Scientia Review, Vol. 02(09), September 2026, pp. 61-66 ISSN: 3107-7919

Scientia Review

Website: www.scientiafoundation.in

POPULAR SCIENCE ARTICLE

From Sequence to Seed: How Functional Genomics Turns Crop Genes into Better Varieties

Manish Kumar[1*], Rudra Pratap Singh[1], Md Ishtiaque[1], Satish Kumar[2], Ashutosh Kumar[2]
[1]Department of Molecular Biology and Biotechnology, Bihar Agricultural University, Sabour, Bhagalpur, Bihar 813210, India
[2]Department of Genetics and Plant Breeding, Bihar Agricultural University, Sabour, Bhagalpur, Bihar 813210, India
[*]Email: manishksharma003@gmail.com

Received: 12 September 2026
Revised: 13 September 2026
Accepted: 16 September 2026
Published online: 17 September 2026

Article ID: SR01162

Citation: Citation: Kumar, M., Singh, R. P., Ishtiaque, M., Kumar, S., & Kumar, A. (2026). From Sequence to Seed: How Functional Genomics Turns Crop Genes into Better Varieties. Scientia Review, 2(9), 61-66

Abstract

Crop breeding has entered a phase in which genome sequence is no longer the limiting resource. The harder question is what individual genes actually do in a plant growing in a field. Harder still is whether changing them improves performance. This article describes how functional genomics answers that question in crops. It follows the path from genetic variation and candidate gene discovery, through expression analysis and mutant-based validation, to genome editing and molecular breeding. Examples are drawn from rice, wheat and tomato. They include genes controlling grain size, tillering, submergence tolerance, disease resistance and fruit quality. The article also examines how validated gene function feeds into precision breeding and how promoter editing generates graded trait variation. It then asks what the idea of designer crops means in practice. Present technical, biological and regulatory constraints are discussed alongside emerging directions such as pan-genomes, single-cell transcriptomics and refined editing chemistries.

Keywords: Crop, gene, genetic, genome, disease, tolerance, tillering

Introduction

Plant breeders have always worked with genes, even when they could not see them. For most of the twentieth century, selection was based on what the plant looked like in the field. That approach produced the semi-dwarf cereals of the Green Revolution and a long list of improved varieties, but it was slow and largely blind to the molecular basis of the traits being selected.

Sequencing changed the starting point. The arrival of next-generation sequencing made genome-wide variation discovery routine and affordable in crops (Varshney *et al.*, 2009). Genome assemblies, resequencing panels and dense marker sets are now available for all major crops and for a growing number of minor ones. The bottleneck has shifted. We no longer struggle to find sequence variation. We struggle to know which of the millions of variants actually change a trait and by how much.

Functional genomics addresses this gap. It asks what genes do in a living plant, under real growing conditions and how altering them changes performance. This article follows that path from gene discovery through functional validation to precision breeding and the emerging idea of designer crops.

What functional genomics means for a breeder

Structural genomics describes the parts list of a genome: sequence, gene models, repeats and structural variants. Functional genomics asks what those parts do. In a crop context the question is more specific. Does this gene influence grain number, panicle architecture, oil composition, nitrogen uptake or resistance to a particular pathogen? And is the effect large enough and stable enough to be useful in breeding?

Three features distinguish the crop version of this discipline. First, the endpoint is an agronomic trait measured in a field, not a molecular phenotype in a growth chamber. Second, most target traits are quantitative, controlled by many loci of modest effect. Third, the useful outcome is often a specific allele rather than simply knowledge that a gene exists. A gene becomes a breeding asset only when a defined allele with a predictable effect can be moved into elite backgrounds. The overall workflow, from genetic variation to deployed cultivar, is summarized in Figure 1.

From gene discovery to gene function

Gene discovery in crops usually begins with a mapping population or a diversity panel.

Classical bi-parental quantitative trait locus (QTL) mapping localises a trait to a chromosomal interval. Fine mapping then narrows the interval until a plausible candidate emerges. Rice provided several textbook cases. *GS3* was identified as a major QTL for grain length and weight and as a minor QTL for grain width and thickness. It encodes a putative transmembrane protein (Fan *et al.*, 2006). *MONOCULM 1* (*MOC1*) was isolated from a rice mutant with a single main culm. It encodes a GRAS (GAI, RGA, SCR) family transcription factor that controls axillary bud formation and therefore tillering (Li *et al.*, 2003).

