

Minireview Article

Genome Sequencing in Field Crops: Unlocking Agricultural Potential

Abstract

Genome sequencing has revolutionized agriculture by providing crucial insights into the genetic makeup of field crops. This paper explores the significance of genome sequencing in unlocking the agricultural potential of various field crops. By decoding the DNA of crops like wheat, maize, rice and soybeans, researchers gain a comprehensive understanding of their genetic diversity, disease resistance and yield-enhancing traits. This knowledge enables the development of precision breeding strategies, leading to the creation of high-yielding, stress-tolerant and nutritionally enhanced crop varieties. Moreover, genome sequencing facilitates the identification of key genes involved in plant-microbe interactions and adaptation to environmental stressors. Such insights can inform sustainable agricultural practices, reduce chemical inputs and enhance crop resilience in the face of climate change. This paper underscores the transformative impact of genome sequencing on crop improvement, food security and global agricultural sustainability.

Keyword: Genome, Sequencing, Diversity, Sustainable

Introduction

Genome sequencing, the process of determining the complete DNA sequence of an organism, has revolutionized numerous fields of study, including agriculture. In recent years, genome sequencing has emerged as a powerful tool in the field of crop improvement and has played a pivotal role in enhancing our understanding of field crops at the molecular level. The need for genetically enhanced, stress-resistant crops is becoming more urgent due to the rising global population and the potential that main food crops would produce less due to climate change. The discovery of genes driving significant agronomic features related to food production and quality has been hastened by the convergence of low-cost genome sequencing, increased computing capacity, and high-throughput molecular phenotyping methods. Here, we explore the history of plant breeding and how modern plant breeding techniques like site-directed nucleases and genomic studies are being used by researchers to improve the features of food crops. These products' deployment from the lab to the field is still hampered by biological and regulatory barriers that need to be resolved [1].

Comment [AS1]: Make-up

Comment [AS2]: Importance

Comment [AS3]: Sequencing

Comment [AS4]: Such as

Comment [AS5]: Add comma before and

Comment [AS6]: Without s

Comment [AS7]: Researchers are gaining ...

Comment [AS8]: Change to in addition

Comment [AS9]: Is facilitating

Comment [AS10]: Change to knowledge

Comment [AS11]: I suggest be increase

Comment [AS12]: highlights

Comment [AS13]: transformative

Comment [AS14]: add s

Comment [AS15]: add ity

Comment [AS16]: Change to become

Comment [AS17]: Delete

Comment [AS18]: Improving

Comment [AS19]: Stress-tolerant

Comment [AS20]: Change to increasingly

Comment [AS21]: Change to as

Comment [AS22]: World's population grows and ...

Comment [AS23]: Climate change threatens to reduce the production of major food crops.

Comment [AS24]: I suggests change to that control important

Comment [AS25]: Traits

Comment [AS26]: I suggest you rewrite it be active voice

Comment [AS27]: However, biological and regulatory hindrances impede the rollout of these products to the field and must be addressed.

Crop genomics continues to be a crucial component of the scientific advancement required to guarantee global food security. The initial sequencing of the rice genome was finished not long after the publication of the sequence of the first plant genome, that of Arabidopsis thaliana. Since then, the genomes of over 100 different crops have been sequenced, plant genome research has advanced on several fronts, and the upcoming years look to see significant advancements motivated by the introduction of new technology and methods. Crop genome sequencing, genetic mapping, and the collection of various levels of biological data are anticipated to continue to advance. There will be fascinating opportunities to combine genome-scale data from various biological scales, which will advance our mechanistic understanding of crop biological processes and increase the motivation for transferring laboratory findings to the field [2].

Comment [AS28]: Change to ensure

Comment [AS29]: Completed shortly

Comment [AS30]: sequencing

Comment [AS31]: more over

Comment [AS32]: crop genomes

Comment [AS33]: suggest a change to 'and advances have been made in plant genome research on several fronts. The introduction of new technologies and methods is expected to drive significant advancements in the coming years.

