

Transcriptomics in crop improvement

Sayali G. Magar, Pravin V. Jadhav, Umesh Shinde, Eknath R. Vaidya

Biotechnology Centre, Department of Agricultural Botany, Dr. Panjabrao Deshmukh Krishi Vidyapeeth, Akola, Maharashtra State, India

✉email:sayalimagar2020@gmail.com;

ABSTRACT

Transcriptomics, a powerful field within genomics, has emerged as a key tool in unraveling the genetic complexity of crop plants. By studying the entire set of RNA transcripts, transcriptomics provides insights into gene expression patterns, regulatory networks, and molecular mechanisms underlying crop traits. This chapter explores the recent advancements and applications of transcriptomics in crop improvement, highlighting its significance in addressing the challenges faced by modern agriculture. We discuss the utilization of high-throughput RNA sequencing, alternative splicing, gene discovery, regulatory networks, functional characterization, multi-omics integration, non-coding RNAs, single-cell transcriptomics, and transcriptomics-assisted breeding. By leveraging transcriptomics, researchers and breeders can accelerate the development of high-yielding, stress-tolerant, and nutritionally improved crop varieties, paving the way for sustainable agricultural practices.

Keywords- Transcriptomics, crop improvement, high-throughput RNA sequencing

I. INTRODUCTION

A. Background on Crop Improvement

Crop improvement is a field of agricultural science that focuses on enhancing the characteristics of crops to increase their productivity, nutritional value, disease resistance, and adaptability to various environmental conditions(1). Over the centuries, farmers and scientists have employed various traditional breeding techniques to develop improved crop varieties. These methods traditionally involved selecting and crossing plants with desirable traits to create offspring with a combination of desired characteristics. However, the advent of modern biotechnology and genomics has revolutionized the field of crop improvement (2,3). Genomics, the study of an organism's complete set of genes, has provided researchers with powerful tools to understand the genetic makeup of crops (4). This knowledge enables them to identify specific genes responsible for important traits, such as yield, quality, and tolerance to stresses like pests, diseases, drought, and salinity. Furthermore, genomics has paved the way for genetic engineering or genetic modification (GM) of crops(5). Through GM techniques, scientists can introduce or modify specific genes in crops to confer desirable traits directly. For example, they can introduce genes from other organisms that provide resistance to pests, diseases, or herbicides(6). This approach has resulted in the development of genetically modified crops, such as insect-resistant Bt cotton and herbicide-tolerant soybeans(1,6). In recent years, advanced genomics tools like CRISPR-Cas9 have emerged, offering precise and efficient gene editing capabilities. CRISPR-Cas9 allows scientists to edit specific genes within the genome, offering immense potential for crop improvement (7). It enables the development of crops with precise alterations, such as improved nutritional content, enhanced shelf life, or increased drought tolerance, without the need to introduce foreign genes(3,8,9). With the advent of Next-Generation Sequencing (NGS) technology, the study of RNA has expanded rapidly, uncovering the vast universe of non-coding RNAs (ncRNA) and enriching our understanding of the RNA world (10,11). Consequently, transcriptomics now encompasses a broader range of research content. Overall, biotechnology has revolutionized crop improvement by providing scientists with a deeper understanding of crop genetics, accelerating breeding through marker-assisted selection, and offering precise gene editing capabilities. These advancements hold great promise for addressing global food security challenges, increasing crop resilience, and producing crops with improved traits to meet the demands of a growing population in the face of changing environmental conditions.

