

Decoding the Cucumber Genome: Functional Genomics and Its Applications in Genetic Improvement

Xuewen Xu, Xiaodong Yang, Xuehao Chen ✉

School of Horticulture and Landscape Architecture, Yangzhou University, Yangzhou, 225009, Jiangsu, China

✉ Corresponding author: xhchen@yzu.edu.cn

International Journal of Horticulture, 2024, Vol.14, No.5 doi: [10.5376/ijh.2024.14.0030](https://doi.org/10.5376/ijh.2024.14.0030)

Received: 18 Jul., 2024

Accepted: 25 Aug., 2024

Published: 20 Sep., 2024

Copyright © 2024 Xu et al., This is an open access article published under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Preferred citation for this article:

Xu X.W., Yang X.D., and Chen X.H., 2024, Decoding the cucumber genome: functional genomics and its applications in genetic improvement, International Journal of Horticulture, 14(5): 283-296 (doi: [10.5376/ijh.2024.14.0030](https://doi.org/10.5376/ijh.2024.14.0030))

Abstract Cucumber (*Cucumis sativus* L.), as an important global economic crop, has significant implications for enhancing yield and quality through the decoding of its genome. The development of functional genomics provides powerful tools for cucumber genome research. Utilizing high-throughput sequencing and gene editing technologies, researchers can gain insights into the structure and function of the cucumber genome. This review uses functional genomics to analyze key genes in the cucumber genome, exploring their roles in disease resistance, fruit quality, and stress tolerance. The results showed that through a comprehensive analysis of the cucumber genome, a series of disease resistance-related genes were identified, and the disease resistance of cucumbers was successfully enhanced through gene editing technology. Additionally, by regulating genes related to fruit development, the quality of cucumber fruits was significantly improved, and the yield under stress conditions was increased. These findings highlight the potential of functional genomics in cucumber improvement, providing new ideas and approaches for crop breeding. This review not only supports theoretical research on gene function in cucumbers but also offers valuable references for functional genomics in other crops.

Keywords Cucumber genome; Functional genomics; Crop improvement; High-throughput sequencing; Disease resistance

1 Introduction

Cucumber (*Cucumis sativus* L.) is a significant vegetable crop cultivated worldwide, valued for its nutritional content and economic importance. It is a rich source of vitamins and minerals, making it a staple in many diets (Han et al., 2022). Additionally, cucumber serves as a model system for studying sex determination and plant vascular biology, further highlighting its scientific relevance (Huang et al., 2009; Li et al., 2019). The crop's relatively small, diploid genome, short life cycle, and self-compatible mating system offer advantages for genetic studies and breeding programs (Wang et al., 2020).

Research on the cucumber genome has advanced significantly over the years. The initial draft genome sequence of *Cucumis sativus* var. *sativus* L. was assembled using a combination of traditional Sanger and next-generation Illumina GA sequencing technologies, achieving 72.2-fold genome coverage (Huang et al., 2009). This foundational work revealed key insights into the cucumber's genetic structure, including the absence of recent whole-genome duplication and the presence of few tandem duplications, which explain the small number of genes in the cucumber (Huang et al., 2009). Subsequent efforts have improved the quality and completeness of the cucumber genome assembly. For instance, a high-quality reference genome was generated using single-molecule real-time (SMRT) long reads, yielding more sequence data and identified novel genes and retrotransposons (Li et al., 2019). These advancements have facilitated the identification of genes and quantitative trait loci (QTL) responsible for key phenotypic traits, aiding in marker-assisted selection and breeding (Wang et al., 2018; Wang et al., 2020).

In this review, we summarized advancements in cucumber genome decoding and its applications for crop improvement. And extensively discussed the identification of genes associated with key horticultural traits such as disease resistance, fruit quality, and stress tolerance through genomic technologies and genetic analysis.

Additionally, we highlighted recent progress in developing molecular markers and genetic maps for breeding programs aimed at enhancing desirable traits. This review provides valuable insights for genetic research, gene discovery, and the development of superior cucumber varieties, thereby supporting the sustainable advancement of cucumber breeding.

2 Cucumber Genome Structure

2.1 Genome size and organization

The cucumber genome is relatively small, with an estimated size of approximately 367 Mbp. Recent advancements in sequencing technologies have enabled the assembly of high-quality reference genomes. For instance, the research team generated a high-quality cucumber reference genome using advanced technologies such as PacBio, 10X Genomics, and Hi-C, resulting in a total length of 226.2 Mb and an N50 of 8.9 Mb (Li et al., 2019). This assembly revealed new features, including 1,374 full-length long terminal retrotransposons and 1,078 novel genes, showcasing many new characteristics of the cucumber genome (Li et al., 2019).

Additionally, the genome of *Cucumis hystrix*, a wild species closely related to cucumber, was assembled to a size of 289 Mb, providing insights into the genetic diversity and potential for introgression in cucumber breeding (Qin et al., 2021). The study indicates that *Cucumis hystrix* is cross-compatible with cultivated cucumber, which offers the potential for enriching the cucumber gene pool and improving traits such as disease resistance and stress tolerance.

2.2 Chromosome mapping

Chromosome mapping in cucumber has been significantly advanced through various genomic and cytogenetic techniques. Li et al. (2019) assembled a high-quality cucumber genome reference, which includes 174 contigs with a total length of 226.2 Mb and an N50 of 8.9 Mb, providing 29.0 Mb more sequence data than previous versions. Through the use of 10X Genomics and Hi-C data, 89 contigs (approximately 211.0 Mb) were directly connected into 7 pseudomolecule sequences. This high-quality genome reveals new features of the cucumber genome and serves as a valuable resource for cucumber and plant comparative genomics studies (Li et al., 2019).

Turek et al. (2023) conducted a structural analysis of the B10v3 cucumber genome by integrating biological and bioinformatics data. The study demonstrated that by aligning sequences with the reference genome, the RagTag program was used to reorder the sequenced contigs, confirming the chromosomal positions of approximately 98% of protein-coding genes. Additionally, BLAST analysis revealed similarities and differences between the B10v3 genome and other cucumber cultivar genomes (Turek et al., 2023). These studies indicate that by integrating advanced genomic sequencing technologies such as SMRT long-read sequencing, 10X Genomics, and Hi-C data, cucumber chromosomal mapping and genome assembly have reached a new level, providing a solid foundation for cucumber genomics research and breeding.

2.3 Gene content and distribution

The cucumber genome contains a diverse array of genes, and significant progress has been made in the functional annotation of protein-coding genes. A recent study identified 94,486 pairs of homologous protein-coding genes between cucumber and 14 other angiosperm species (Song et al., 2018). These homologous genes were used as proxies for functional annotation, significantly improving the accuracy of gene function prediction in cucumber. The study ultimately assigned Gene Ontology (GO) terms to 10,885 cucumber protein-coding genes, demonstrating improved annotation accuracy compared to existing methods (Song et al., 2018).

Genome-wide association studies (GWAS) have identified regions associated with important horticultural traits, providing valuable resources for crop improvement (Wang et al., 2018). Additionally, the identification of QTL and molecular marker genes has further deepened the understanding of the genetic basis of key phenotypic traits in cucumber (Wang et al., 2020).

3 Functional Genomics of Cucumber

Functional genomics is a field of molecular biology that attempts to describe gene functions and interactions. In cucumbers, functional genomics encompasses various omics technologies, including transcriptomics, proteomics, and metabolomics, to understand the complex biological processes and improve crop traits.

