises will become effective under field conditions, which has not yet occurred.

The efforts continue; however, over more than 40 years, knowledge in molecular biology has shown that the expression of phenotypes is far more complicated than originally thought. DNA alone does not hold all the answers.

The prospects of gene editing in solving agricultural problems

Dr. C.L. Auerbach (1899-1994), a German geneticist who made a decisive contribution to the development of induced mutations in the 1940s, quickly realized that, regardless of the technique used, mutations occurred entirely at random. Due to this, Dr. Auerbach once said: "If humans could direct mutations, they would no longer be slaves to the past, but masters of the future." Had she been alive today, she would be astonished by what is now being imagined and achieved with gene editing technology.

Countless relevant insights gained from molecular biology, as discussed earlier, were fundamental for two scientists, J. Doudna and E. Charpentier, who won the 2020 Nobel Prize in Chemistry. In 2014, they published the article *The New Frontier of Genome Engineering with CRISPR-Cas9* in Science (Doudna and Charpentier 2014). Once again, it seemed like a new revolution in molecular biology, with broad potential applications in plants and animals. Since then, the number of publications on gene editing has grown dramatically. This is because the new technique enables the alteration of gene expression in individuals and, according to some molecular biologists, offers several advantages, including significantly speeding up the development of new cultivars.

Although it is not often discussed, the significant advantage of using gene editing over GMOs, as previously mentioned, is the ability to always use genes that have been in vitro edited from the same species. This is a huge advantage. Nevertheless, it should not be considered an easily applicable technology. Literature reports frequently highlight successes achieved in certain species (Ahmad 2023).

Reflections of Mendel's Laws in the production of grains, fruits, fibers, and energy on the planet

Available data suggest that our species, *Homo sapiens*, emerged approximately 400,000 years ago. In 1865, the year of the publication of Mendel's Laws, the global population was 1.378 billion. From 1900, when Mendel's Laws were rediscovered, to 2025, in just 125 years, the population increased by 6.7 billion people, growing from 1.5 billion in 1900 to 8.2 billion in 2025. The magnitude of this population growth within a relatively short period becomes even more surprising considering the various factors that contributed to reducing its scale. Among these, the occurrence of several pandemics, two world wars, numerous smaller conflicts, the use of condoms, and the discovery of birth control pills stand out.

After all, what do Mendel's laws and genetics have to do with this? At first, genetics contributed to advances in medical knowledge. This, in turn, significantly increased life expectancy over time due to, among other factors, the development of antibiotics, vaccines, and various technologies. Genetics played a crucial role in generating information that directly or indirectly contributed to the creation of these products. Additionally, for the population to maintain good health, it is necessary to have sufficient quantity and quality of food, as well as other essential products like fiber and energy. Moreover, it is important to remember that the well-being of the population depends on producing agricultural products with minimal environmental impact.

In brief, the success and increased efficiency in food production over the past 160 years, aided by knowledge provided

by genetics, will be discussed. However, it is impossible to cover all advancements in productivity and food quality, the focus will be on a few species of greater importance to humanity, including rice and wheat, which played a crucial role during the greatest population growth between 1960 and 1980, in what came to be known as the Green Revolution. This revolution was based on three pillars: the use of fertilizers, irrigation, and short-statured cultivars. Rice production worldwide, before 1960 and today, has predominantly taken place in flooded fields. The cultivars available at that time were very tall, and when nitrogen fertilizers (N) were used, the plants grew excessively and lodging. Under such conditions, instead of increasing grain yield, it was reduced. Everything changed when the most commonly used cultivar at the time, 'Peta' was crossed with the short-statured line 'Dee-gee-woo-gen'. After selection, lines were obtained that led to the development of cultivars like 'IR-8'. These cultivars combined high grain productivity with an intermediate stature, preventing lodging at harvest time, even when using nitrogen fertilizers. From then on, the success of rice cultivation was enormous. A similar phenomenon occurred with wheat.

Over time, these two species have undergone numerous other genetic technologies that have been applied, such as hybrid production. In the case of rice, it all started in China around 1973. In this context, the work of Chinese breeder Yuan Longping (1929-2021) must be mentioned. He discovered the first male sterile lines, which made the commercial production of hybrids viable. These hybrids spread to most rice-producing countries. In recognition of his work, in 1999, the International Astronomical Union named one of the discovered asteroids "Yuan Longping Asteroid" (Xin et al. 2023), honoring his lifelong contribution to preventing world hunger. The use of most new technologies in various areas of genetics is present in rice and wheat breeding programs aimed at increasing the efficiency of developing new cultivars worldwide.