Plant biology research has changed significantly over the 20th century and will likely continue to do so. This has been made possible, in part, by the use of genomics techniques to identify the genetic makeup of several plant species, as well as to resolve population-level genome variations affecting thousands of individuals. Since the publishing of the first plant genome sequence, that of Arabidopsis thaliana, in 2000, genomics technology has significantly improved. For the purpose of creating genome, transcriptome, and epigenome databases for model and crop species that have allowed for in-depth conclusions about plant biology, plant genomics researchers have rapidly accepted new algorithms, technologies, and methodologies. Ploidy, heterozygosity, and paralogy are difficulties in sequencing any genome, but they are exacerbated in plant genomes compared to animal genomes due to their larger sizes, high number of repetitive sequences, and frequent whole- or segmental genome duplication. De novo transcriptome assemblies can be created, offering an alternate method to get around these intricate genomes and reach these resistant species' gene spaces. Technology advancements in sequencing platforms are driving the area of genomics, but software and algorithm development has lagged behind decreases in sequencing costs, increased throughput, and better quality. The length and caliber of output from sequencing technologies are expected to keep growing, and the complementing algorithms and bioinformatic tools required to manage huge, repetitive genomes are also expected to get better [3].

Comment [AS34]: I suggest "these methodologies also facilitate the resolution of population-level genome variations that affect numerous individuals.

Comment [AS35]: To generate

Comment [AS36]: Swiftly embraced novel

Comment [AS37]: Genome sequencing

Comment [AS38]: Which are

Comment [AS39]: abundant

Comment [AS40]: Delete

Comment [AS41]: However, the growth in sequencing speed, quality, and affordability has not been matched by parallel software and algorithmic developments.

Comment [AS42]: Delete

Unraveling Crop Genomes

Genome sequencing allows scientists to decode the entire genetic blueprint of a field crop, revealing its DNA sequence, genetic variations and functional elements. This wealth of genetic information provides invaluable insights into the crop's genetic diversity, traits and

potential for improvement. ~~By~~ studying the genome, researchers can identify and analyze genes responsible for desirable traits such as disease resistance, yield potential and nutritional quality.

Comment [AS43]: Through

Enhancing Crop Breeding Programs

Genome sequencing accelerates traditional crop breeding methods by enabling the identification of genes associated with desired traits. This information facilitates marker-assisted selection (MAS), a technique where ~~specific~~ genetic markers linked to favorable traits are used to ~~select and breed~~ plants with improved characteristics. MAS streamlines the breeding process, making it more efficient and precise, ultimately leading to the development of high-yielding and resilient crop varieties.

Comment [AS44]: Delete

Comment [AS45]: Enhanced traits

Comment [AS46]: Resilient and high-yielding crop

Marker-Assisted Selection

Molecular markers are complementary tools to ~~traditional~~ selection, used to select parental genotypes in breeding programs, eliminate linkage drag in back-crossing and select for traits that are difficult to measure using phenotypic assays. They can increase our understanding of phenotypic characteristics and their genetic association, which may modify the breeding strategy. MAS ~~allows~~ the breeder to achieve early selection of a trait in a breeding program and it is particularly useful when the trait is under complex genetic control, or when phenotypic trials are unreliable or expensive. With the development of molecular techniques, MAS ~~is now used to~~ enhance traditional breeding programs to improve crops and modern plant breeding is dependent on molecular markers for the rapid and precise analysis of germplasm and ~~for~~ trait mapping [4].

Comment [AS47]: Change to conventional

Comment [AS48]: enables

Comment [AS49]: has enhanced

Comment [AS50]: Delete

Application of Marker Assistant Selection

The elite Indian rice variety, Naveen is highly susceptible to major biotic and abiotic stresses such as blast, bacterial blight (BB), gall midge (GM) and drought which limit its productivity in rainfed areas. In the current study, three major genes—Pi9 for blast, Xa21 for bacterial blight (BB), and Gm8 for gall midge (GM)—as well as three major QTLs—qDTY1.1, qDTY2.2, and qDTY4.1—that confer increased yield under drought in the Naveen background—were introgressed. Gene-based/linked markers were employed for the foreground selection of biotic and abiotic stress-tolerant genes/QTLs at each step of progression. To find lines with a high level of resistance to blast, BB, GM, and drought tolerance without yield penalty in a non-stress condition, intensive phenotype-based selections were carried out in the field. In order to improve yield performance compared to Naveen, a collection of 8 MAFB lines and 12 MABC lines with 3 to 6 genes/QTLs and resistance/tolerance to biotic stressors as well as reproductive stage drought stress were