3.1 Transcriptomics

Transcriptomics involves the study of the entire set of RNA transcripts produced by the genome under specific conditions or within specific cells. In cucumber, transcriptome analysis has played a significant role in identifying gene expression patterns related to various traits such as fruit development, stress response, and disease resistance. For example, the 'Chinese Long' cucumber variety has undergone whole-genome sequencing, and transcriptomic, proteomic, and metabolomic data have been obtained. These research data provide scientists with a comprehensive understanding of the physiological characteristics and molecular developmental mechanisms of 'Chinese Long' cucumber, particularly in the improvement of fruit quality (Han et al., 2022).

Moreover, the WRKY gene family plays an essential role in stress response. Extensive research using transcriptomics has revealed specific genes that respond to biotic and abiotic stresses (Chen et al., 2020). The study identified 61 *WRKY* genes in cucumber and explored their roles in responding to various biotic and abiotic stresses (Figure 1). Through gene expression analysis, it was found that *CsWRKY27*, *CsWRKY50*, and *CsWRKY52* were significantly expressed under all stress treatments, indicating that these genes might play a crucial role in stress tolerance in cucumber (Chen et al., 2020). Additionally, *CsWRKY46* was expressed only under biotic stress, while *CsWRKY57* showed a significant response to salt and heat stress. These findings provide new insights into the functional roles of *WRKY* genes in cucumber stress responses.

3.2 Proteomics

Proteomics is the large-scale study of proteins, particularly their structures and functions. In cucumber, proteomic studies have been employed to understand the protein expression profiles and their modifications under different conditions. One study by Hao et al. (2020) analyzed the differentially expressed proteins in cucumber fruits under nitrogen deficiency using proteomics methods, finding that these proteins were mainly associated with carbon metabolism, amino acid synthesis, ascorbate metabolism, and the proteasome. Nitrogen deficiency enhanced the glucose phosphorylation process while inhibiting the pentose phosphate pathway, significantly affecting carbon metabolism and the synthesis of most amino acids, and possibly leading to the accumulation of ascorbate in cucumber fruits.

Another study provided updated functional annotations for protein-coding genes in the cucumber genome. By comparing the genomes of 15 species, the study identified 94,486 orthologous protein-coding gene pairs (OPPs) between cucumber and 14 other angiosperms and assigned GO terms to 10,885 cucumber protein-coding genes. This study proposed an effective strategy for transferring functional information from protein-coding genes in model plants to newly sequenced or "non-model" plant species (Song et al., 2018). Additionally, the development of an EMS mutant library facilitated the identification of proteins associated with various phenotypic traits, thereby deepening the understanding of the genetic basis of these traits (Chen et al., 2018).

3.3 Metabolomics

Metabolomics is the study of chemical processes involving metabolites, the small molecule substrates, intermediates, and products of metabolism. In cucumbers, metabolomic analyses have been used to profile the metabolites associated with fruit quality and stress responses. The integration of metabolomic data with transcriptomic and proteomic data has provided a comprehensive understanding of the metabolic pathways and their regulation. For example, the complete sequencing of the "Chinese Long" cucumber genome has enabled the identification of key metabolites involved in fruit quality and stress tolerance, which are crucial for molecular breeding programs aimed at improving cucumber varieties (Kumar et al., 2020; Han et al., 2022).

Additionally, Miao et al. (2019) demonstrated that grafting has a significant impact on the flavor of cucumber fruits. After grafting with different rootstocks, there were notable changes in the metabolites, such as sugars and amino acids, within cucumber fruits. Through an integrated analysis of metabolomics and transcriptomics, candidate genes related to sugar metabolism and volatile compound synthesis were identified, providing scientific evidence for improving the fruit flavor of grafted cucumbers (Miao et al., 2019).

The integration of transcriptomics, proteomics, and metabolomics in cucumber research has significantly advanced our understanding of the functional genomics of this important crop. These omics technologies provide valuable insights into the genetic and molecular mechanisms underlying key traits, thereby facilitating the development of improved cucumber varieties through molecular breeding techniques.

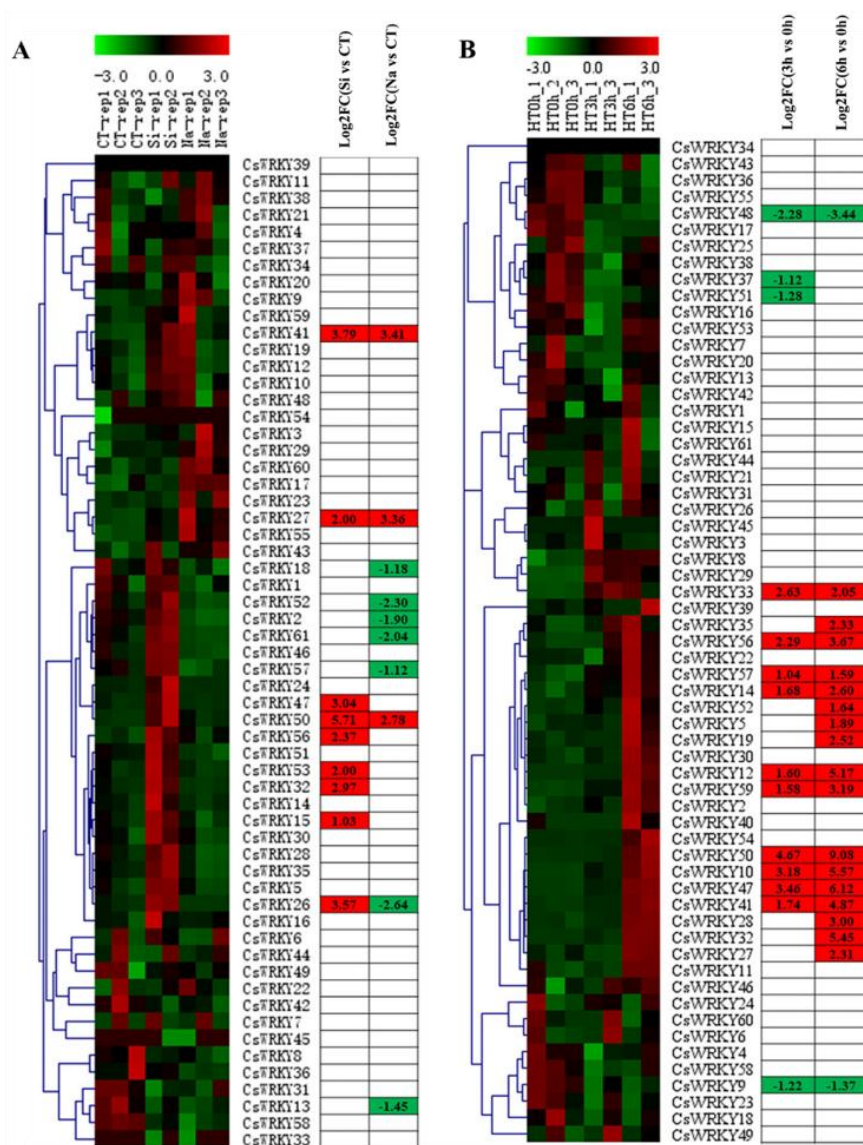


Figure 1 Expression profiles of *CsWRKY* genes in response to various abiotic stress treatments (Adopted from Chen et al., 2020)
 Image caption: The figure shows the expression patterns of cucumber *WRKY* genes under different abiotic stresses (salt stress and heat stress), with changes in gene expression before and after stress treatments displayed in the form of a heatmap. The color gradient from green to red represents a decrease or increase in gene expression levels. The results indicate that several *WRKY* genes exhibit significant expression changes under salt and heat stress, particularly *WRKY27*, *WRKY50*, and *WRKY52*, which respond significantly to both types of stress. This suggests that these genes may play a key role in cucumber's response to abiotic stress, revealing the potential function of *WRKY* genes in enhancing cucumber's stress tolerance and providing important insights for further exploration of their specific mechanisms in stress responses (Adapted from Chen et al., 2020)