B. Decoding the Transcriptome: Techniques for Studying Gene Expression Dynamics

- a. **RNA Sequencing (RNA-seq):** RNA-seq is a powerful technique used in transcriptomics research. It involves converting RNA molecules into complementary DNA (cDNA) fragments, which are then sequenced using high-throughput sequencing technologies(12,13). By analyzing the resulting sequence reads, researchers can quantify gene expression levels, identify alternative splicing events, discover novel transcripts, and characterize non-coding RNA species(10,14,15).
- b. **Microarray Analysis:** Microarrays have been widely used in transcriptomics, although they are gradually being replaced by RNA-seq(10,16). Microarrays involve the hybridization of labeled cDNA or RNA samples to a chip containing thousands of known DNA sequences(3). The intensity of hybridization signals provides information about gene expression levels. While not as comprehensive as RNA-seq, microarrays remain useful for studying gene expression patterns in specific contexts (17,18)
- c. **RT-qPCR (Reverse Transcription-quantitative PCR):** RT-qPCR is a widely used technique for quantifying the expression levels of specific genes(19). It involves converting RNA into cDNA and then using quantitative PCR to amplify and measure the amount of cDNA. RT-qPCR provides high sensitivity and specificity, making it valuable for validating gene expression findings from other transcriptomic methods(20).

II. Transcriptomics in Agriculture

- A. Nurturing Agricultural Progress through Science and Innovation:** Transcriptomics plays a vital role in agricultural research and crop improvement by providing insights into gene expression patterns associated with agronomic traits, stress responses, nutritional enhancement, crop-pathogen interactions, developmental processes, and crop-microbiome interactions (21). This knowledge contributes to the development of improved crop varieties, sustainable agricultural practices, and resilient farming systems, addressing the challenges of food security, environmental sustainability, and the changing needs of agriculture. Here are some key aspects highlighting the importance of transcriptomics in agriculture:
- a. **Crop Improvement:** Transcriptomics provides a comprehensive understanding of gene expression patterns associated with important agronomic traits, such as yield, quality, disease resistance, and stress tolerance (22). Researchers can identify genes and regulatory networks associated with desirable traits by analyzing the transcriptomes of different crop varieties or breeding lines (18). This knowledge helps in the development of improved crop varieties through marker-assisted selection (MAS) and gene editing techniques, facilitating the breeding of crops with enhanced productivity, nutritional content, and resilience to biotic and abiotic stresses (12).
 - b. **Stress Responses:** Transcriptomics aids in unraveling the molecular mechanisms underlying crop responses to various stresses, including drought, salinity, heat, cold, pests, and diseases (23,24). By studying the gene expression changes triggered by these stresses, researchers can identify stress-responsive genes and pathways (18). This information is crucial for understanding crops' physiological and biochemical adaptations to cope with adverse conditions(25). It helps in developing stress-tolerant crop varieties through breeding or genetic engineering approaches, allowing farmers to cultivate crops in challenging environments and mitigate yield losses (22,26).
 - c. **Nutritional Enhancement:** Transcriptomics is instrumental in studying the expression of genes involved in nutrient uptake, transport, and accumulation in crops (27). By examining gene expression profiles related to nutrient metabolism, researchers can identify genes responsible for traits such as enhanced nutrient content, improved bioavailability, and reduced antinutritional factors (28,29). This knowledge aids in the development of nutritionally enriched crop varieties, addressing micronutrient deficiencies (e.g., biofortification) and improving the overall nutritional quality of food crops (30).
 - d. **Crop-Pathogen Interactions:** Transcriptomics helps in deciphering the molecular interactions between crops and pathogens. By studying the changes in gene expression during pathogen infection, researchers can identify genes involved in plant defense responses, pathogen recognition, and resistance mechanisms(Pantha et al., 2021; Schumpp, 2022). This knowledge is vital for understanding host-pathogen interactions and developing crop varieties with enhanced resistance to pests and diseases, reducing the dependence on chemical pesticides and promoting sustainable agriculture practices (15,30,33).
 - e. **Developmental Processes:** Transcriptomics provides insights into the gene expression dynamics during various stages of crop development, including seed germination, plant growth, flowering, and fruit ripening (34,35). By analyzing gene expression profiles, researchers can identify genes involved in developmental processes, signaling pathways, and hormone regulation (36,37). This knowledge helps in understanding the molecular basis of crop development and optimizing agronomic practices to maximize yield and quality (38).
 - f. **Crop-Microbiome Interactions:** Transcriptomics assists in understanding the interactions between crops and beneficial microorganisms, such as plant growth-promoting rhizobacteria and mycorrhizal fungi (39,40). By studying the gene expression changes in crops in response to microbial colonization, researchers can identify genes involved in plant-microbe interactions and nutrient acquisition(40). This knowledge contributes to the development of sustainable agricultural practices that harness the potential of beneficial microorganisms to enhance crop productivity and nutrient uptake.