Corn is another grass species of immense importance for the world's food supply. It is one of the species with the greatest involvement of genetic technologies to improve the efficiency of developing new cultivars. Taking the example of corn cultivation in the United States, it is easy to see the importance of genetics in its agricultural success and the generation of numerous commercial products. It originated in Mexico. Corn samples were brought to Europe when the Americas were discovered in 1492. There, it was cultivated and likely subjected to cycles of visual selection. Its return to the United States occurred at the end of the 18th century.

A detailed account of what happened later to corn grain productivity in the United States was provided by Nielsen (2023). He noted that productivity estimates have been available since 1866, when it was 1,340 kg ha⁻¹. After 70 years, the increase was very small, as in 1937, the productivity was only 1,600 kg ha⁻¹. According to the author, from then on, the first "miracle" occurred with the adoption of hybrid corn. The rapid adoption by farmers was facilitated because, that year, the drought was intense, and the double hybrid used was not only productive but also drought-tolerant. With the commercialization of hybrids, seed companies were created, and along with the public sector, research involving genetics and management was intensified. The second miracle occurred just a few years later, in 1955, with the availability of single hybrids. For the period from 1956 to 2022, for example, the linear regression coefficient (b) between the variables, years of data and grain productivity was estimated. The estimated b value was 126.9 kg ha⁻¹year⁻¹, with a coefficient of determination (R²) of 0.93. To reinforce what was stated, it is important to mention that the average grain productivity in 2022 was 11,621 kg ha⁻¹, a value 3.65 times higher than in 1957. Nielsen (2023) also emphasizes that although the average productivity varies significantly from year to year, there is no trend of stagnation.

What happened with corn cultivation in Brazil also deserves reflection. In the early 20th century, although there were difficulties in obtaining reliable productivity estimates, it was similar to what was reported in the United States at the same time, according to reports by Benjamin Hunnicutt in 1914. Brazil was the second country to adopt the use of double-hybrid corn. The work to develop hybrids began at the Agronomic Institute of Campinas (IAC), and the first results were obtained in 1939. They were commercialized in the 1953 harvest, about 16 years after the United States. The first private company focused on corn production was Agroceres, founded in 1945.

A remarkable fact in corn cultivation in Brazil is that until 1980, its cultivation in the Southern and Southeastern regions, where the largest production in the country was concentrated, was sown between August and September in the South and from October 15 to November 15 in the Southeast. When sowing was done after the latter date, productivity reductions were significant. Surprisingly, in the state of Paraná and later in other regions, especially in the Central-West, corn cultivation began after the soybean harvest, at the end of January, often extending into March. In

this condition, the risk to the crop was enormous, particularly due to reduced precipitation (rain), especially when it coincided with the grain-filling period. As expected, the area was initially small and productivity was low. However, by combining the development of new hybrids with some management techniques, the production of the first and second harvests became similar in less than 20 years, contrary to all expectations. Subsequently, in the second harvest, the area and grain productivity continued to grow. For example, in the 2023/2024 period, the area cultivated in the first harvest was 3.9 million hectares, while in the second, it was 16.4 million hectares. Since the productivity was similar, the grain production in the second harvest was four times higher than in the first.

Soybeans, the most widely cultivated legume in the world, are an important source of protein and oil. It is believed that they originated in Manchuria, a region in northern China with high latitude. A typical short-day plant, soybeans are highly sensitive to photoperiod. Flowering occurs prematurely when introduced into regions with lower latitudes, resulting in low biomass production. This initially limited their spread to other regions of the planet, particularly those with lower latitudes. The understanding of photoperiod sensitivity began when Gardner and Allard, in 1920, first discovered that day length is a dominant environmental factor influencing soybean flowering, associated with genes involved in the process. Over the past century, researchers have sought to understand how the genetic control of this trait works. It was found to be oligogenic. By working with these genes, it became possible to develop new soybean lines that allowed the cultivation of this legume in regions with significant variation in latitude, i.e., differing day lengths during the growing season (Zhai et al. 2022, Sun et al. 2024).

The evolution of soybean cultivation in Brazil is worth telling, highlighting the contributions of Mendelian genetics and subsequent genetic knowledge. It is believed that soybean seeds were first introduced to Brazil in the 19th century, but with little success. Later, introductions were made using germplasm selected from the United States. However, the crop became viable only in the state of Rio Grande do Sul (RS). In fact, at the end of 2024, the soybean introduction centenary was celebrated in Santa Rosa (RS). Its spread to the rest of the country, especially to regions closer to the Equator, only became feasible in the 1970s when plants with a long juvenile period were selected. These plants vegetate first and then flower, ensuring that the grain production justifies the economic viability of the crop. The success was immense, mainly because today, it is possible to produce soybeans in Brazilian regions close to the Equator, at zero latitude, in a relatively short period, thanks to breeding programs.