4 Key Genes and Pathways in Cucumber

4.1 Disease resistance genes

Cucumber plants have developed various genetic mechanisms to resist diseases, particularly those caused by pathogens such as the root-knot nematode (*Meloidogyne incognita*). Research has identified several key genes and pathways involved in this resistance. For instance, the comparative transcriptomic analysis between susceptible cucumber inbred line Q24 and resistant *Cucumis metuliferus* (CM) revealed that genes associated with Ca²⁺ signaling, salicylic acid (SA)/jasmonate (JA) signaling, and auxin (IAA) signaling pathways play crucial roles in mediating resistance. Genes for calmodulin and calcium-binding proteins were upregulated in CM, while SA/JA synthesis and signal transduction genes were markedly activated, and IAA signaling pathway genes were inhibited upon infection (Li et al., 2021).

Cheng et al. (2019) conducted a genome-wide analysis and identified several candidate genes within QTL regions associated with resistance to root-knot nematodes in cucumber. These genes are primarily involved in the expression of disease resistance proteins and the regulation of plant hormone signaling pathways, such as the SA and JA signaling pathways. Additionally, WRKY transcription factors have been shown to respond to biotic stresses, including infections from downy mildew and powdery mildew, indicating their role in disease resistance (Chen et al., 2020).

4.2 Growth and development pathways

The growth and development of cucumber plants are regulated by various transcription factors and gene families. The GRAS transcription factors, for example, are involved in regulating plant growth and development. *CsGRAS2* (DELLA) and *CsGRAS26* (LISCL) have been identified as key regulators in cucumber, influencing growth through their response to phytohormones such as gibberellin (GA) and abscisic acid (ABA) (Li et al., 2020a).

Similarly, the WRKY gene family has been implicated in plant growth and organ development, with certain WRKY genes showing organ-specific expression patterns (Chen et al., 2020). The *DUF966* genes also play a significant role in fruit development, with *CsDUF966_4.c* and *CsDUF966_7.c* being strongly selected during cucumber breeding programs for their regulation of fruit growth (Tian et al., 2022).

4.3 Stress response mechanisms

Cucumber plants encounter various abiotic stresses, such as salinity, heat, and drought, which necessitate robust stress response mechanisms. The GRAS transcription factors, particularly *CsGRAS2* and *CsGRAS26*, are involved in the response to multiple abiotic stresses, including low and high temperatures and salinity (Li et al., 2020b). The WRKY gene family also plays a critical role in abiotic stress responses, with several *WRKY* genes responding to salt and heat stresses (Chen et al., 2020).

Additionally, the basic/helix-loop-helix (bHLH) transcription factor family, specifically *CsbHLH041*, has been shown to enhance tolerance to salt and ABA in transgenic Arabidopsis and cucumber seedlings, indicating its importance in abiotic stress response (Li et al., 2020b). The PYL gene family, which is central to ABA signal transduction, also contributes to stress responses, with most PYL genes being upregulated under ABA, PEG, and salt stress treatments (Zhang et al., 2022). Furthermore, the HSP20 gene family is involved in heat stress response, with many *CsHSP20* genes being upregulated under high-temperature conditions (Huang et al., 2022).

5 Genetic Variation and Breeding

5.1 Sources of genetic variation

Genetic variation is the cornerstone of breeding programs aimed at crop improvement. In cucumber, significant genetic diversity has been identified through various genomic studies. The USDA cucumber collection, for instance, has been characterized using high-throughput genotyping-by-sequencing (GBS) technology, revealing over 23,000 high-quality single-nucleotide polymorphisms (SNPs) that highlight the genetic diversity and population structure within the collection (Wang et al., 2018). Additionally, whole-genome re-sequencing has identified numerous insertion and deletion (InDel) markers, which are associated with genetic variation and can be

used for hybrid seed purity testing and germplasm classification (Adedze et al., 2021). These genetic resources are crucial for understanding the genetic basis of important traits and for the development of new cucumber varieties with desirable characteristics.

5.2 Marker-assisted selection

Marker-assisted selection (MAS) is a powerful tool that enhances breeding efficiency by allowing precise selection of desirable traits based on molecular markers. In cucumber, MAS has been successfully applied to improve various traits, including disease resistance and fruit quality. For example, SNP markers have been developed for the selection of cucumber vein yellowing virus (CVYV) resistance, enabling the rapid generation of resistant lines (Kahveci et al., 2021). Similarly, a set of 327 common background selection markers (BMs) has been developed for efficient MAS in cucumber breeding programs (Lee et al., 2022). These markers facilitate the selection of plants with the desired genetic background, thereby accelerating the breeding process. Furthermore, the identification of QTLs associated with important traits, such as pre-harvest sprouting and fruit shape, provides valuable targets for MAS (Zhang et al., 2019; Cao et al., 2021).

5.3 Genomic selection

Genomic selection (GS) is an advanced breeding approach that uses genome-wide markers to predict the performance of breeding lines. This method has been increasingly applied in cucumber breeding to improve complex traits. A study on genomic prediction in cucumber demonstrated the effectiveness of GS in predicting 12 commercially important traits, with predictive abilities ranging from 0.38 to 0.95 under cross-validation strategies (Liu et al., 2021). The study highlighted the importance of incorporating non-additive genetic effects into the prediction models to enhance their accuracy. By leveraging the genetic information from a large number of markers, GS enables the selection of superior breeding lines at an early stage, thus reducing the breeding cycle and increasing the efficiency of developing new cucumber varieties.

The integration of genetic variation studies, marker-assisted selection, and genomic selection provides a comprehensive framework for the improvement of cucumber crops. These advanced genomic tools and techniques are instrumental in developing cucumber varieties with enhanced traits, ultimately contributing to sustainable agriculture and food security.

6 Biotechnological Approaches in Cucumber Improvement

6.1 Genetic engineering

Genetic engineering has been a cornerstone in the field of crop improvement, providing tools to introduce new traits and enhance existing ones in cucumbers. *Agrobacterium tumefaciens*-mediated transformation is a widely used method for genetic engineering in cucumbers. This technique involves the transfer of desired genes into the cucumber genome, leading to the development of transgenic plants with improved traits such as disease resistance, enhanced nutritional content, and better yield. For example, Fan et al. (2020) developed a new method to generate composite cucumber plants with transgenic roots through an *Agrobacterium*-mediated one-step process. This method is simple and efficient, making it suitable for root functional analysis and other root-related studies.

However, the efficiency of this transformation method can be low, and various factors such as genotype, explant source, and co-cultivation conditions need to be optimized to improve transformation rates (Liu et al., 2023). Despite these challenges, genetic engineering remains a powerful tool for cucumber improvement, enabling the introduction of novel traits that are difficult to achieve through traditional breeding methods.