B. Importance of transcriptomics in gene expression

Transcriptomics is a branch of genomics that focuses on studying the transcriptome, which is the complete set of RNA molecules produced by the genes of an organism. It involves the analysis of gene expression patterns, including the identification and quantification of different types of RNA molecules, such as messenger RNA (mRNA), non-coding RNA, and small RNA molecules. Understanding gene expression is crucial because it provides insights into how genes are regulated and how they contribute to various biological processes and phenotypic traits. Here are some key aspects highlighting the importance of transcriptomics in understanding gene expression:

- a. **Gene Regulation:** Transcriptomics helps uncover the intricate mechanisms by which genes are turned on (activated) or off (repressed) in different cells, tissues, and developmental stages(24). By studying the RNA transcripts, researchers can identify which genes are actively transcribed and at what levels, providing valuable information about gene regulatory networks and pathways(41).
- b. **Functional Annotation of Genomes:** Transcriptomics assists in annotating genomes by identifying the protein-coding genes and non-coding RNAs present in an organism. It allows researchers to determine the precise boundaries of genes, locate alternative splicing events, and identify different isoforms of genes (42). This knowledge helps in understanding the functional elements within the genome and provides a foundation for further functional studies.
- c. **Disease Research:** Transcriptomics plays a significant role in understanding the molecular mechanisms underlying diseases (18,42). By comparing gene expression profiles between healthy and diseased tissues, researchers can identify differentially expressed genes associated with specific diseases or conditions (21). This information helps in unraveling disease pathways, identifying potential biomarkers, and discovering therapeutic targets for drug development.
- d. **Environmental responses:** In different environmental conditions, transcriptomics allows researchers to identify genes and pathways involved in the response to specific environmental cues (38). This information can shed light on the molecular mechanisms underlying an organism's adaptation, survival, and defense strategies. Transcriptomics also aids in the identification of stress-responsive genes and regulatory networks, providing insights into the genetic basis of stress tolerance and resilience(43). Additionally, transcriptomics enables the characterization of gene expression changes in response to environmental factors, helping to unravel the intricate interactions between organisms and their environment

(44). This knowledge can contribute to the development of strategies for mitigating environmental challenges, improving crop resilience, and understanding the impacts of environmental changes on ecosystems(45). Overall, transcriptomics plays a crucial role in advancing our understanding of the molecular mechanisms driving environmental responses and adaptation in organisms across various domains of life.

III. Type of RNA-Seq

RNA-sequencing (RNA-seq) has revolutionized the field of genomics by enabling comprehensive analysis of the transcriptome. By capturing the entirety of RNA molecules present in a sample, RNA-seq provides valuable insights into gene expression patterns, alternative splicing events, and the identification of novel RNA species. This chapter aims to explore the principle behind RNA-seq and highlight its key advantages in deciphering the complexity of gene regulation. It encompasses various approaches and techniques for studying gene expression at the transcriptome level. Here are some notable types of RNA-Seq, Figure 1 illustrates the general protocol for RNA-seq, along with some notable types of RNA-Seq:

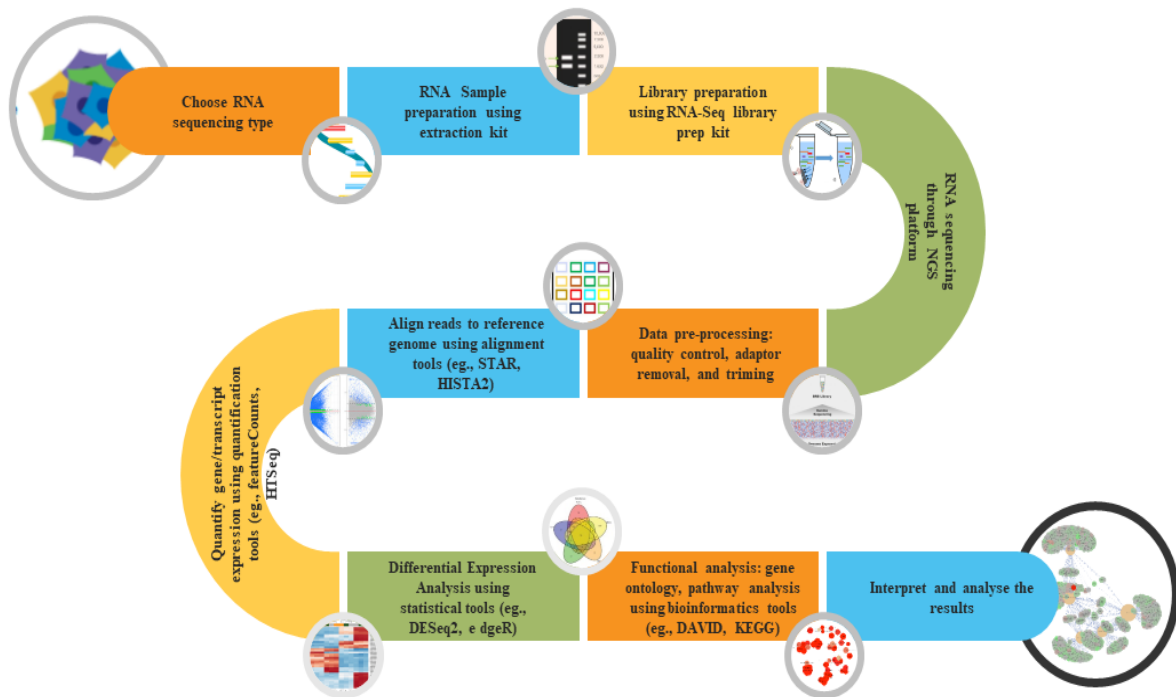


Figure 1: Schematic presentation of steps involved in Transcriptomic approach

- A. Bulk RNA-seq:** Bulk RNA-seq involves sequencing the entire RNA population from a sample, providing an overview of gene expression levels across the transcriptome (46). It provides an overview of the transcriptome, which encompasses all the RNA molecules present in a particular cell, tissue, or organism(47). It is widely used to compare gene expression between different conditions or tissues.
- RNA Extraction:** Collect the sample of interest (e.g., cells, tissues) and preserve it appropriately to maintain RNA integrity. Extract total RNA using a suitable method (e.g., TRIzol, column-based kits) according to the manufacturer's instructions. Assess the RNA quality and quantity using spectrophotometry (e.g., NanoDrop) and electrophoresis (e.g., Agilent Bioanalyzer).
 - Library Preparation:** Purify and concentrate the extracted RNA, if necessary, using RNA cleanup kits. Optionally, perform rRNA depletion or mRNA enrichment to selectively target mRNA molecules. Perform reverse transcription to convert RNA into complementary DNA (cDNA). Use oligo-dT primers or random primers for cDNA synthesis. Fragment the cDNA into shorter segments, typically by enzymatic fragmentation or sonication. Add sequencing adapters to the cDNA fragments, including barcode sequences for sample multiplexing.
 - Library Quality Control:** Assess the quality and size distribution of the prepared cDNA library using methods such as qPCR, Bioanalyzer, or TapeStation. Quantify the library concentration accurately using qPCR or fluorometric assays (e.g., Qubit).
 - Sequencing:** Submit the validated and quantified cDNA library for sequencing on a suitable platform (e.g., Illumina HiSeq, NovaSeq). Choose appropriate sequencing parameters (e.g., read length, sequencing depth) based on the research goals and budget.
 - Data Analysis:** Here are some commonly used bioinformatics tools and software for the analysis of bulk RNA-seq data:
 - **Alignment and Mapping:**
 - **STAR:** A fast and accurate RNA-seq aligner that maps sequencing reads to a reference genome.