Association mapping widened the search. A genome-wide association study (GWAS) of 14 agronomic traits across 517 rice landraces showed that population-scale resequencing could recover known loci and point to new ones, without constructing a dedicated cross (Huang *et al.*, 2010). Such studies work best under three conditions. The trait must be variable in the germplasm being scanned. Population structure must be properly modelled. Marker density must be sufficient to approach causal variants.

Association signals are correlations, not proof of function. That distinction matters. A significant marker may sit thousands of base pairs from the causal polymorphism or may tag a haplotype block containing several plausible genes. Independent functional evidence is required before a candidate is treated as validated.

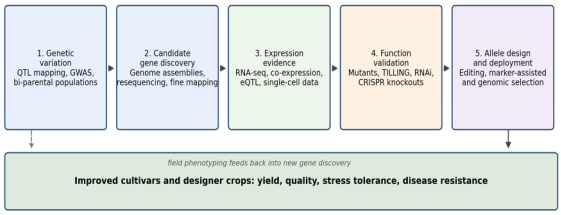

Fig 1. The functional genomics pipeline in crop improvement. Genetic variation is surveyed by QTL mapping and association studies, candidate genes are identified from assemblies and fine mapping, expression evidence narrows the list, function is confirmed using mutants and genome editing and validated alleles are deployed through editing and marker-based or genomic selection. Field phenotyping of the resulting material returns new questions to the discovery stage.

The main experimental tools

Transcriptomics remains the most widely used entry point. RNA sequencing (RNA-seq) measures which genes respond to drought, cold, pathogen attack or a developmental transition. Co-expression analysis then groups genes into modules that often share regulators. Single-cell and spatial transcriptomics extend this resolution to individual cell types and tissue positions. That is valuable for traits governed by a small number of cells, such as meristem size or vascular development. In plants these methods are still limited by protoplast isolation efficiency, cell wall interference and incomplete reference atlases for most crops (Chen *et al.*, 2023).

Forward genetics starts from a phenotype and works towards the gene. Reverse genetics starts from a sequence and asks what happens when it is disrupted. Targeting Induced Local Lesions IN Genomes (TILLING) made the reverse approach practical in crops. It screens chemically mutagenised populations for point mutations in a chosen gene (McCallum *et al.*, 2000). It yields non-transgenic allelic series, including weak alleles that can be more useful than a full knockout. Polyploidy complicates the picture, because redundant homoeologues buffer single-gene mutations. That is why polyploid-adapted TILLING platforms were developed for wheat (Uauy *et al.*, 2009).

Clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated (Cas) systems have made validation faster and more direct. A knockout can now be generated in a target cultivar rather than in a model genotype. This removes much of the ambiguity of extrapolating between backgrounds. In hexaploid wheat, simultaneous editing of all three *TaMLO* homoeoalleles produced heritable resistance to powdery mildew, confirming the conserved susceptibility role of Mildew resistance locus O (*MLO*) (Wang *et al.*, 2014). Base editing and prime editing add nucleotide-level precision. They allow specific base substitutions and in the case of prime editing short insertions and deletions, without a donor template or homologous recombination (Shimatani *et al.*, 2017; Lin *et al.*, 2020). A comparison of the major approaches, including what each can and cannot establish, is given in Table 1.

Table 1. Major functional genomics approaches used in crop improvement and the type of evidence each provides.

Approach	What it measures	Strength	Main limitation
QTL mapping and GWAS	Association between markers and trait variation	Genome-wide; uses natural or induced variation	Gives intervals or correlations, not gene function (Huang *et al.*, 2010)
RNA sequencing and co-	Transcript abundance across	Rapid prioritisation of	Expression change does not prove

expression	tissues, stages and treatments	candidates and regulatory modules	causality
Single-cell and spatial transcriptomics	Expression at cell-type or tissue-position resolution	Resolves traits driven by few cells	Technically demanding in plants; few crop atlases (Chen *et al.*, 2023)
TILLING and mutant populations	Phenotypic effect of induced point mutations	Non-transgenic allelic series, including weak alleles	Background mutations; redundancy in polyploids (McCallum *et al.*, 2000)
RNA interference	Phenotype after transcript knockdown	Useful for essential genes and gene families	Partial and sometimes variable silencing
CRISPR knockout	Phenotype of a defined loss-of-function allele	Direct causal test in the target cultivar	Transformation and regeneration limit many crops
Base and prime editing	Effect of specific nucleotide changes	Precise substitutions without donor DNA	Efficiency depends on target and delivery (Lin *et al.*, 2020)
Promoter and cis-regulatory editing	Quantitative shifts in gene expression	Generates graded trait variation	Requires knowledge of regulatory architecture (Rodríguez-Leal *et al.*, 2017)