6.2 CRISPR/Cas9 technology

The CRISPR/Cas9 system has revolutionized the field of genetic engineering by providing a precise and efficient method for genome editing. This technology allows for targeted modifications in the cucumber genome, enabling the creation of new mutants with desirable traits. For instance, CRISPR/Cas9 has been used to create compact plant architectures in cucumbers by targeting specific genes responsible for plant growth (Xin et al., 2022). The simplicity and versatility of CRISPR/Cas9 make it an attractive tool for cucumber improvement, as it can be used

to enhance traits such as disease resistance, stress tolerance, and yield (Jaganathan et al., 2018; Schulze and Lammers, 2020; Li et al., 2021).

Studies have shown that disrupting the function of the *eIF4E* gene in cucumber using CRISPR/Cas9 technology has led to the development of resistance to Cucumber Vein Yellowing Virus (CVYV) and several other viruses. This non-transgenic method enables the rapid generation of disease-resistant mutants without affecting plant development (Chandrasekaran et al., 2016). Additionally, the development of optimized transformation protocols has further enhanced the efficiency of CRISPR/Cas9-mediated genome editing in cucumber, making it a promising approach for future crop improvement efforts (Xin et al., 2022).

6.3 RNA interference (RNAi)

RNAi is another powerful tool for functional genomics and crop improvement. This technology involves the downregulation of specific genes through the introduction of small interfering RNAs (siRNAs), which can silence target genes without affecting other genes in the genome. RNAi has been successfully used in cucumbers to develop resistance against various biotic and abiotic stresses, improve nutritional quality, and modify plant architecture (Rajam, 2020). The non-transgenic nature of RNAi-based approaches makes them particularly appealing, as they do not involve the introduction of foreign DNA into the plant genome, thereby reducing regulatory hurdles and public concerns associated with genetically modified organisms (GMOs) (Távora et al., 2022). The combination of RNAi with other biotechnological approaches, such as CRISPR/Cas9, holds great potential for the comprehensive improvement of cucumber crops (Rajput et al., 2021; Távora et al., 2022).

Biotechnological approaches such as genetic engineering, CRISPR/Cas9 technology, and RNA interference offer powerful tools for the improvement of cucumber crops. These technologies enable precise modifications in the cucumber genome, leading to the development of plants with enhanced traits that can contribute to increased yield, disease resistance, and stress tolerance. The continued optimization and integration of these approaches will be crucial for meeting the growing demand for food and ensuring sustainable agricultural practices.

7 Case Studies

7.1 Enhancing disease resistance in cucumber

Enhancing disease resistance in cucumber is a key objective of crop improvement. Lee et al. (2020) conducted a genetic diversity assessment of 264 cucumber germplasm resources and, combined with GWAS, analyzed three horticultural traits: powdery mildew resistance, pedicel color, and spine color. The study found significant genetic diversity among cucumber germplasm resources, which were divided into three subpopulations based on their geographic distribution. The East Asian population exhibited a genetic background that was significantly different from cucumber germplasm resources in other regions. Additionally, the study identified key SNPs and candidate genes associated with powdery mildew resistance and pedicel and spine color, such as *Csa5G453160* and *Csa5G471070*, which are valuable references for cucumber breeding (Figure 2).

The USDA cucumber collection has also played an important role in identifying genetic variations related to disease resistance traits. The research team utilized high-throughput genome sequencing technology to analyze 1,234 cucumber samples, ultimately developing a core germplasm collection of 395 representative resources that captured at least 96% of genetic diversity (Wang et al., 2018). Furthermore, the development of an EMS mutant library has provided valuable germplasm resources, including mutants with altered disease resistance traits, further advancing the genetic improvement of cucumber (Chen et al., 2018).

7.2 Improving fruit quality

Improving the quality of cucumber fruits involves enhancing traits such as flavor, texture, and nutritional content. Specific genes identified through functional genomics, such as those involved in the biosynthesis of raffinose family oligosaccharides (RFOs), have shown potential to improve fruit quality under stress conditions. For example, the overexpression of *CsGo/SI* has been demonstrated to enhance assimilate transport and fruit growth under cold stress (Dai et al., 2022).

The study by Dai et al. (2022) explored the role of *Galactinol synthase 1* (*CsGolS1*) in cucumber under cold stress. The results indicated that *CsGolS1* is expressed in the minor veins of cucumber leaves and enhances the efficiency of assimilate transport by promoting the synthesis of RFOs. This improvement in assimilate transport under cold stress led to better overall plant performance. By specifically overexpressing *CsGolS1* under cold stress conditions, the growth rate of cucumber leaves, fruits, and the entire plant was significantly increased (Figure 3). This suggests that *CsGolS1* alleviates the negative impact of cold stress on assimilate transport speed by increasing the concentration of RFOs in phloem sap, thereby promoting plant growth and photosynthesis.