It should also be highlighted that soybean cultivation in Brazil expanded primarily into regions with Cerrado vegetation. These areas had little tradition in grain production, mainly due to high aluminum saturation, very low phosphorus levels, and low organic matter. Cultivation in the region only became feasible by selecting plants and *Bradyrhizobium* strains with higher nitrogen fixation efficiency capable of tolerating the conditions of the Brazilian Cerrado biome.

Combining genetic knowledge with appropriate crop management techniques, Brazil became the largest soybean producer in the world. The cultivated area, which was 1.32 million hectares in 1970, grew to 47.4 million hectares in 2023, a 35.9-fold increase over 54 years. It is important to note that, thanks to nitrogen fixation through the soybean-*Bradyrhizobium* symbiosis, nitrogen fertilizers are practically unnecessary in soybean cultivation in Brazil.

The renowned genetics and plant breeding professor at ESALQ/USP, Ernesto Paterniani (1928–2009), often said in his classes that “plant selection is the greatest tool humanity has to adapt plants to human needs.” All available evidence supports this assertion. Mendel’s work played a pivotal role in making this statement a reality. Particularly in Brazil, numerous challenges have been and continue to be overcome. There is no shortage of examples, both in quantity and quality, of products developed to meet the demand for vegetables, fruits, and ornamental plants. Additionally, the production of new sugarcane cultivars has supported food production and biofuel ethanol. Equally significant is the development of new forage cultivars, which sustain the vast dairy and beef cattle herds. Lastly, cultivating selected clones, especially with *Eucalyptus* genus, for fiber production has fulfilled the demand for cellulose, charcoal, and raw wood, thereby minimizing the exploitation of native species.

REFERENCES

- Abelson JA (1980) A revolution in biology. *Science* **209**: 1319-1321.
- Ahmad M (2023) Plant breeding advancements with “CRISPR-Cas” genome editing technologies will assist future food security. *Frontiers Plant Science* **12**: 1-18.
- Andorf C, Beavis WD, Hufford M, Smith S, Suza WP, Wang K, Woodhouse M, Yu J and Lübberstedt T (2019) Technological advances in maize