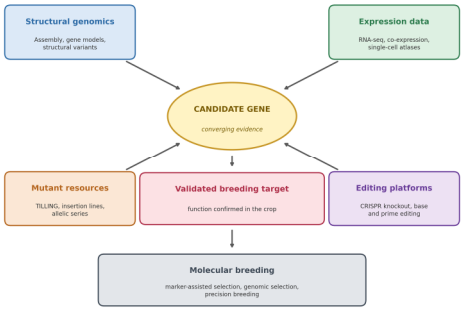

Fig 2. Integration of functional genomics technologies around a candidate gene. Structural genomics, expression data, mutant resources, editing platforms and downstream molecular breeding each contribute a different type of evidence. Convergence of several independent lines of evidence is what converts a candidate into a validated breeding target.

Genes behind real agronomic traits

Functional genomics has already delivered genes that breeders use. *Sub1A* is an ethylene-response-factor-like gene from the Indian landrace FR13A. It confers tolerance of complete submergence in rice and its introgression is associated with restrained shoot elongation and reduced carbohydrate consumption during flooding (Xu *et al.*, 2006). Its identification allowed marker-assisted introgression of the *SUB1* locus into widely grown mega-varieties, one of the most frequently cited successes of MAS in a public breeding programme (Kumar *et al.*, 2025). This is an example of a single cloned gene changing management options for farmers in flood-prone lowlands.

Disease resistance illustrates a second logic. Instead of adding a resistance gene, susceptibility genes can be removed. Targeted mutagenesis of the ERF transcription factor gene *OsERF922* increased blast resistance in rice. The edited lines showed no significant penalty in the agronomic traits measured in that study (Wang *et al.*, 2016). The *MLO* work in wheat follows the same principle (Wang *et al.*, 2014).

Quality traits are also tractable, because they often depend on defined biosynthetic steps. Multiplexed editing of five genes in the gamma-aminobutyric acid (GABA) pathway raised GABA content in tomato fruit (Li *et al.*, 2018). Selected genes and their trait associations are listed in Table 2.

Table 2. Examples of functionally characterized genes and the crop traits they influence.

Gene	Crop	Trait	Nature of evidence	Reference
GS3	Rice	Grain length and weight	Fine mapping and cloning of a major QTL	Fan *et al.* (2006)
MOC1	Rice	Tillering and axillary bud outgrowth	Mutant analysis and gene cloning	Li *et al.* (2003)
Sub1A	Rice	Submergence tolerance	Positional cloning and transgenic validation	Xu *et al.* (2006)
OsERF922	Rice	Blast resistance (susceptibility factor)	CRISPR-Cas9 knockout	Wang *et al.* (2016)
TaMLO-A1, B1, D1	Bread wheat	Powdery mildew resistance	Editing of three homoeoalleles	Wang *et al.* (2014)
GABA pathway genes	Tomato	Fruit GABA accumulation	Multiplexed CRISPR editing of five pathway genes	Li *et al.* (2018)
SlCLV3, *SlS*, *SlER* promoters	Tomato	Fruit size, inflorescence branching, architecture	Promoter editing generating allelic series	Rodríguez-Leal *et al.* (2017)

Functional genomics and precision breeding

Knowing a gene is useful only if the knowledge enters a breeding programme. Two routes dominate. The first is marker-assisted selection (MAS) using markers designed on the causal polymorphism itself. This removes the recombination risk associated with linked markers, which is the central rationale for converting validated genes into diagnostic, gene-based markers (Kumar *et al.*, 2025). The second is genomic selection. Here genome-wide marker effects are estimated in a training population and used to predict the performance of untested candidates (Crossa *et al.*, 2017). Functional information can improve prediction indirectly, by identifying markers that deserve greater weight. It also helps directly, by supplying validated major-effect loci that can be handled separately from the polygenic background, for example through marker-assisted backcrossing or gene pyramiding (Kumar *et al.*, 2025).

Genome editing adds a third route. Rather than searching germplasm for a favourable allele, the allele can be created in an elite line. The advantage is speed. The constraint is that editing works only where the underlying biology is genuinely understood and where transformation and regeneration are efficient in the target genotype. A structured comparison of conventional, genomics-assisted and precision breeding is presented in Table 3.

Table 3. Comparison of conventional breeding, genomics-assisted breeding and precision breeding based on functional genomics.