- HISAT2: A splice-aware aligner that aligns RNA-seq reads to a reference genome, taking into account alternative splicing.
- Bowtie2: A widely used aligner for mapping short reads to a reference genome.
- Read Quantification and Transcript Assembly:
 - featureCounts: A tool that assigns reads to genes or other genomic features, providing read counts for each feature.
 - HTSeq: A tool for counting reads mapped to features in a genomic annotation.
 - StringTie: A software for transcript assembly and quantification, providing gene and isoform expression levels.
- Differential Expression Analysis:
 - DESeq2: A widely used R package for differential gene expression analysis, taking into account the sample-specific variability.
 - edgeR: An R package for differential expression analysis, particularly useful for small sample sizes.
 - limma: An R package for linear modeling and differential expression analysis, often used for RNA-seq data from complex experimental designs.
- Functional Enrichment Analysis:
 - Gene Ontology (GO) Enrichment Analysis: Tools like GOrse, topGO, or clusterProfiler perform enrichment analysis to identify overrepresented biological processes, molecular functions, or cellular components associated with differentially expressed genes.
 - KEGG Pathway Analysis: Tools such as KOBAS or clusterProfiler can analyze gene sets for enrichment in specific KEGG pathways.
 - DAVID: A web-based tool that allows functional annotation and enrichment analysis of gene lists.
- Visualization and Data Exploration:
 - R/Bioconductor: A collection of open-source R packages, including ggplot2, pheatmap, and ComplexHeatmap, for generating publication-quality plots, heatmaps, and visualizations.
 - Integrative Genomics Viewer (IGV): A desktop application for interactive visualization and exploration of genomic data, including RNA-seq data.

These are just a few examples of the bioinformatics tools commonly used in the analysis of bulk RNA-seq data. Depending on specific research goals and preferences, there are many other tools and software available that offer similar functionalities for processing, analysis, and interpretation of RNA-seq data. Bulk RNA-seq has broad applications in transcriptomics research. It enables the identification of differentially expressed genes, the discovery of novel transcripts, and the characterization of gene expression patterns (48). It is often used to compare gene expression profiles between healthy and diseased tissues, different developmental stages, or samples exposed to different experimental conditions. By providing a comprehensive view of gene expression across the transcriptome, bulk RNA-seq offers valuable insights into the molecular mechanisms underlying various biological processes and diseases.

- B. Single-cell RNA-seq:** Single-cell RNA-seq allows for the profiling of gene expression in individual cells, providing insights into cellular heterogeneity and identifying rare cell types (Sunaga-Franze et al., 2022). This technique enables the characterization of gene expression patterns at a single-cell resolution(50,51). Unlike bulk RNA-seq, which provides an average gene expression profile for a population of cells, scRNA-seq allows researchers to study gene expression patterns at a single-cell resolution(47). The scRNA-seq workflow involves multiple steps to capture and sequence RNA from individual cells:
- a. Cell Isolation: Cells of interest are isolated from a tissue or sample using various methods such as fluorescence-activated cell sorting (FACS), microfluidics-based techniques, or manual picking under a microscope. The goal is to obtain a population of single, viable cells for further analysis.
 - b. Cell Lysis and RNA Extraction: The isolated cells undergo lysis to release their cellular contents, including RNA. RNA extraction is then performed to isolate the RNA molecules from each individual cell. Various RNA extraction methods can be employed, such as column-based purification or magnetic bead-based protocols.
 - c. Reverse Transcription and cDNA Amplification: The isolated RNA is converted into complementary DNA (cDNA) through reverse transcription. This process involves using reverse transcriptase enzymes and primers to generate DNA molecules that are complementary to the RNA templates. Additionally, a unique molecular identifier (UMI) is often incorporated during reverse transcription to tag individual cDNA molecules, allowing for accurate quantification of gene expression later in the analysis.
 - d. Library Preparation and Sequencing: The cDNA is then prepared into sequencing libraries by fragmenting the cDNA, attaching sequencing adapters, and amplifying the library using PCR. Single-cell barcoding is typically incorporated during library preparation to label each cDNA fragment with a cell-specific barcode. This barcode enables the identification and demultiplexing of sequences from different cells during data analysis. The prepared libraries are then subjected to high-throughput sequencing, generating short sequence reads for each cDNA fragment.
 - e. Data Analysis: The generated sequencing data is processed and analyzed bioinformatically to extract meaningful information. This involves aligning the sequencing reads to a reference genome or transcriptome, quantifying gene expression levels, and identifying differentially expressed genes. The expression profiles of individual cells are used to cluster cells based on similarities in gene expression patterns and to infer cell types or states within the sample. Advanced computational methods and algorithms are employed to handle the large-scale, high-dimensional data generated by scRNA-seq experiments.
- Data Preprocessing and Quality Control:
 - Cell Ranger: A widely used pipeline provided by 10x Genomics for preprocessing and quality control of single-cell RNA-seq data.
 - Seurat: An R package for single-cell RNA-seq data analysis, including data preprocessing, quality control, and normalization.
 - Scanpy: A Python package for single-cell analysis that provides tools for preprocessing, filtering, and quality control.