Comment [AS51]: Blast resistance (?)

established. Lines created using both MAFB and MABC performed better than those created using simply MAFB. This work serves as an excellent illustration of the value of a combined method to forward and backcross breeding with marker assistance for improving numerous biotic and abiotic stress tolerance in the background of well-liked mega varieties[5].

One of the key factors limiting rice yield globally is drought. Improved drought resilience is a result of the robust shoot and extensive root system. Genome-wide association studies (GWAS) are the method of choice nowadays for identifying QTLs for complex characteristics like drought and root tolerance. In the current study, 114 rice genotypes were tested under water stress circumstances for a variety of root and shoot properties. Different root and shoot characteristics varied significantly across all genotypes. According to a correlation study, root length, root volume, fresh root weight, and dry root weight are all related to high dry shoot weight and fresh shoot weight. For diverse root, shoot, and drought tolerance characteristics, a total of 11 significant marker-trait relationships were found, with the coefficient of determination (R^2) ranging from 18.99 to 53.41%. Markers RM252 and RM212 revealed associations with three root characteristics, indicating potential for root system enhancement. In the current study, RM127 was linked to a unique QTL for root length that accounted for 19.30% of variance. By utilising marker-assisted selection, it is possible to increase root and drought resistance qualities by taking advantage of marker alleles with increasing phenotypic impacts[6].

Genetic Mapping

Molecular markers have revolutionized genome mapping over the last two decades and the high density of markers that can now be generated from second-generation sequence data offers the potential for generating very high-density genetic maps. These markers can be used to develop haplotypes for genes or regions of interest and complete genome mapping is now becoming a reality. Genetic mapping places molecular genetic markers in linkage groups based on their co-segregation in a population. The genetic map predicts the linear arrangement of markers on a chromosome and maps are prepared by analyzing populations derived from crosses of genetically diverse parents and estimating the recombination frequency between genetic loci. Many types of markers can be used for map construction, with population size and marker density being important for map resolution [7].

The sequential assignment of loci to a particular region on a chromosome is known as gene mapping. Species-specific genetic maps are made up of genes, genomic markers, and/or genomic distances between each marker. These distances are not determined by the actual

Comment [AS52]: Delete

Comment [AS53]: Delete

Comment [AS54]: features

Comment [AS55]: delete

Comment [AS56]: Delete

Comment [AS57]: Delete

Comment [AS58]: Add an article (the)

Comment [AS59]: I suggest change to "the use of marker-assisted selection may take advantage of marker alleles with increasing phenotypic impacts and increase root and drought resistance qualities"

Comment [AS60]: delete

Comment [AS61]: has

Comment [AS62]: change to allowing for

Comment [AS63]: Delete

Comment [AS64]: I suggest change to "the linear arrangement of markers on a chromosome is predicted by the genetic map. Maps are prepared ..."

Comment [AS65]: Delete

position of the chromosomes, but rather by the number of chromosomal crossings that take place during meiosis. Dense genetic marker maps for humans are already available, and the development of next-generation sequencing technology has accelerated the creation of genetic maps for other species. Linkage mapping, a technique used to map disease genes or trait loci, requires the use of genetic maps. Genetic study has found that combining disease gene mapping and genetic mapping with next-generation sequencing is an effective approach[8].

Comment [AS66]: Add an article (The)

Haplotype-based Breeding (HBB) Advances in plant genome biology have inspired innovative approaches to expedite the progress of assembling desirable phenotype in cropbreeding programs (figure 1).