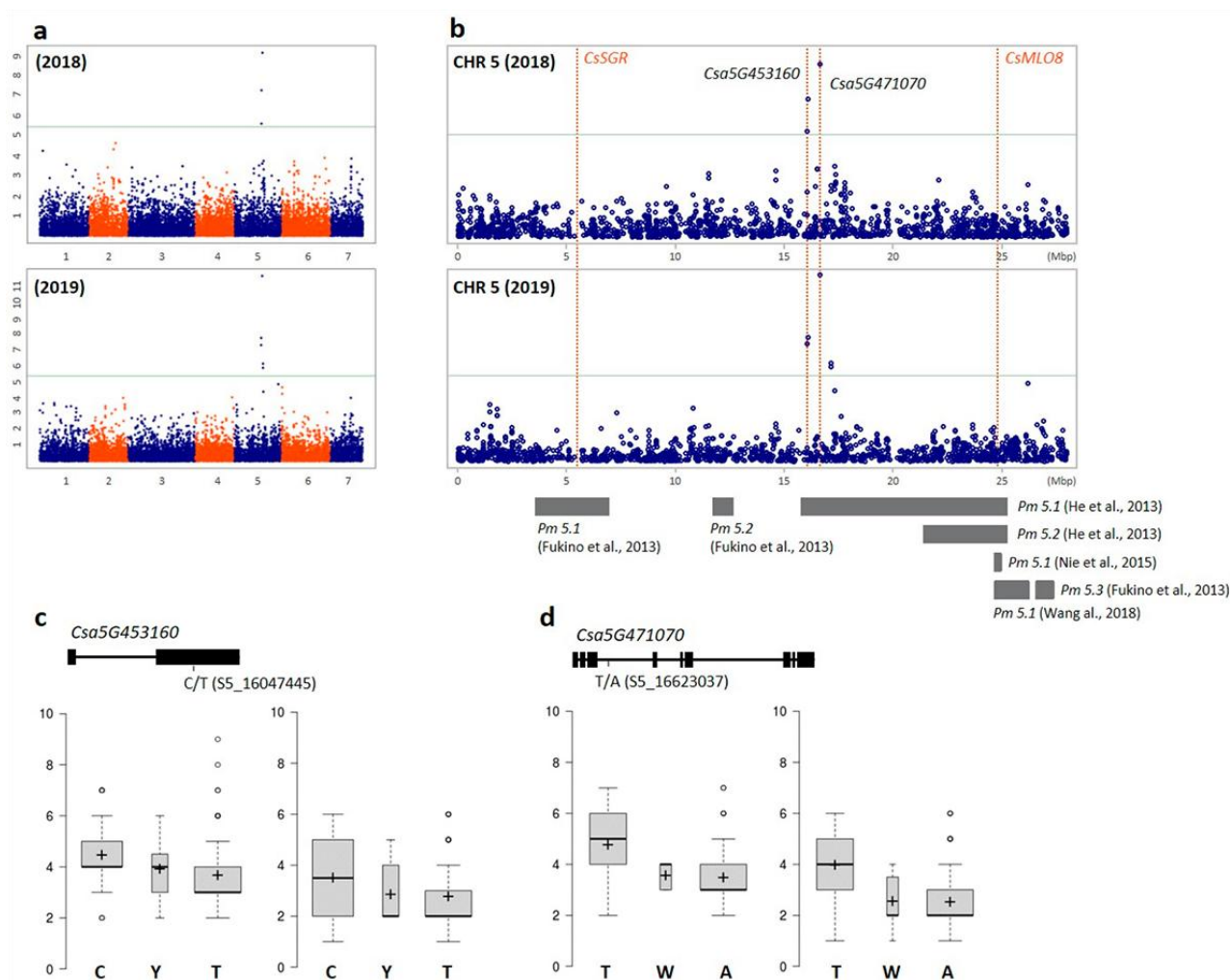


Figure 2 Significant SNPs and candidate genes correlating cucumber PM resistance from GWAS (Adopted from Lee et al., 2020)
 Image caption: a shows the Manhattan plots for the years 2018 and 2019, highlighting the SNPs significantly associated with powdery mildew resistance, primarily concentrated in the 16 to 17 Mbp region on chromosome 5. b further illustrates the positional relationship between these significant SNPs, known QTLs, and candidate genes. c and d validate the potential role of candidate genes *Csa5G453160* and *Csa5G471070* in powdery mildew resistance through gene structure and SNP polymorphism analysis. These findings provide important genetic markers for molecular breeding of powdery mildew resistance (Adapted from Lee et al., 2020)

7.3 Increasing yield under stress conditions

Increasing cucumber yield under stress conditions is essential for ensuring stable production. The identification of salt tolerance genes through GWAS has provided insights into the molecular mechanisms underlying stress resistance. Seven loci associated with salt tolerance have been identified, and candidate genes such as *CsaV3_2G035120* and *CsaV3_3G023710* have been predicted to play significant roles in enhancing stress tolerance (Liu et al., 2022). In the future, functional validation of these genes, such as through overexpression or

CRISPR technology, could further elucidate their precise roles in salt tolerance. This would provide valuable insights for the development of high-yielding, high-quality, and salt-tolerant cucumber varieties.

Additionally, a comprehensive characterization of the GRAS gene family has revealed its role in regulating responses to various abiotic stresses, including salinity and extreme temperatures. Using RNA-seq and qRT-PCR techniques, the study found that specific *GRAS* genes, such as *CsGRAS2* (DELLA) and *CsGRAS26* (LISCL), exhibited significant regulatory activity under all abiotic stress conditions and plant hormone treatments, such as gibberellin and abscisic acid. This suggests that these genes may play a critical regulatory role in the interaction of multiple signaling pathways. Notably, when comparing cucumber varieties with different levels of stress tolerance (such as CG104 and CG37), the expression of *CsGRAS2* and *CsGRAS26* was found to significantly influence the plant's stress response to environmental changes (Li et al., 2020a). These findings provide valuable genetic resources for future molecular breeding efforts aimed at enhancing cucumber's environmental adaptability.

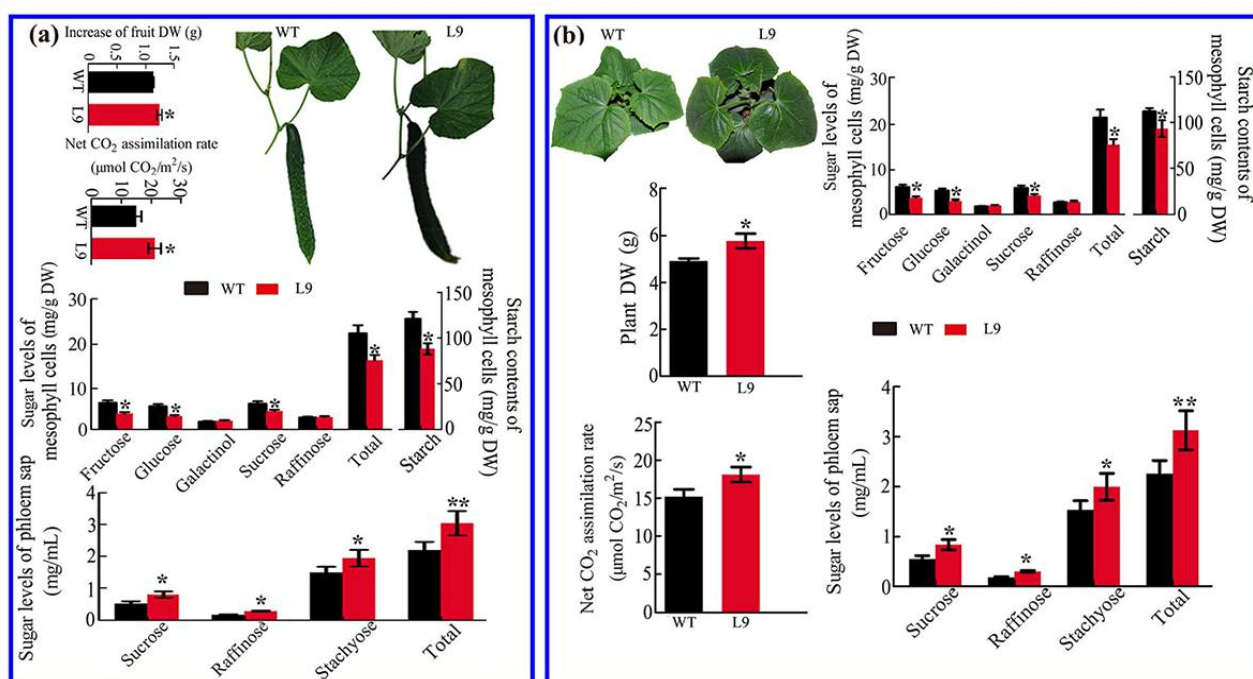


Figure 3 Cold-induced intermediary cell-specific overexpression of *CsGolS1* promoted fruit growth, whole plant development, and source leaf photosynthesis under cold stress (Adopted from Dai et al., 2022)

Image caption: The figure shows the effect of cold-induced minor vein-specific overexpression of *CsGolS1* on cucumber fruit growth and whole-plant development under cold stress. The results indicate that the fruit growth rate of transgenic cucumbers under cold stress is significantly faster than that of the wild type, and the accumulation of dry matter in the whole plant is also notably increased. Further analysis reveals that this growth advantage is associated with higher sugar content in the phloem sap, lower sucrose and starch levels in mesophyll cells, and higher photosynthetic efficiency. The figure demonstrates that *CsGolS1* enhances assimilate translocation efficiency, promoting overall growth and fruit development in cucumbers under cold stress, and confirms the potential of *CsGolS1* in improving cold tolerance in plants (Adapted from Dai et al., 2022)

8 Applications in Crop Improvement

8.1 Development of new cucumber varieties

The development of new cucumber varieties has been significantly advanced through the application of genomics and molecular breeding techniques. The complete sequencing of the "Chinese Long" cucumber genome has facilitated the identification of key genetic markers and the development of genetic maps, which are crucial for breeding programs aimed at improving fruit quality and other desirable traits (Han et al., 2022). Genomic prediction models have also been employed to estimate the general combining ability (GCA) of inbred lines, enhancing the efficiency of breeding programs by predicting complex traits with high accuracy (Liu et al., 2021).