Feature	Conventional breeding	Genomics-assisted breeding	Precision breeding
Basis of selection	Phenotype in the field	Genome-wide markers and predicted breeding values	Validated genes, causal alleles and designed edits
Source of variation	Existing germplasm and crosses	Existing germplasm, selected more efficiently	Existing germplasm plus targeted novel alleles
Quantitative traits	Selection on aggregate performance	Prediction across many small-effect loci (Crossa *et al.*, 2017)	Tuning of expression by cis-regulatory edits (Rodríguez-Leal *et al.*, 2017)
Time to a fixed trait	Many generations of crossing and selection	Reduced by early-generation selection	Can be short for a defined single-locus target
Main requirement	Field capacity and experienced	Genotyping platform and	Gene-level knowledge and efficient
	selection	training population	transformation
Main limitation	Slow; low resolution for complex traits	Predictions are population and environment specific	Restricted to genes whose function is well established

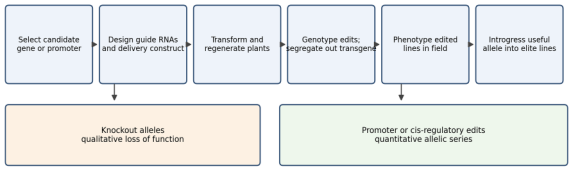

Fig 3. CRISPR-based functional validation and trait improvement. A candidate gene or promoter is targeted, edited plants are regenerated and genotyped, transgene-free segregants are recovered and phenotyped and useful alleles are moved into elite backgrounds. Coding-sequence knockouts give qualitative loss of function, whereas promoter and cis-regulatory edits generate quantitative allelic series suitable for fine-tuning traits.

From candidate genes to designer crops

The phrase designer crop is often used loosely. A more careful definition is a cultivar in which several trait-determining loci have been deliberately combined or adjusted using validated gene function, rather than assembled by chance through recombination.

Two results give the idea concrete meaning. Editing the promoters of tomato genes controlling fruit size, inflorescence branching and plant architecture generated a series of alleles with graded effects. This allows quantitative traits to be dialled up or down instead of switched off (Rodríguez-Leal *et al.*, 2017). It changed how editing is used, since most agronomic traits need adjustment rather than elimination. The second result is de novo domestication. Working in wild allotetraploid rice, researchers assembled the genome, established a transformation and editing system and edited a small set of known domestication genes. The targeted traits included shattering, awn length, plant height, grain characters and flowering (Yu *et al.*, 2021). The lines produced were early-stage material and far from finished cultivars. Even so, the demonstration matters. Functional knowledge from a domesticated crop can be transferred to a wild relative that already carries useful stress tolerance and biomass. Both examples rest on the same foundation. Editing is only as good as the gene function behind it.

Challenges and future perspectives

Several constraints deserve honest statement. The gap between association and causation

remains the central problem. The majority of reported marker-trait associations in crops are never validated at the gene level and unvalidated candidates should not be described as trait genes. Pleiotropy and background dependence complicate translation. A gene that increases grain number may reduce grain filling. An allele that performs well in one background or environment may fail in another.

Technical bottlenecks persist. Transformation and regeneration are inefficient in many elite genotypes and in several legume and cereal crops. Polyploid genomes require simultaneous modification of homoeologues, which adds difficulty even when the target is clear (Wang *et al.*, 2014). Delivery of editing reagents without stable transgene integration is improving, but it is not yet routine across species (Zhang *et al.*, 2016). Integrating transcriptome, proteome, metabolome and phenotype data is also computationally demanding and sensitive to experimental design.

Regulation shapes what reaches farmers. Countries differ in whether transgene-free edited plants are treated as genetically modified organisms and these differences influence research investment, product development and trade. Responsible deployment requires off-target assessment, transparent description of the edits made and attention to access for smallholder farmers.

Looking ahead, pan-genomes will supplement single reference assemblies and capture structural variation missed by earlier resources. Single-cell and spatial methods will link gene activity to specific cell types in crop organs (Chen *et al.*, 2023). Machine learning will help prioritise candidates, although predictions will still need experimental confirmation. Editing chemistries will keep broadening (Shimatani *et al.*, 2017; Lin *et al.*, 2020). Greater use of cis-regulatory editing should make quantitative trait tuning more routine. Nutrient use efficiency, root architecture, heat tolerance during flowering and nutritional quality are likely priorities.