Comment [AS67]: Add 's'

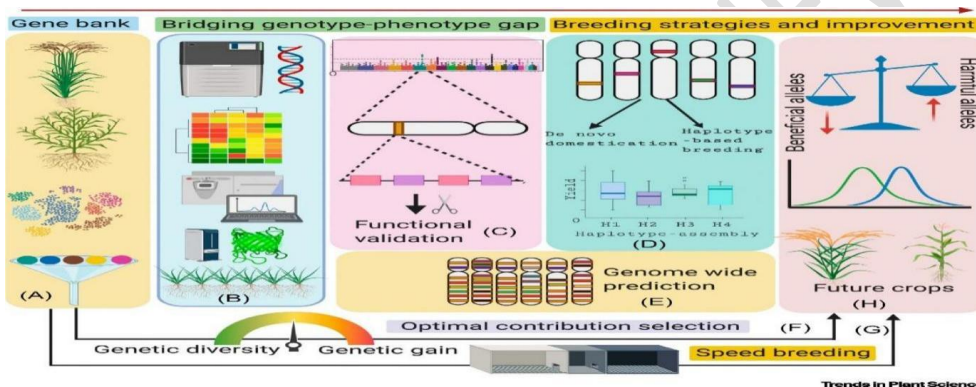


Figure 1. An Overview of GAB 2.0, Genomics-Assisted Breeding's Second-Generation Approach to Developing Future Crops. In order to create future crops, the image depicts a comprehensive strategy that tries to build up advantageous alleles or eliminate harmful alleles in plant genomes. (A) Both superior (useful) and detrimental impact (destructive) alleles can be found in gene bank collections of germplasm. (B) Field phenotyping and multi-omics tests, along with high-throughput sequencing, offer a potent way to link genetic changes with significant phenotypes. (C) The discovery of a causal gene follows functional confirmation of a gene-trait connection. (D) Knowledge of the genes responsible for important plant properties opens the door to haplotype-based breeding, genomic breeding, or de novo domestication. (E) Breeding program decisions can also be made using a genomic prediction technique based on data from genome-wide genotyping. (F) For pre-breeding and breeding purposes, techniques like optimum contribution selection (OCS) that maintain a balance between the rate of genetic increase and genetic diversity/inbreeding will be essential. (G) Speed breeding will hasten the development of crop breeding. (H) Using these novel breeding

tools and methods will enhance the genetic advantages of breeding programs by allowing breeding populations to accumulate advantageous alleles or eliminate deleterious alleles. Future crop designs will benefit from this breeding approach [9].

Association Studies

Association mapping is a ~~further~~ statistical method ~~to identify genetic loci~~ associated with phenotypic trait variation. Association mapping shares much in common with quantitative trait loci (QTL) mapping. QTL mapping generally involves the use of structured populations and relatively distant markers can segregate with the QTL, providing a wide genetic region within which the gene is located. The use of unstructured populations in association mapping means that they represent many more recombination events and are often many generations from a common ancestor, providing the potential of a greater resolution for a set population size. The advances in genome sequencing technology, allowing the production of millions of markers, provide an increasing ability to generate large quantities of molecular marker genotyping data, which favors association studies over traditional QTL mapping, because of this, association studies are likely to become more common.

Comment [AS68]: Change to Used to identify genetic loci that are associated with variations in phenotypic traits.

Uncovering Genetic Variation

Genome sequencing ~~allows~~ the identification and characterization of genetic variations within field crop populations. This information is crucial for understanding the genetic diversity of crops and developing strategies for ~~crop improvement~~. ~~By studying genetic variations,~~ breeders can select plants with diverse genetic backgrounds to enhance resilience, adaptability and tolerance to environmental stresses such as drought, pests and diseases.

Comment [AS69]: facilitates

Comment [AS70]: for their improvement, enables breeders ..