Additionally, the creation of an EMS mutant library has provided valuable germplasm resources for functional genomic research and the improvement of cucumber genetics (Chen et al., 2018).

8.2 Sustainable agricultural practices

Sustainable agricultural practices in cucumber cultivation have been bolstered by the integration of genomics-assisted breeding (GAB). GAB has enabled the targeted manipulation of allelic variation, facilitating the development of cucumber varieties with enhanced stress tolerance and disease resistance (Varshney et al., 2021). The genetic diversity within cucumber germplasm collections, characterized through high-throughput genotyping, has been leveraged to identify genome regions associated with important horticultural traits, such as powdery mildew resistance and fruit quality (Wang et al., 2018; Lee et al., 2020). These advancements contribute to more resilient cucumber varieties that require fewer chemical inputs, promoting sustainable agricultural practices.

8.3 Future prospects in cucumber breeding

The future of cucumber breeding looks promising with the continued advancements in functional genomics and molecular breeding techniques. The chromosome-scale genome assembly of *Cucumis hystrix*, a wild species interspecifically cross-compatible with cultivated cucumber, offers new opportunities for introgression breeding, potentially introducing novel traits such as enhanced disease resistance and stress tolerance into cultivated varieties (Qin et al., 2021). Furthermore, the systematic review and documentation of molecularly tagged genes and QTL provide a comprehensive resource for marker-assisted selection, streamlining the breeding process and accelerating the development of superior cucumber varieties (Wang et al., 2020). As genomics-assisted breeding evolves, it is expected to play a crucial role in designing climate-smart, high-yielding, and nutritionally superior cucumber cultivars (Varshney et al., 2021).

9 Challenges and Opportunities

9.1 Technical challenges in cucumber genomics

The field of cucumber genomics faces several technical challenges that hinder the full exploitation of genomic data for crop improvement. One significant challenge is the limited availability of genomic resources and tools for functional genomics studies in cucumber. Despite the advancements in sequencing technologies, the construction of comprehensive mutant libraries and the optimization of mutagenesis conditions remain complex and resource-intensive tasks (Chen et al., 2018). Additionally, the acquisition of phenotypic data continues to be a bottleneck in functional genomics studies. High-throughput phenotyping technologies are still evolving, and their integration with genomic data is crucial for bridging the phenotype-genotype gap (Yang et al., 2020). Furthermore, the narrow genetic base of cultivated cucumber poses a challenge for breeding programs, necessitating the exploration of wild species for introgression of desirable traits (Qin et al., 2021).

9.2 Opportunities for innovation in cucumber research

Despite the challenges, there are numerous opportunities for innovation in cucumber research. The availability of high-throughput sequencing technologies and the development of genomic prediction models offer promising avenues for enhancing cucumber breeding programs. Genomic prediction, which involves the use of genetic markers to predict complex traits, has shown potential in improving the accuracy and efficiency of breeding efforts (Liu et al., 2021). Additionally, the integration of multi-omics approaches, such as transcriptomics, proteomics, and metabolomics, can provide a comprehensive understanding of the molecular mechanisms underlying important traits, thereby facilitating the development of superior cucumber varieties (Han et al., 2022). The establishment of core collections and the characterization of genetic diversity within germplasm collections also present opportunities for identifying novel traits and enhancing crop resilience (Wang et al., 2018).

9.3 Integrating genomics with traditional breeding

Integrating genomics with traditional breeding methods can significantly accelerate the development of improved cucumber varieties. MAS and QTL mapping are powerful tools that can be used to identify and select for desirable traits in breeding programs (Wang et al., 2020). The systematic review and documentation of

molecularly tagged genes and QTLs provide valuable resources for breeders to enhance disease resistance, fruit quality, and other agronomically important traits (Wang et al., 2020). Furthermore, the use of genomic data to inform breeding decisions can lead to the development of varieties with enhanced yield, stress tolerance, and other beneficial traits (Kumar et al., 2020). The combination of traditional breeding techniques with modern genomic tools holds great promise for the future of cucumber crop improvement.

10 Concluding Remarks

The research on the cucumber genome has yielded significant insights into its genetic diversity, population structure, and the identification of key genes and QTL associated with important horticultural traits. High-throughput GBS technology has been instrumental in characterizing the genetic variation within the U.S. National Plant Germplasm System (NPGS) cucumber collection, identifying over 23,000 SNPs and genome regions associated with 13 horticulturally important traits. Additionally, the development of a core collection representing 96% of the genetic variation in the NPGS has provided a valuable resource for future genetic analyses and crop improvement.

The complete sequencing of the "Chinese Long" cucumber genome and subsequent transcriptome, proteome, and metabolome analyses have facilitated the application of molecular breeding techniques to improve fruit quality and stress resistance. GWAS have identified significant SNPs and candidate genes for traits such as powdery mildew resistance, spine color, and fruit stalk-end color, which are crucial for cucumber breeding. Furthermore, the systematic review and documentation of molecularly tagged genes and QTL have provided a comprehensive resource for marker-assisted selection in cucumber breeding.

The advancements in cucumber genomics have profound implications for cucumber breeding. The identification of genetic markers and QTL associated with key traits enables breeders to employ marker-assisted selection, thereby accelerating the development of improved cucumber varieties with enhanced disease resistance, stress tolerance, and fruit quality. The introgression of beneficial traits from wild species such as *Cucumis hystrix* into cultivated cucumber varieties offers a promising avenue for broadening the genetic base and improving the resilience of cucumber crops.

Genomic prediction models have shown potential in predicting complex traits, providing a framework for selecting superior breeding lines with desirable traits. The establishment of an EMS mutant library has expanded the genetic resources available for functional genomics studies, facilitating the identification of genes responsible for important phenotypic traits. The high-quality chromosome-scale genome assembly of cucumber provides a robust reference for comparative genomics and genetic research, further enhancing the precision and efficiency of cucumber breeding programs.

Future research should focus on the functional validation of candidate genes and QTL identified through GWAS and other genomic studies to elucidate their roles in trait expression and regulation. The integration of multi-omics approaches, including transcriptomics, proteomics, and metabolomics, will provide a more comprehensive understanding of the molecular mechanisms underlying important horticultural traits. Efforts should also be directed towards the development of advanced genomic tools and resources, such as CRISPR/Cas9-based gene editing, to facilitate precise genetic modifications and the creation of novel cucumber varieties with improved traits. Additionally, expanding the genetic diversity of cucumber breeding lines through the introgression of traits from wild relatives and other *Cucumis* species will be crucial for enhancing the adaptability and resilience of cucumber crops in the face of changing environmental conditions.

In conclusion, the ongoing advancements in cucumber genomics and functional genomics hold great promise for the future of cucumber breeding, enabling the development of superior cucumber varieties that meet the demands of both growers and consumers.

Acknowledgment

The authors express their sincere gratitude to the two anonymous peer reviewers who contributed to the review of this manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (32172570, 32030093 and 32341043).

