J Sreekumar, C Mohan,
N Krishna Radhika and
T Makeshkumar

Advancing tuber crops research: New initiatives

Tuber crops like cassava, sweet potato, yam, and aroids are vital for global food security, and their improvement is being revolutionized by genome editing, omics, and AI. Genome editing, especially CRISPR/Cas, enables precise trait modification for disease resistance, nutrition, and climate resilience. Omics technologies viz., genomics, metabolomics, and phenomics provide insights into genetic variation, biochemical pathways, and trait expression, while integrated multi-omics enhances predictive breeding and biomarker discovery. Traditional breeding remains essential but it is strengthened by these tools. AI and machine learning accelerate data integration, linking genotype to phenotype, predicting complex traits, and enabling climate-smart, resilient, nutritionally rich and high-yielding tuber crop development.

Keywords: AI and Machine Learning, CRISPR/Cas, Genome editing, Omics technologies

WITH the increasing global demand for food and the challenges posed by climate change, innovative solutions are essential to secure the future of agriculture. Among the crops of utmost importance are tuber crops such as cassava, sweet potato, yam, and aroids known for their crucial role in food security and income generation. As researchers work to boost the productivity, quality, and resilience of these crops, new initiatives are emerging, driven by groundbreaking advancements in genome editing, omics technologies, artificial intelligence and integration with traditional breeding. Together, these initiatives are reshaping the future of tuber crop research, providing faster, more precise, and sustainable approaches to crop improvement.

GENOME EDITING

Revolutionizing crop improvement

Genome editing is a precise and simple technology that allows the improvement of crops through genetic modifications. Gene editing has three pre-requisites. The availability of genomic data, an efficient transformation system and a regeneration efficiency for the crop. Genome editing is basically relying on the inherent ability of the DNA to repair itself whenever a double strand break happens in the DNA. This is a precision of genome editing because the exact gene can be targeted to cleave. Based on the type of DNA repair, genome editing is categorised primarily into three groups based on the mode of editing adopted. SDNI in which a random mutation is created precisely in the gene of interest. Here the repair of DNA consequent to Cas mediated cleavage of the DNA is random, creating a loss of function mutation of the gene. This loss of function mutation can be due to addition or deletion or substitution of bases in the DNA which happens through NHEJ pathway. This mutation created is random and simulates natural mutation. In SDNII, a mutation is created in such a way that the DNA repair consequent to the double strand breakage after the Cas protein cleaves the DNA, happens according to the template provided along with the CRISPR/Cas construct and the base alteration happens as directed by the sequence provided. This is homology directed repair. This technology is more precise and the genetic modification can be an addition or deletion or substitution of bases as directed by the sequence provided. In the SDNIII entire gene is replaced with a gene of interest as the repair happens. In India, SDN1 and SDNII genome-edited plants, which do not contain any foreign DNA, are exempt from certain regulatory provisions under the Environment (Protection) Act, 1986. Recently 2 genome edited improved rice varieties has been released in India.

At the cutting edge of modern agriculture is genome editing, particularly the CRISPR/Cas technology. This tool offers an unprecedented ability to precisely modify plant genomes, enabling researchers to target specific genes that control critical traits such as disease resistance, starch composition, nutritional quality and stress tolerance. The main advantage of genome editing is its precision, researchers can alter genes directly without introducing foreign DNA, making it faster and more efficient than traditional breeding methods. This technology has immense potential for improving tuber crops.

- **Disease resistance:** Genome editing can be used to target and modify genes associated with resistance to viral diseases like Cassava Mosaic Disease (CMD) or late blight in potatoes and taro.
- **Starch composition:** By editing genes in the starch biosynthesis pathway, genome editing can improve the quality of starches, waxy starch and high amylose starch, in crops like cassava making them better suited for both food consumption and industrial uses.
- **Climate resilience:** Genetic alterations in genes involved in drought tolerance and heat stress can create tuber varieties that are better equipped to withstand the unpredictable challenges posed by climate change.
- **Improving nutritional quality:** Nutrient-rich tubers are crucial for addressing global malnutrition. Genome editing enables the enhancement of nutritional traits like vitamin A, minerals, and antioxidants. Sweet potatoes, for instance, could be edited to increase their beta-carotene levels, providing a natural, biofortified food source that combats vitamin A deficiencies, especially in developing regions.

With genome-edited crops already being approved for cultivation in certain regions, this technology is positioned to revolutionize tuber crop production by significantly accelerating the breeding process. The public acceptance including policy intervention is on the positive track with respect to genome editing technology compared to older transgenic approaches. The advent of genomics provides a platform to sequence any plant and therefore whole genome data become cheaper, can be attained for any crop. Plant tissue culture transformation and regeneration protocols are available for most tuber crops.