Numerous uncommon genetic variations linked to Mendelian illnesses have been quickly identified thanks to developments in sequencing technology, cohort formation, and data dissemination[10, 11]. This success is mostly due to their genetic designs, which are very straightforward; Mendelian disorders are typically brought on by harmful alleles that are concentrated within a few ~~number~~ of genomic sites. ~~Nevertheless, among diagnosed instances,~~ clinical heterogeneity is frequently seen[10, 12, 13]. For instance, Marfan Syndrome, an autosomal dominant illness brought on by mutations in the FBN1 gene, is linked to abnormalities in the heart, eyes, skeleton, and ~~even the~~ lungs. Rarely do carriers of pathogenic FBN1 alleles experience all of the accompanying symptoms[13], and even members of the same family can exhibit different phenotype [15]. Allelic heterogeneity ~~accounts for~~ some of the clinical variation seen among Mendelian illness cases [10,

Comment [AS71]: Change to a plural noun. Add 's:

Comment [AS72]: However

Comment [AS73]: Delete

Comment [AS74]: Delete

Comment [AS75]: Contributes to

12]although other lines of evidence also point to a role for environmental and genetic background effects [13, 16, 17, 18,19, 20].

The low prevalence of these disorders places intrinsic constraints on the discovery of specific factors that alter the severity of Mendelian diseases. Construction of cohorts of afflicted patients large enough to uncover genetic and environmental modifiers is typically challenging (but not impossible [21, 22]), especially if they have very small effect sizes. This constraint has led to the use of model organisms [23, 24] or the incorporation of orthogonal analyses [25,26] in many research that examine modifier effects. As an alternative, some Mendelian illnesses may reflect the severely afflicted end of a spectrum of pathologic variation, according to our hypothesis and that of others. This link is well established for diseases including familial hypercholesterolemia [27], hereditary breast cancer [28], and long QT syndrome [29], and new analyses of huge bio-bank datasets have allowed researchers to investigate the interaction between uncommon pathogenic variation and common polymorphisms [30,31]. The condition of interest in these samples, however, translated to a univariate, quantitative phenotype, making the analyses conceivable. Investigating the interaction between common and unusual genetic variation becomes significantly more challenging for Mendelian illnesses that instead map to high-dimensional arrays of diverse symptoms.

Uniting Genomics and Phenomics

The integration of genome sequencing with phenomics, the study of an organism's physical and biochemical traits, is transforming crop improvement efforts. By combining genomic information with high-throughput phenotyping technologies, researchers can correlate genetic variations with observable traits, facilitating the identification of key genes underlying complex traits. This knowledge can guide the development of targeted breeding strategies and the design of precision agriculture practices.

Low narrow-sense heritability of economically significant traits, a long breeding cycle, practical challenges for extensive phenotyping of breeding populations, a large, complex polyploid genome with high heterozygosity, and genotype-environment-management interaction effects have all been linked to low genetic gain rates. More precisely, because of the huge biomass of sugarcane plants, good comprehensive phenotyping is logistically very difficult, which reduces the precision of selection. This is especially true in the early phases of selection, which are complicated by the strong effects of interplot competition brought on by tiny single- or double-row plots. As a result, a key obstacle to

Comment [AS76]: Consequently,

Comment [AS77]: Delete

Comment [AS78]: I suggest a change to “for mendelian illnesses that are instead mapped to high-dimensional arrays of diverse symptoms, the investigation of the interaction between common and unusual genetic variation becomes considerably more challenging”.

Comment [AS79]: Delete

Comment [AS80]: Delete

expediting sugarcane development, accurate, cost-effective, and high-throughput phenotyping presents an outstanding opportunity for more accurate evaluation of the genuine yield potential of sugarcane clones in breeding experiments [32].

Accelerating Crop Domestication

Genome sequencing plays a vital role in studying the genetic history and domestication of field crops. ~~By comparing the genomes of wild relatives with domesticated varieties,~~ researchers can identify genomic regions associated with key domestication traits, such as increased ~~seed~~ size or loss of seed shattering. This knowledge can be utilized to accelerate crop domestication ~~efforts,~~ ~~leading~~ to the development of improved crop varieties with enhanced agronomic traits.

Comment [AS81]: Delete. In my opinion, it's unnecessary

Comment [AS82]: Delete

Comment [AS83]: Add "thus leading to .."