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Adedze Y.M.N., Lu X., Xia Y., Sun Q., Nchongboh C.G., Alam M.A., Liu M.H., Yang X., Zhang W.T., Deng Z.J., Li W.H., and Si L.T., 2021, Agarose-resolvable InDel markers based on whole genome re-sequencing in cucumber, *Scientific Reports*, 11(1): 3872.
<https://doi.org/10.1038/s41598-021-83313-x>
- Cao M.M., Li S.J., Deng Q., Wang H.Z., and Yang R.H., 2021, Identification of a major-effect QTL associated with pre-harvest sprouting in cucumber (*Cucumis sativus* L.) using the QTL-seq method, *BMC Genomics*, 22: 1-11.
<https://doi.org/10.1186/s12864-021-07548-8>
- Chandrasekaran J., Brumin M., Wolf D., Leibman D., Klap C., Pearlsman M., Sherman A., Arazi T., and Gal-On A., 2016, Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology, *Molecular Plant Pathology*, 17(7): 1140-1153.
<https://doi.org/10.1111/mpp.12375>
- Chen C., Chen X., Han J., Lu W., and Ren Z., 2020, Genome-wide analysis of the WRKY gene family in the cucumber genome and transcriptome-wide identification of WRKY transcription factors that respond to biotic and abiotic stresses, *BMC Plant Biology*, 20: 1-19.
<https://doi.org/10.1186/s12870-020-02625-8>
- Chen C., Cui Q.Z., Huang S.W., Wang S.H., Liu X.H., Lu X.Y., Chen H.M., and Tian Y., 2018, An EMS mutant library for cucumber, *Journal of Integrative Agriculture*, 17(7): 1612-1619.
[https://doi.org/10.1016/S2095-3119\(17\)61765-9](https://doi.org/10.1016/S2095-3119(17)61765-9)
- Cheng C., Wang X., Liu X., Yang S., Yu X., Qian C., Li J., Lou Q., and Chen J., 2019, Candidate genes underlying the quantitative trait loci for root-knot nematode resistance in a *Cucumis hystrix* introgression line of cucumber based on population sequencing, *Journal of Plant Research*, 132: 813-823.
<https://doi.org/10.1007/s10265-019-01147-1>
- Dai H., Zhu Z., Wang Z., Zhang Z., Kong W., and Miao M., 2022, *Galactinol synthase 1* improves cucumber performance under cold stress by enhancing assimilate translocation, *Horticulture Research*, 9: uhab063.
<https://doi.org/10.1093/hr/uhab063>
- Fan Y., Xu F., Zhou H., Liu X., Yang X., Weng K., Sun X., and Lyu S., 2020, A fast, simple, high efficient and one-step generation of composite cucumber plants with transgenic roots by *Agrobacterium rhizogenes*-mediated transformation, *Plant Cell, Tissue and Organ Culture (PCTOC)*, 141: 207-216.
<https://doi.org/10.1007/s11240-020-01781-x>
- Han D., Ma X., Zhang L., Zhang S., Sun Q., Li P., Shu J., and Zhao Y., 2022, Serial-Omics and molecular function study provide novel insight into cucumber variety improvement, *Plants*, 11(12): 1609.
<https://doi.org/10.3390/plants11121609>
- Hao J., Li Q., Yu H., Wang H., Chai L., Miao T., and Jiang W., 2021, Comparative proteomic analysis of cucumber fruits under nitrogen deficiency at the fruiting stage, *Horticultural Plant Journal*, 7(1): 59-72.
<https://doi.org/10.1016/j.hpj.2020.12.002>
- Huang J., Hai Z., Wang R., Yu Y., Chen X., Liang W., and Wang H., 2022, Genome-wide analysis of HSP20 gene family and expression patterns under heat stress in cucumber (*Cucumis sativus* L.), *Frontiers in Plant Science*, 13: 968418.
<https://doi.org/10.3389/fpls.2022.968418>
- Huang S., Li R., Zhang Z., Li L., Gu X., Fan W., Lucas W., Wang X., Xie B., Ni P., Ren Y., Zhu H., Li J., Lin K., Jin W., Fei Z., Li G., Staub J., Kilian A., Vossen E., Wu Y., Guo J., He J., Jia Z., Ren Y., Tian G., Lu Y., Ruan J., Qian W., Wang M., Huang Q., Li B., Xuan Z., Cao J., Wu A., Wu Z., Zhang J., Cai Q., Bai Y., Zhao B., Han Y., Li Y., Li X., Wang S., Shi Q., Liu S., Cho W., Kim J., Xu Y., Heller-Uszyńska K., Miao H., Cheng Z., Zhang S., Wu J., Yang Y., Kang H., Li M., Liang H., Ren X., Shi Z., Wen M., Jian M., Yang H., Zhang G., Yang Z., Chen R., Liu S., Li J., Ma L., Liu H., Zhou Y., Zhao J., Fang X., Li G., Fang L., Li Y., Liu D., Zheng H., Zhang Y., Qin N., Li Z., Yang G., Yang S., Bolund L., Kristiansen K., Zheng H., Li S., Zhang X., Yang H., Wang J., Sun R., Zhang B., Jiang S., Wang J., Du Y., and Li S., 2009, The genome of the cucumber, *Cucumis sativus* L., *Nature Genetics*, 41: 1275-1281.
<https://doi.org/10.1038/ng.475>
- Jaganathan D., Ramasamy K., Sellamuthu G., Jayabalan S., and Venkataraman G., 2018, CRISPR for crop improvement: an update review, *Frontiers in Plant Science*, 9: 985.
<https://doi.org/10.3389/fpls.2018.00985>