Omics technologies: Unlocking the secrets of plant biology

While genome editing offers a powerful tool for targeted genetic improvements, omics technologies, including genomics, metabolomics, and phenomics, play a complementary role by providing deeper insights into the biological systems that drive plant traits. These technologies enable researchers to explore the plant's genetic makeup, metabolic processes, and physical traits in ways that were previously not possible.

Genomics: Mapping genetic blueprints for success

Genomics focuses on understanding the entire genetic code of an organism, allowing scientists to pinpoint key genes responsible for traits like yield, disease resistance, and nutrient content. By conducting Genome-Wide Association Studies (GWAS) and QTL mapping, researchers can identify genetic variations linked to specific traits, thereby speeding up the breeding process and ensuring the development of high-performing varieties. For example, genomics is essential for identifying resistance genes in tuber crops that protect against pests such as the sweet potato weevil. Once these genes are identified, they can be incorporated into breeding programs.

Metabolomics: Understanding the chemical composition

Metabolomics, the study of small molecules produced within plants, provides valuable information about how genetic variations affect plant biochemistry. By profiling metabolites such as sugars, fatty acids, flavonoids, and antioxidants, metabolomics offers insights into a plant's nutritional quality, stress responses, and disease resistance. In tuber crops, metabolomic profiling can help uncover bioactive compounds like beta-carotene in sweet potatoes or anthocyanins in sweet potato and yams, both of which contribute to nutritional value and consumer appeal. This technology also plays a vital role in identifying metabolites linked to climate resilience and disease resistance, facilitating the development of improved varieties.

Phenomics: Measuring traits in action

Phenomics is the study of observable traits or phenotypes, like tuber size, shape, and colour which are the result of complex interactions between the plant's genotype and the environment. The phenomics tools available today, including drones, infrared cameras, and automated imaging systems, allow researchers to track and measure these traits at large scales, enabling high-throughput phenotyping in both laboratory and field conditions. For instance, automated imaging systems can monitor plant responses to environmental stressors such as drought or disease, allowing researchers to track changes in leaf colour, lesion development, or canopy temperature in real-time. This data can then be integrated with genomic and metabolomic insights to accelerate the development of resilient and high-yielding tuber varieties.

Traditional breeding: Building on natural variability

Despite the advances in genome editing and omics technologies, traditional breeding remains a cornerstone of crop improvement. Traditional breeding methods involve selecting plants with desirable traits and crossing them to develop new varieties over several generations. While this process is slower than genome editing, it remains essential for introducing genetic diversity and stabilizing new traits in crops. Traditional breeding, combined with the power of omics tools, can be used to stack multiple beneficial traits, such as disease resistance and improved starch and with nutritional quality into a single variety. Genomics helps identify the best candidates for breeding, while metabolomics and phenomics can track how those traits express themselves in the field.

Data integration: The power of combined omics

While genomics, metabolomics, and phenomics each provide valuable insights into plant biology, the true power comes from integrating these data sets into a unified framework. By combining the genetic information from genomics, the biochemical profiles from metabolomics, and the physical measurements from phenomics, researchers can create a more comprehensive understanding of how tuber crops respond to environmental stresses, pests, and diseases. For example, when genomics identifies a gene associated with disease resistance, metabolomics can show how that gene affects the production of defensive compounds like phytoalexins, and phenomics can assess how this translates into a reduction in disease symptoms. This multi-layered approach enables more targeted, faster, and precise improvements in crop traits.

Additionally, integrating these datasets allows for the creation of predictive models that can forecast how crops will perform under various environmental conditions. Such models are critical for climate-smart breeding, enabling the development of tuber varieties that are resilient and high-yielding even in changing climates.

Integrating genomics, phenomics, and metabolomics offers a powerful strategy to bridge the gap between genetic variation and observable plant traits. By linking genes to traits through biochemical pathways, researchers can trace the flow of biological information from genotype to phenotype. This pathway typically involves genetic variants influencing metabolic processes, which in turn shape the physical and physiological characteristics of the plant. Understanding this continuum provides deep insights into the molecular mechanisms underlying complex traits in crops.

One of the key benefits of integration is the enhancement of predictive models for breeding. Genomic selection, which uses DNA markers to predict trait performance, becomes significantly more accurate when enriched with phenotypic measurements and metabolite data. This multi-omics approach captures more of the biological variation, improving the reliability of trait prediction and enabling breeders to make more informed selection decisions. Another major advantage is the ability to discover robust biomarkers for traits or responses to stress. Metabolomics can identify specific compounds that signal the presence of drought tolerance, disease resistance, or nutrient efficiency. These biomarkers can then serve as early indicators in breeding pipelines or crop management, facilitating faster and more precise selection or intervention.