Wild plant species were tamed and farmed near their origins when human civilization was established, and they were then spread to other regions of the world. Numerous mutations happened throughout the years as they were being raised in the wild and on farms, adding additional variants to their DNA. The enhanced agricultural plants of today are the product of decades of artificial selection for a small number of those mutations, sometimes in conjunction with conscious selection for desirable recombinants that emerged naturally or were created by selective breeding. Crop species' genomes bear the marks of artificial phenotypic selection. Researchers have discovered ~~a number of~~ genes and causative mutations linked to domestication events during the past three decades, providing a clearer knowledge of how our ancestors and foremothers modified plant growth to fulfil their demands for food and fodder. Our capacity to use effective genome editing techniques to script complicated genetic information has allowed us to make considerable strides in speeding up crop domestication. The advantages of using genome editing technologies to domesticate wild and semi-domesticated species, the requirements for altering wild genomes, and potential future target loci for quickly introducing domestication syndrome in wild plant species have all been covered in this study. In order to sustainably satisfy our present and future demands, genome editing technologies may enable us to introduce wild and partially tamed agricultural species into conventional agriculture[33].

Comment [AS84]: For clarity, I suggest you use several

Genomic-Assisted Crop Protection

Genome sequencing aids in understanding the genetic basis of plant diseases and pests, enabling the development of effective control strategies. By studying crop genomes, scientists can identify resistance genes and markers associated with disease resistance, facilitating the breeding of resistant crop varieties. Additionally, genome sequencing

enhances our understanding of the genetic interactions between crops and pathogens, allowing for targeted approaches in pest and disease management. The conventional breeding of crops fails to keep up with the rising demand for food and the pests and diseases ~~constant adaptation~~. The difficulty of finding superior crop cultivars has increased as a result of the additional complexity that global climatic changes have put on biological systems. ~~These mandate~~ the creation and use of cutting-edge technologies, such as genome editing (GE), that provide focused and quick breeding ~~programmes~~ in crops with improved resistance to viruses and pests. With no need for crosses, GE expedites the entire breeding process by preventing the introduction of ~~unwanted~~ characteristics through linkage in apex varieties. ~~The direct targeting of plant susceptibility (S) genes or virulence factors of pests and pathogens using GE technologies can also enhance plant protection.~~ This can be ~~done~~ by directly editing the genome of the pest or by integrating the GE machinery into the genome of the plant or microorganisms acting as biocontrol agents (BCAs). With the introduction of CRISPR/Cas, GE technology has advanced much further throughout the years. Here, we examine the most recent developments in GE plant protection, with a particular emphasis on CRISPR/Cas-based genome editing of crops, pests, and diseases. We address how CRISPR/Cas might work in conjunction with other technologies, like as host-induced gene silencing (HIGS) and the use of BCAs, to hasten the creation of green policies that would support sustainable agriculture in the future[34].

Comment [AS85]: Delete

Comment [AS86]: Consequently,

Comment [AS87]: Delete

Comment [AS88]: Change the spelling: "programs"

Comment [AS89]: Undesirable

Comment [AS90]: I suggest a change to "Additionally, GE technologies can improve plant protection by directly targeting plant susceptibility (S) genes or virulence factors of pests and pathogens."

Comment [AS91]: Achieved

Transposon Insertion Sequencing Application

Numerous plant-associated bacteria can ~~influence plant development in a favorable way~~, and there is rising interest in using these bacteria in agricultural settings to lessen the need for fertilizers and pesticides. Our ability to use microorganisms in this way, however, is now constrained by our incomplete understanding of the molecular mechanisms underlying bacterial-plant interactions. The vast majority of bacterial genes' activities are either unknown or poorly understood since traditional techniques of researching molecular interactions have worked slowly to characterise one gene at a time. Efforts to optimise microbial communities and generate microbe-based products will be facilitated by new methods to enhance and expedite studies into the activities of bacterial genes in agricultural systems. It is quite likely that methods for high-throughput gene functional analysis, including transposon insertion sequencing studies, will be used extensively to identify important elements of plant-bacterial interactions. Transposon insertion sequencing is a technique that combines high-throughput sequencing with saturation transposon mutagenesis to concurrently examine the function of