Conclusion

Functional genomics has moved crop science from cataloguing sequence to explaining performance. Sequencing supplies candidates. Expression and mutant analyses narrow them. Genome editing tests them directly in the crop of interest. Cloned genes such as *Sub1A*, *GS3*, *MOC1* and *TaMLO* show that this chain can end in real agronomic gain, rather than in a publication alone. The idea of designer crops is best understood as a direction rather than an achieved state. Progress will come from combining validated gene function with sound breeding practice, careful field evaluation and workable regulation. Handled that way, functional genomics offers one of the more dependable routes to crops that are more productive, more resilient and better suited to changing growing conditions.

Conflict of interest

The authors declare no competing or conflict of interest.

References

Chen, C., Ge, Y., & Lu, L. (2023). Opportunities and challenges in the application of single-cell and spatial transcriptomics in plants. *Frontiers in Plant Science*, *14*, 1185377.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., De Los Campos, G., ... & Varshney, R. K. (2017). Genomic selection in plant breeding: methods, models and perspectives. *Trends in plant science*, *22*(11), 961-975.

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., ... & Zhang, Q. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theoretical and applied genetics*, *112*(6), 1164-1171.

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., ... & Han, B. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature genetics*, *42*(11), 961-967.

Kumar, A., Kumar, B., Kumari, P. and Saket, K. (2025) Marker-assisted selection (MAS). In: Principles of Quantitative and Biometrical Genetics, Chapter 13, pp. 137-151. Agri Root Publication House, India. ISBN 978-81-990999-3-7.

Li, R., Li, R., Li, X., Fu, D., Zhu, B., Tian, H., Luo, Y. and Zhu, H. (2018). Multiplexed CRISPR/Cas9-mediated metabolic engineering of γ-aminobutyric acid levels in Solanum lycopersicum. *Plant biotechnology journal*, *16*(2), 415-427.

Li, X., Qian, Q., Fu, Z., Wang, Y., Xiong, G., Zeng, D., Wang, X., Liu, X., Teng, S., Hiroshi, F., Yuan, M., Luo, D., Han, B. and Li, J. (2003). Control of tillering in rice. *Nature*, *422*(6932), 618-621.

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A.V., Raguram, A., Doman, J.L., Liu, D.R. and Gao, C. (2020). Prime genome editing in rice and wheat. *Nature biotechnology*, *38*(5), 582-585.

McCallum, C. M., Comai, L., Greene, E. A., & Henikoff, S. (2000). T argeting I nduced L ocal L esions IN G enomes (TILLING) for plant functional genomics. *Plant physiology*, *123*(2), 439-442.

Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., & Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell*, *171*(2), 470-480.

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., Teramura, H., Yamamoto, T.,

Komatsu, H., Miura, K., Ezura, H., Nishida, K., Ariizumi, T. and Kondo, A. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nature biotechnology*, *35*(5), 441-443.

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R.K., Tsai, H., Berardi, S., Comai, L. and Dubcovsky, J. (2009). A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC plant Biology*, *9*(1), 115.

Varshney, R. K., Nayak, S. N., May, G. D., & Jackson, S. A. (2009). Next-generation sequencing technologies and their implications for crop genetics and breeding. *Trends in biotechnology*, *27*(9), 522-530.

Wang, F., Wang, C., Liu, P., Lei, C., Hao, W., Gao, Y., Liu, Y.-G. and Zhao, K (2016). Enhanced rice blast resistance by CRISPR/Cas9-targeted mutagenesis of the ERF transcription factor gene OsERF922. *PloS one*, *11*(4), e0154027.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., & Qiu, J. L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nature biotechnology*, *32*(9), 947-951.

Xu, K., Xu, X., Fukao, T., Canlas, P., Maghirang-Rodriguez, R., Heuer, S., Ismail, A.M., Bailey-Serres, J., Ronald, P.C. and Mackill, D.J. (2006). Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice. *Nature*, *442*(7103), 705-708.

Yu, H., Lin, T., Meng, X., Du, H., Zhang, J., Liu, G., Chen, M., Jing, Y., Kou, L., Li, X., Gao, Q., Liang, Y., Liu, X., Fan, Z., Liang, Y., Cheng, Z., Chen, M., Tian, Z., Wang, Y., Chu, C., Zuo, J., Wan, J., Qian, Q., Han, B., Zuccolo, A., Wing, R.A., Gao, C., Liang, C. and Li, J. (2021). A route to de novo domestication of wild allotetraploid rice. *Cell*, *184*(5), 1156-1170.

Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., Qiu, J.-L. and Gao, C. (2016). Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nature communications*, *7*(1), 12617.