- Kahveci E., Devran Z., Özkaynak E., Hong Y., Studholme D.J., and Tör M., 2021, Genomic-assisted marker development suitable for Cscvy-1 selection in cucumber breeding, *Frontiers in Plant Science*, 12: 691576.
<https://doi.org/10.3389/fpls.2021.691576>
- Kumar R., Jorben J., Yadav R.R., and Senapati M., 2020, Genomics and its application in crop improvement, *Journal of Pharmacognosy and Phytochemistry*, 9(1): 547-552.
<https://doi.org/10.22271/phyto.2020.v9.i1i.10488>
- Lee E.S., Yang H.B., Kim J., Lee H.E., Lee Y.R., and Kim D.S., 2022, Development of SNP marker sets for marker-assisted background selection in cultivated cucumber varieties, *Agronomy*, 12(2): 487.
<https://doi.org/10.3390/agronomy12020487>
- Lee H., Kim J., Kang B., and Song K., 2020, Assessment of the genetic diversity of the breeding lines and a genome-wide association study of three horticultural traits using worldwide cucumber (*Cucumis* spp.) germplasm collection, *Agronomy*.
<https://doi.org/10.3390/agronomy10111736>
- Li C.X., Dong S.Y., Liu X.P., Bo K.L., Miao H., Beckles D.M., Zhang S.P., and Gu X.F., 2020a, Genome-wide characterization of cucumber (*Cucumis sativus* L.) GRAS genes and their response to various abiotic stresses, *Horticulturae*, 6(4): 110.
<https://doi.org/10.3390/horticulturae6040110>
- Li J., Wang T., Han J., and Ren Z., 2020b, Genome-wide identification and characterization of cucumber bHLH family genes and the functional characterization of CsbHLH041 in NaCl and ABA tolerance in Arabidopsis and cucumber, *BMC Plant Biology*, 20: 1-20.
<https://doi.org/10.1186/s12870-020-02440-1>
- Li Q., Li H., Huang W., Xu Y., Zhou Q., Wang S., Ruan J., Huang S., and Zhang Z., 2019, A chromosome-scale genome assembly of cucumber (*Cucumis sativus* L.), *GigaScience*, 8(6): giz072.
<https://doi.org/10.1093/gigascience/giz072>
- Li X., Sun Y., Yang Y., Yang X., Xue W., Wu M., Chen P., Weng Y., and Chen S., 2021, Transcriptomic and histological analysis of the response of susceptible and resistant cucumber to *Meloidogyne incognita* infection revealing complex resistance via multiple signaling pathways, *Frontiers in Plant Science*, 12: 675429.
<https://doi.org/10.3389/fpls.2021.675429>
- Liu C., Liu X., Han Y., Wang X.A., Ding Y., Meng H., and Cheng Z., 2021, Genomic prediction and the practical breeding of 12 quantitative-inherited traits in cucumber (*Cucumis sativus* L.), *Frontiers in Plant Science*, 12: 729328.
<https://doi.org/10.3389/fpls.2021.729328>
- Liu D., Dong S., Miao H., Liu X., Li C., Han J., Zhang S., and Gu X., 2022, A large-scale genomic association analysis identifies the candidate genes regulating salt tolerance in cucumber (*Cucumis sativus* L.) seedlings, *International Journal of Molecular Sciences*, 23(15): 8260.
<https://doi.org/10.3390/ijms23158260>
- Liu H., Zhao J., Chen F., Wu Z., Tan J., Nguyen N., Cheng Z., and Weng Y., 2023, Improving *Agrobacterium tumefaciens*-mediated genetic transformation for gene function studies and mutagenesis in cucumber (*Cucumis sativus* L.), *Genes*, 14(3): 601.
<https://doi.org/10.3390/genes14030601>
- Miao L., Di Q., Sun T., Li Y., Duan Y., Wang J., Yan Y., He C., Wang C., and Yu X., 2019, Integrated metabolome and transcriptome analysis provide insights into the effects of grafting on fruit flavor of cucumber with different rootstocks, *International Journal of Molecular Sciences*, 20(14): 3592.
<https://doi.org/10.3390/ijms20143592>
- Qin X., Zhang Z., Lou Q., Xia L., Li J., Li M., Zhou J., Zhao X., Xu Y., Li Q., Yang S., Yu X., Cheng C., Huang S., and Chen J., 2021, Chromosome-scale genome assembly of *Cucumis hystris*—a wild species interspecifically cross-compatible with cultivated cucumber, *Horticulture Research*, 8: 1-9.
<https://doi.org/10.1038/s41438-021-00475-5>
- Rajam M.V., 2020, RNA silencing technology: A boon for crop improvement, *Journal of Biosciences*, 45(1): 1-5.
<https://doi.org/10.1007/s12038-020-00082-x>
- Rajput M., Choudhary K., Kumar M., Vivekanand V., Chawade A., Ortiz R., and Pareek N., 2021, RNA interference and CRISPR/Cas gene editing for crop improvement: Paradigm shift towards sustainable agriculture, *Plants*, 10(9): 1914.
<https://doi.org/10.3390/plants10091914>
- Schulze S., and Lammers M., 2020, The development of genome editing tools as powerful techniques with versatile applications in biotechnology and medicine: CRISPR/Cas9, ZnF and TALE nucleases, RNA interference, and Cre/loxP, *ChemTexts*, 7(1): 3.
<https://doi.org/10.1007/s40828-020-00126-7>
- Song H., Lin K., Hu J., and Pang E., 2018, An updated functional annotation of protein-coding genes in the cucumber genome, *Frontiers in Plant Science*, 9: 325.
<https://doi.org/10.3389/fpls.2018.00325>
- Tian J., Li Y., Hu Y., Zhong Q., Yin J., and Zhu Y., 2022, Mining the roles of cucumber DUF966 genes in fruit development and stress response, *Plants*, 11(19): 2497.
<https://doi.org/10.3390/plants11192497>
- Turek S., Płader W., Hoshi Y., Skarzyńska A., and Pawełkiewicz M., 2023, Insight into the organization of the B10v3 cucumber genome by integration of biological and bioinformatic data, *International Journal of Molecular Sciences*, 24(4): 4011.
<https://doi.org/10.3390/ijms24044011>

- Távora F., Diniz F., Rêgo-Machado C., Freitas N., Arraes F., Andrade E., Furtado L., Osiro K., Sousa N., Cardoso T., Henning L., Molinari P., Feingold S., Hunter W., Sá M., Kobayashi A., Nepomuceno A., Santiago T., and Molinari H., 2022, CRISPR/Cas-and topical RNAi-based technologies for crop management and improvement: Reviewing the risk assessment and challenges towards a more sustainable agriculture, *Frontiers in Bioengineering and Biotechnology*, 10: 913728.
<https://doi.org/10.3389/fbioe.2022.913728>
- Varshney R.K., Bohra A., Yu J., Graner A., Zhang Q., and Sorrells M.E., 2021, Designing future crops: genomics-assisted breeding comes of age, *Trends in Plant Science*, 26(6): 631-649.
<https://doi.org/10.1016/j.tplants.2021.03.010>
- Wang X., Bao K., Reddy U., Bai Y., Hammar S., Jiao C., Wehner T., Ramírez-Madera A., Weng Y., Grumet R., and Fei Z., 2018, The USDA cucumber (*Cucumis sativus* L.) collection: genetic diversity, population structure, genome-wide association studies, and core collection development, *Horticulture Research*, 5: 1-10.
<https://doi.org/10.1038/s41438-018-0080-8>
- Wang Y., Bo K., Gu X., Pan J., Li Y., Chen J., Wen C., Ren Z., Ren H., Chen X., Grumet R., and Weng Y., 2020, Molecularly tagged genes and quantitative trait loci in cucumber with recommendations for QTL nomenclature, *Horticulture Research*, 7: 1-10.
<https://doi.org/10.1038/s41438-019-0226-3>
- Xin T., Tian H., Ma Y., Wang S., Yang L., Li X., Zhang M., Chen C., Wang H., Li H., Xu J., Huang S., and Yang X., 2022, Targeted creation of new mutants with compact plant architecture using CRISPR/Cas9 genome editing by an optimized genetic transformation procedure in cucurbit plants, *Horticulture Research*, 9: uhab086.
<https://doi.org/10.1093/hr/uhab086>
- Yang W., Feng H., Zhang X., Zhang J., Doonan J., Batchelor W., Xiong L., and Yan J., 2020, Crop phenomics and high-throughput phenotyping: past decades, current challenges, and future perspectives, *Molecular Plant*, 13(2): 187-214.
<https://doi.org/10.1016/j.molp.2020.01.008>
- Zhang T., Li X., Yang Y., Guo X., Feng Q., Dong X., and Chen S., 2019, Genetic analysis and QTL mapping of fruit length and diameter in a cucumber (*Cucumis sativus* L.) recombinant inbred line (RIL) population, *Scientia Horticulturae*, 250: 214-222.
<https://doi.org/10.1016/j.scienta.2019.01.062>
- Zhang Z., Luo S., Liu Z., Wan Z., Gao X., Qiao Y., Yu J., and Zhang G., 2022, Genome-wide identification and expression analysis of the cucumber PYL gene family, *PeerJ*, 10: e12786.
<https://doi.org/10.7717/peerj.12786>

Disclaimer/Publisher's Note

The statements, opinions, and data contained in all publications are solely those of the individual authors and contributors and do not represent the views of the publishing house and/or its editors. The publisher and/or its editors disclaim all responsibility for any harm or damage to persons or property that may result from the application of ideas, methods, instructions, or products discussed in the content. Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.