Comment [AS92]: Favorably influence plant development

every non-essential gene in a bacterial genome. This method may be applied for both in vitro and in vivo investigations to find the genes responsible for pathogen virulence, microbe-plant interactions, and stress tolerance. The knowledge gained from such research would significantly speed up the process of determining the function of bacterial genes and offer insights into the genes and pathways that are responsible for the biotic interactions, metabolism, and survival of bacteria that are important for agriculture. This information might be used to create crop inoculants, create crop protection products, or choose the best bacteria for boosting plant development under a particular set of circumstances. This paper introduces transposon insertion sequencing, describes how it has been used to investigate microbes linked with plants, and suggests new agricultural uses for these methods[35].

Conclusion

Genome sequencing has transformed the field of crop improvement, empowering scientists and breeders with a wealth of genetic information to **develop high-yielding**, disease-resistant and climate-resilient field crops. The application of genomic technologies in agriculture holds great promise for addressing global food security challenges by accelerating crop improvement efforts and ensuring sustainable agricultural practices. As genome sequencing technologies continue to advance, their integration into crop breeding programs will play an increasingly critical role in shaping the future of agriculture.

Comment [AS93]: I suggest a change to “create crops that generate a high yield”

References

1. Steinwand, M.A., Ronald, P.C. Crop biotechnology and the future of food. *Nat Food* 2020, **1**, 273–283. <https://doi.org/10.1038/s43016-020-0072-3>.
2. Purugganan, M.D., Jackson, S.A., Advancing crop genomics from lab to field. *Nat Genet.* 2021, **53**, 595–601. <https://doi.org/10.1038/s41588-021-00866-3>.
3. Hamilton JP and Buell CB, Advances in plant genome sequencing. *The Plant Journal*, 2012, **70**, 177–190.
4. Koebner, R., Summers, R., The impact of molecular markers on the wheat breeding paradigm. *Cell Mol Biol Lett.* 2002, *c7(2B)*, 695-702.
5. Janaki Ramayya P, Vinukonda VP, Singh UM, Alam S, Venkateshwarlu C, Vipparla AK, *et al.* Marker-assisted forward and backcross breeding for improvement of elite Indian rice variety Naveen for multiple biotic and abiotic stress tolerance. *PLoS ONE*, 2021, **16(9)**:e0256721. <https://doi.org/10.1371/journal.pone.0256721>.

6. Verma, H., Sarma, R.N. Identification of Markers for Root Traits Related to Drought Tolerance Using Traditional Rice Germplasm. *Mol Biotechnol* 2021, **63**, 1280–1292. <https://doi.org/10.1007/s12033-021-00380-1>.
7. Grover, A., Sharma, P.C., Development and use of molecular markers: past and present. *Crit. Rev. Biotechnol.* 2016;**36**(2), 290-302. doi: 10.3109/07388551.2014.959891.
8. Kelly L. Williams, Gene Mapping. *Encyclopedia of Bioinformatics and Computational Biology*, 2019, **2**, 242-250. <https://doi.org/10.1016/B978-0-12-809633-8.20233-1>.
9. Varshney, R.K., Bohra, A., Yu, J., Graner, A., Zhang, Q., and Sorrells, M.E., Designing Future Crops: Genomics-Assisted Breeding Comes of Age. *Trends in Plant Sci.* 2021, **26**(6), 631-649.
10. Chong JX,*et al.* The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am. J. Hum. Genet.*, 2015, 97, 199–215.
11. Landrum MJ,*et al.* ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.*2018, 46, D1062–D1067.
12. Wright CF, FitzPatrick DR & Firth HV, Paediatric genomics: diagnosing rare disease in children. *Nat. Rev. Genet.*2018, 19, 253–268.
13. Posey JE,*et al.* Insights into genetics, human biology and disease gleaned from family based genomic studies. *Genet. Med. J. Am. Coll. Med. Genet.* 2019, 21, 798–812.
14. Wenger BM,*et al.* A genotype-first approach to exploring Mendelian cardiovascular traits with clear external manifestations. *Genet. Med. J. Am. Coll. Med. Genet.* 2021, 23, 94–102 .
15. Akhurst RJ. TGF β signaling in health and disease. *Nat. Genet.* 2004, 36, 790–792 (2004).
16. Chen R,*et al.* Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases. *Nat. Biotechnol.* 2016, 34, 531–538.
17. Tarailo-Graovac M, Zhu JYA, Matthews A, van Karnebeek CDM & Wasserman WW. Assessment of the ExAC data set for the presence of individuals with pathogenic genotypes implicated in severe Mendelian pediatric disorders. *Genet. Med. J. Am. Coll. Med. Genet.* 2017, 19, 1300–1308.
18. Bastarache L, *et al.* Phenotype risk scores identify patients with unrecognized Mendelian disease patterns. *Science*, 2018, 359, 1233–1239.