Summary of multi-omics integration in tuber crops

Application	Tuber crop examples	Genomics approach	Metabolomics approach	Phenomics approach	Outcome
Starch quality and yield stability	Cassava, sweet potato	GWAS/QTL for GBSS, SBE, SSII	Starch precursors, sugars, ADP-glucose	NIR, viscosity, tuber sizing	High-quality, environment-stable starch varieties
Viral disease resistance	Sweet potato, cassava	NBS-LRR gene mapping, GWAS	Phenolics, antioxidants, Défense amino acids	UAV imaging, early symptom quantification	Early screening and durable resistance breeding
Late blight resistance	Taro	R-gene/QTL mapping	Alkaloids, phytoalexins post-infection	Lesion area quantification via imaging	Durable, multi-gene resistance pipelines
Flavour, aroma, nutrition	Sweet potato, yam	Carotenoid and volatile gene mapping	Pigments, volatiles, sugars, antioxidants	Hyperspectral & sensory phenotyping	Biofortified and market-preferred varieties
Drought/heat stress tolerance	Cassava, sweet potato	ABA, HSP, and root trait gene mapping	Proline, trehalose, antioxidants	IR canopy temp, root depth imaging	Climate-resilient varieties
Pest resistance	Sweet potato	Defensive metabolite QTLs	Cyanogenic glycosides, alkaloids	Damage scoring automation	Reduced pesticide use and enhanced resistance
Plant–pathogen interaction networks	cassava	PTI/ETI gene mapping	SA, JA, ET hormones, defense secondary metabolites	Time-lapse, chlorophyll imaging	Systems-level immune response design
Underground architecture	Cassava, yam, sweet potato	Root/tuber trait QTL mapping	Root exudates and signaling metabolites	X-ray CT, GPR, rhizotrons	Improved nutrient use and pathogen avoidance

AI and Machine Learning for linking genomics and phenomics in tuber crops

With rapid advances in genotyping and high-throughput phenotyping, it is now possible to measure genomes and traits across thousands of plant varieties. Facilities like plant phenotyping centre enable detailed measurements of growth, shape, and yield under controlled conditions. However, traditional statistical genetics often fails when dealing with complex, multigenic

traits, creating a bottleneck for breeding programs that aim to develop resilient, high-yielding crops. To address this, computational methods are being developed to integrate genomic data with decades of accumulated biological knowledge on genes, pathways, interactions, and processes. Such integration not only helps explain genotype–phenotype links in biological terms but also aids in grouping and prioritizing genetic variants that underlie complex traits.

Artificial intelligence (AI) and machine learning (ML) play a key role in this shift. They can process large, noisy, and heterogeneous data sets, extracting meaningful information that traditional methods cannot. Applications already range from analysing raw phenotyping data (images, hyperspectral signals) to predicting complex traits directly from genomic data. For instance, AI/ML models have been used to predict tissue-specific gene expression from sequence features, link hyperspectral reflectance patterns in soybean to yield, and improve genomic prediction through feature selection or deep learning architectures such as "genomic images."

Looking forward, AI/ML offer tremendous opportunities to integrate data across scales, from molecular (single-cell omics) to field (remote sensing and weather data). Developing methods to efficiently represent and merge such heterogeneous data will be crucial, with the potential to accelerate discoveries in plant biology and enable more precise, knowledge-driven breeding of resilient crops. The success of ML critically depends on the availability of sufficiently large numbers of samples that have reasonable quality, are representative of the population and that share a sufficient number of common features. Moreover, compared to more traditional statistical methods, for various applications a disadvantage of ML is that estimating confidence in the predictions is less straightforward, and that model interpretation is not always feasible. Methodology to deal with these issues is currently under active development. Note that model interpretation by dissecting AI/ML models is useful for revealing the underlying biological mechanisms, and is also important for revealing quality issues and potential biases in the data. An additional, important consideration is that new ML methodologies are needed for problems with complex, heterogeneous and/or variable data which current methods find difficult to handle.

CONCLUSION

The new initiatives in tuber crops research, including genome editing, omics technologies, and traditional breeding, represent a holistic approach to agricultural improvement. Data integration across these fields is enabling researchers to unlock the full potential of tuber crops, speeding up the development of resilient, nutritious, and high-yielding varieties. These innovations are not only making it possible to tackle the challenges of climate change, pest resistance, and food security but also positioning tuber crops at the center of future agricultural solutions. By leveraging the combined power of genomics, metabolomics, phenomics, and genome editing, researchers can create more efficient and sustainable farming systems that benefit both farmers and consumers. The future of tuber crop research is not just about individual breakthroughs but about how these diverse technologies can work together to create a brighter, more secure food future.

For further information, please contact:
ICAR-Central Tuber Crops Research Institute, Sreekariyam, Thiruvananthapuram, 695 017, Kerala