- p>breeding: past, present and future.
- Theoretical and Applied Genetics**
- 132**
- : 817-849.
- Bellen HJ and Yamamoto S (2015) Morgan's legacy: Fruit flies and the functional annotation of conserved genes. **Cell** **163**: 12-14.
- Berlan JP (2021) Hybrid corn beyond heterosis: reading George Shull's hybrid corn articles (1908-1909). **Journal Genetics** **100**: 72.
- Bernardo R (2009) Genome wide selection for rapid introgression of exotic germplasm in maize. **Crop Science** **49**: 419-425.
- Bevan MW, Uauy C, Wulff BBH, Zhou J, Krasileva K and Clark MD (2017) Genomic innovation for crop improvement. **Nature** **543**: 346-354.
- Bezuidt OK, Pierneef R, Gomri AM, Adesioye F, Makhalanyane TP, Kharroub K and Cowan DA (2016) The geobacillus pan-genome: Implications for the evolution of the Genus. **Frontiers in Microbiology** **7**: 723.
- Bhat JA, Yu D, Bohra A, Ganie SA and Varshney RK (2021) Features and applications of haplotypes in crop breeding. **Communications Biology** **4**: 1266.
- Cannon EK, Portwood JL, Hayford RK, Haley OC, Gardiner JM, Andorf CM and Woodhouse MR (2024) Enhanced pan-genomic resources at the maize genetics and genomics database. **Genetics** **227**: iyae036.
- Chen NWG, Thareau V, Ribeiro T, Magdelenat G, Ashfield T, Innes RW, Pedrosa-Harand A and Geffroy V (2018) Common bean subtelomeres are hot spots of recombination and favor resistance gene evolution. **Frontiers in Plant Science** **9**: 1185.
- Cohen SN, Chang ACY, Boyer HW and Helling RB (1973) Construction of biologically functional bacterial plasmids in vitro. **Proceedings of the National Academy of Science USA** **70**: 3240-3244.
- Comstock RE and Robinson HF (1948) The components of genetic variance in population of biparental progenies and their use in estimating the average degree of dominance. **Biometrics** **4**: 254-266.
- Cortinovis G, Vincenzi L, Anderson R, Marturano G, Marsh JI, Bayer PE, Rocchetti L, Frascarelli G, Lanzavecchia G, Pieri A, Benazzo A, Bellucci E, Vittori VD, Nanni L, Fernández JJF, Rossato M, Aguilar OM, Morrell PL, Rodriguez M, Gioia T, Kerstin Neumann K, Diaz JCA, Gratias A, Klopp C, Bitocchi E, Geffroy V, Delledonne M, Edwards D and Papa R (2024) Adaptive gene loss in the common bean pan-genome during range expansion and domestication. **Nature Communications** **15**: 6698.
- Doudna JA and Charpentier E (2014) The new frontier of genome engineering with CRISPR-Cas9. **Science** **340**: 1258096.
- Fu H and Dooner HK (2002) Intrasppecific violation of genetic colinearity and its implications in maize. **Proceedings of the National Academy of Science USA** **99**: 9573-9578.
- Gerstein MB, Bruce C, Rozowsky JS, Zheng D, Du J, Korbel JO, Emanuelsson O, Zhang ZD, Weissman S and Snyder M (2007) What is a gene, post-ENCODE? History and updated definition. **Genome Research** **17**: 669-681.
- Griffiths AJF, Wessler SR, Carroll SB and Doebley J (2016) **Introdução à genética**. Guanabarra Koogn, Rio de Janeiro, 756p.
- Hallauer AR, Carena MJ and Miranda Filho JB (2010) **Quantitative genetics in maize breeding**. Springer, New York, 663p.
- Khan S, Anwar S, Yu S, Sun M, Yang Z and Gao ZQ (2019) Development of drought-tolerant transgenic wheat: Achievements and limitations. **International Journal of Molecular Sciences** **20**: 3350.
- Liu G, Qiu D, Lu Y, Wu Y, Han X, Jiao Y, Wang T, Yang J, You A, Chen J and Zhang Z (2023) Identification of superior haplotypes and haplotype combinations for grain size- and weight-related genes for breeding applications in rice (*Oryza sativa* L.). **Genes** **14**: 2201.
- Lötter A, Duong TA, Candotti J, Mizrahi E, Wegrzyn JL, Alexander A and Myburg AA (2023) Haplogenome assembly reveals structural variation in Eucalyptus interspecific hybrids. **GigaScience** **12**: 1-15.
- Ni L, Liu Y, Ma X, Lui T, Yang X, Wang Z, Liang Q, Liu S, Zhang M, Wang Z, Shen Y and Tian Z (2023) Pan-3D genome analysis reveals structural and functional differentiation of soybean genomes. **Genome Biology** **24**: 1-26.
- Nielsen RL (2023) **Historical corn grain yields in the U.S.**. Corny News Network, New York, p. 1-5.
- Pearson H (2006) What is a gene? **Nature** **441**: 398-401.
- Pray L and Zhaurova K (2008) Barbara McClintock and the discovery of jumping genes (transposons). **Nature Education** **1**: 169.
- Rai M and Tyagi W (2022) Haplotype breeding for unlocking and utilizing plant genomics data. **Frontiers in Genetics** **14**: 1-5.
- Ramalho MAP, Santos JB, Pinto CABP, Souza EA, Gonçalves FMA and Souza JC (2021) **Genética na agropecuária**. UFLA, Lavras, 508p.
- Rédei GP (2002) Vignette of the history of genetics. In Kang MS (ed) **Quantitative genetics, genomics and plant breeding**. CABI Publishing, New York, p. 1-20.
- SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melker-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z and Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of maize genome. **Science** **274**: 765-768.
- Shull GH (1948) What is "heterosis"? **Genetics** **33**: 439-446.
- Souza AP, Burgess SJ, Doran L, Hansen J, Manukan L, Maryn N, Dhananjay G, Leonelli L, Niyogi KK and Long SP (2020) Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. **Science** **377**: 851-854.
- Souza Júnior CL (1989) **Componentes de variância genética e suas implicações no melhoramento vegetal**. FEALQ, Piracicaba, 98p.
- Sun Z, Yuan L, Wang Y, Fang R, Lin X, Li H, Chen L, Wu Y, Huang X, Kong F, Liu B, Lu S and Kong L (2024) Post-flowering photoperiod sensitivity of soybean in pod-setting responses. **Biology** **13**: 1-20.
- Tettelin H, Masignani V and Cieslewicz MJ (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". **Proceedings of the National**

Academy of Science USA 102: 13950-13955.

Van Dijk PJ, Jessop AP and Ellis THN (2022) How did Mendel arrive at his discoveries? **Nature Genetics 54:** 926-933.

Varshney RK, Sinha P, Singh VK, Kumar A, Zhang Q and Bennetzen JL (2020) 5Gs for crop genetic improvement. **Current Opinion in Plant Biology 56:** 190-196.

Xin Y, Zhang Z, JHuang J and Luo L (2023) Yuan Longping, a great world

hunger fighter. **Plant Breeding 143:** 83-85.

Yang X, Lee WP, Ye K and Lee C (2019) One reference genome is not enough. **Genome Biology 20:** 1-3.

Zhai H, Wan Z, Jiao S, Zhou J, Xu K, Nan H, Liu Y, Xiong S, Fan R, Zhu J, Jiang W, Pang T, Luo X, Wu H, Yang G, Bai X, Kong F and Xia Z (2022) GmMDE genes bridge the maturity gene E1 and florigens in photoperiodic regulation of flowering in soybean. **Plant Physiology 189:** 1021-1036.