19. Rahit KMTH & Tarailo-Graovac M. Genetic modifiers and rare Mendelian disease. *Genes*, 2020, 11, 239.
20. Grange T, et al. Quantifying the genetic basis of Marfan syndrome clinical variability. *Genes*, 2020, 11, 574.
21. Corvol H, et al. Genome-wide association meta-analysis identifies five modifier loci of lung disease severity in cystic fibrosis. *Nat. Commun.* 2015, 6, 8382 (2015).
22. Pottier C, et al. Potential genetic modifiers of disease risk and age at onset in patients with frontotemporal lobar degeneration and GRN mutations: a genome-wide association study. *Lancet Neurol.* 2018, 17, 548–558.
23. Bonyadi, M. et al. Mapping of a major genetic modifier of embryonic lethality in TGF beta 1 knockout mice. *Nat. Genet.* 15, 207–211 (1997).
24. Zhang S, Binari R, Zhou R & Perrimon N. A genomewide RNA interference screen for modifiers of aggregates formation by mutant Huntingtin in *Drosophila*. *Genetics* 2010. 184, 1165–1179.
25. Aubart M, et al. Association of modifiers and other genetic factors explain Marfan syndrome clinical variability. *Eur. J. Hum. Genet.* 2018, 26, 1759–1772 (2018).
26. Pemov A, et al. Genetic modifiers of neurofibromatosis type 1-associated Café-au-Lait Macule count identified using multi-platform analysis. *PLoS Genet.* 2014, 10, e1004575.
27. Sturm, A. C. et al. Clinical genetic testing for familial hypercholesterolemia: JACC scientific expert panel. *J. Am. Coll. Cardiol.* 2018, 72, 662–680.
28. Hindorf LA, Gillanders EM & Manolio TA. Genetic architecture of cancer and other complex diseases: lessons learned and future directions. *Carcinogenesis* 2011, 32, 945–954.
29. Ingles, J. & Semsarian, C. Time to rethink the genetic architecture of long QT syndrome. *Circulation* 2020, 141, 440–443.
30. Nomura A, et al. Polygenic risk scores for low-density lipoprotein cholesterol and familial hypercholesterolemia. *J. Hum. Genet.* 2021, 66, 1079–1087.
31. Goodrich JK, et al. Determinants of penetrance and variable expressivity in monogenic metabolic conditions across 77,184 exomes. *Nat. Commun.* 2021. 12, 3505.
32. Natarajan S, Basnayake J, Wei X, Lakshmanan P. High-throughput phenotyping of indirect traits for early-stage selection in sugarcane breeding. *Remote Sensing.* 2019;11(24): 2952.

33. Pattnaik, D., Avinash, S.P., Panda, S. Accelerating crop domestication through genome editing for sustainable agriculture. *J. Plant Biochem. Biotechnol.* 2023. <https://doi.org/10.1007/s13562-023-00837-1>
34. Rato, C., Carvalho, M.F., Azevedo, C. *et al.* Genome editing for resistance against plant pests and pathogens. *Transgenic Res.* 2021, **30**, 427–459.
35. Fabian, B.K., Tetu, S.G., Paulsen, I.T., Application of Transposon Insertion Sequencing to Agricultural Science. *Frontiers in Plant Science*, 2020, 11, 1664-462X.

UNDER PEER REVIEW