- Dimensionality Reduction and Visualization:
 - Principal Component Analysis (PCA): Tools such as Seurat, Scanpy, or scikit-learn provide PCA implementation for dimensionality reduction and visualization of single-cell RNA-seq data.
 - t-SNE and UMAP: Algorithms like t-Distributed Stochastic Neighbor Embedding (t-SNE) and Uniform Manifold Approximation and Projection (UMAP) are commonly used for visualizing and clustering single-cell RNA-seq data.
 - Cell Clustering and Identification:
 - Seurat: Seurat provides various algorithms for cell clustering, including graph-based clustering, hierarchical clustering, and k-means clustering.
 - Scanpy: Scanpy offers clustering methods like Louvain clustering and Leiden clustering for cell type identification and clustering analysis.
 - Differential Expression Analysis:
 - Seurat: Seurat provides tools for identifying differentially expressed genes between cell clusters or conditions, such as the FindMarkers function.
 - MAST: A popular R package for differential expression analysis in single-cell RNA-seq data, particularly suitable for small sample sizes.
 - Trajectory Analysis:
 - Monocle: An R package for ordering and visualizing single-cell RNA-seq data along developmental trajectories and pseudotime analysis.
 - Slingshot: A tool for trajectory inference and visualization in single-cell RNA-seq data.
 - Functional Enrichment Analysis:
 - clusterProfiler: An R package for gene set enrichment analysis, including functional annotation and pathway analysis of single-cell RNA-seq data.
 - Visualization and Data Exploration:
 - R/Bioconductor: Various R packages, such as ggplot2, ComplexHeatmap, and ggpubr, are used for generating visualizations and plots from single-cell RNA-seq data.
 - Loupe Cell Browser: A visualization software provided by 10x Genomics for interactive exploration and visualization of single-cell RNA-seq data. There are many other specialized tools, pipelines, and software available depending on specific analysis needs and preferences.
- C. Total RNA-seq:** Total RNA-seq is a transcriptomics technique that aims to capture and sequence the complete RNA content present in a sample. Unlike other RNA-seq methods that target specific RNA subsets, such as mRNA or small non-coding RNAs, Total RNA-seq provides a comprehensive view of the transcriptome by capturing all RNA species, including both coding and non-coding RNA molecules (42,52). The workflow for Total RNA-seq involves several key steps:
- a. RNA Extraction: Total RNA is extracted from the sample of interest, which can be derived from cells, tissues, or organisms. The extraction process involves breaking open cells to release their contents, including RNA molecules.
 - b. RNA Integrity Assessment: The extracted RNA is assessed for its quality and integrity using various methods, such as agarose gel electrophoresis or automated RNA quality analyzers. This step ensures that the RNA is intact and suitable for downstream analysis.
 - c. Library Preparation: The extracted RNA undergoes library preparation, where the RNA molecules are converted into complementary DNA (cDNA) fragments. This is typically achieved through reverse transcription, which uses reverse transcriptase enzymes and random primers or oligo-dT primers to convert RNA into cDNA.
 - d. Fragmentation and Adapter Ligation: The cDNA fragments are then fragmented into smaller pieces to facilitate sequencing. Sequencing adapters are added to the cDNA fragments, which allow for binding to the sequencing platform.
 - e. Sequencing: The prepared cDNA library is subjected to high-throughput sequencing, generating short sequence reads. The sequencing platforms can vary, including Illumina sequencing platforms, which are commonly used for Total RNA-seq.
 - f. Data Analysis: The generated sequencing data is processed and analyzed bioinformatically. The reads are aligned to a reference genome or transcriptome to determine the origin of each read. This step allows for the identification of known transcripts and the discovery of novel transcripts. Additionally, quantification methods are employed to determine the expression levels of the identified transcripts. The bioinformatics tools and software commonly used for the analysis of total RNA-seq data largely overlap with those employed for bulk RNA-seq analysis. Many of these tools can also be applied for analyzing total RNA-seq, and additional tools are available specifically for alternative splicing analysis.
- Alternative Splicing Analysis:
- rMATS: A tool for the identification and quantification of alternative splicing events, including skipped exon, alternative 5' or 3' splice site, and more.
 - SUPPA: A tool for alternative splicing analysis, focusing on differentially spliced events between conditions.
- D. mRNA-Seq:** mRNA-seq specifically targets and sequences messenger RNA (mRNA) molecules, which carry the genetic information for protein synthesis(53,54). It focuses on protein-coding transcripts and provides information about gene expression levels and isoform variations. The bioinformatics tools, software, and workflow commonly used for the analysis of mRNA-seq data are akin to those employed in bulk RNA-seq analysis.
- E. Small RNA-Seq:** Small RNA-seq is designed to capture and sequence small non-coding RNA molecules, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs)(55). It helps in understanding the regulatory roles of these small RNAs in gene expression and cellular processes (56,57). The workflow for Small RNA-seq involves several key steps to capture and sequence small non-coding RNA molecules, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs). Here is a general outline of the Small RNA-seq workflow:

- a. RNA Extraction: Total RNA is extracted from the sample of interest, which can be derived from cells, tissues, or biofluids. The extraction process is designed to capture small RNA molecules, including miRNAs and siRNAs. Various RNA extraction methods can be employed, such as column-based purification or phenol-chloroform extraction.
 - b. Size Selection: The extracted RNA is subjected to size selection to enrich for small RNA molecules. This is typically achieved using gel electrophoresis or size-exclusion columns to remove larger RNA species, such as ribosomal RNA (rRNA) and messenger RNA (mRNA).
 - c. 5' and 3' Adaptor Ligation: Adaptors containing specific sequences are ligated to the small RNA molecules. This process involves the addition of a 5' adaptor to one end and a 3' adaptor to the other end of the small RNA molecules. These adaptors provide the necessary sequences for subsequent steps in library preparation.
 - d. Reverse Transcription and PCR Amplification: The ligated small RNA molecules undergo reverse transcription, converting the RNA into complementary DNA (cDNA) using reverse transcriptase enzymes and specific primers. The cDNA is then amplified using polymerase chain reaction (PCR) to generate sufficient material for sequencing.
 - e. Library Preparation: The amplified cDNA is prepared into sequencing libraries by adding sequencing adapters. These adapters contain specific sequences required for the binding of the cDNA fragments to the sequencing platform. Unique molecular identifiers (UMIs) may also be incorporated during library preparation to tag individual cDNA molecules and enable accurate quantification later in the analysis.
 - f. Sequencing: The prepared library is subjected to high-throughput sequencing using platforms such as Illumina sequencing machines. Sequencing generates short sequence reads that represent the small RNA molecules present in the sample.
 - g. Data Analysis:
 - Adapter Trimming and Quality Control:
 - Cutadapt: A tool for removing adapter sequences from sequencing reads.
 - FastQC: A quality control tool for assessing the quality and composition of sequencing reads.
 - Alignment and Annotation:
 - Bowtie: A popular aligner for aligning short sequencing reads to a reference genome or small RNA databases.
 - miRBase: A comprehensive database of known microRNAs, often used for annotating and identifying miRNAs.
 - Rfam: A database of RNA families that can be utilized for annotating other small RNA species.
 - Expression Quantification:
 - featureCounts: A tool for assigning reads to genomic features and quantifying their expression levels, such as miRNAs or other small RNA molecules.
 - HTSeq: A Python library that performs read counting by assigning reads to genomic features based on the genomic position.
 - Differential Expression Analysis:
 - DESeq2: A popular R package for differential expression analysis of small RNA-seq data, taking into account sample-specific variability.
 - edgeR: An R package for differential expression analysis, particularly useful for small sample sizes.
 - Functional Enrichment Analysis:
 - miRPathDB: A database and web server for analyzing the biological functions and pathways associated with differentially expressed miRNAs.
 - DIANA-miRPath: A tool for miRNA target prediction and functional enrichment analysis, providing insights into the potential target genes and biological processes regulated by miRNAs.
 - Visualization and Data Exploration:
 - R/Bioconductor: Various R packages, such as ggplot2, ComplexHeatmap, or pheatmap, are commonly used for generating visualizations and plots from Small RNA-seq data.
 - Integrative Genomics Viewer (IGV): A desktop application for interactive visualization and exploration of genomic data, including Small RNA-seq data.
- F. Ribosome profiling (Ribo-seq):** Ribo-seq involves sequencing the RNA fragments protected by ribosomes during translation (58,59). It provides insights into actively translated mRNA regions, allowing for the investigation of translational efficiency and ribosome occupancy (60–62). The workflow for Ribosome Profiling (Ribo-seq), also known as ribosome foot-printing, involves several key steps to capture and sequence the ribosome-protected fragments, providing insights into translation dynamics and protein synthesis (63,64). Here is a general outline of the Ribo-seq workflow:
- a. Sample Preparation: Cells or tissues of interest are collected and lysed to release the cellular contents, including ribosomes and mRNA.
 - b. Ribosome Capture and Fragmentation: Ribosomes are captured from the lysate using either chemical cross-linking or nuclease treatment. The captured ribosomes protect the mRNA fragments that are actively being translated. The ribosome-mRNA complexes are then treated with ribonucleases to digest the unprotected mRNA, resulting in the isolation of ribosome-protected fragments (RPFs).
 - c. RNA Extraction and Library Preparation: The RPFs, along with other RNA species, are extracted from the sample. Total RNA extraction methods, such as phenol-chloroform extraction or column-based purification, can be employed. The extracted RPFs undergo size selection to enrich for the desired fragment length (typically around 28–30 nucleotides). Subsequently, sequencing libraries are prepared from the selected RPFs, including steps such as adapter ligation and reverse transcription.
 - d. Sequencing: The prepared library is subjected to high-throughput sequencing using platforms such as Illumina sequencing machines. This generates short sequence reads that represent the ribosome-protected fragments.
 - e. Data Analysis: The generated sequencing data undergoes bioinformatics analysis to infer translation dynamics and identify translated regions:

- Read Trimming: The sequencing reads are processed to remove adapter sequences and low-quality bases.
 - Alignment to Reference Genome: The trimmed reads are aligned to a reference genome or transcriptome, allowing mapping of the ribosome-protected fragments.
 - Read Quantification: The aligned reads are used to quantify the abundance of RPFs in specific regions of the mRNA transcripts, reflecting the density of ribosomes at different sites.
 - Ribosome Profiling Analysis: Various computational methods, such as RiboTools, RiboTaper, or ORFscore, are used to analyze the ribosome profiling data. This includes identifying translated regions (open reading frames), calculating ribosome density profiles, detecting ribosome pauses or frameshifts, and inferring translation efficiency.
 - Integration with Transcriptomics: Ribo-seq data can be integrated with transcriptomic data, such as RNA-seq, to study the relationship between transcription and translation and gain insights into gene regulation and protein synthesis dynamics. The Ribo-seq workflow allows researchers to investigate translation dynamics and protein synthesis at a genome-wide scale. By analyzing ribosome-protected fragments, Ribo-seq provides valuable information on translation initiation sites, translation efficiency, ribosome occupancy, and other aspects of protein synthesis. This technique has broad applications in understanding gene regulation, studying translational control mechanisms, and characterizing the coding potential of the genome.
- G. Nascent RNA-seq:** Nascent RNA-seq captures newly transcribed RNA molecules, providing insights into the transcriptional dynamics and regulation of gene expression (62). It allows the study of immediate early responses and temporal gene expression patterns. Here is a general outline of the Nascent RNA-seq workflow:
- a. Cell Fixation and Isolation: Cells of interest are fixed using cross-linking agents, such as formaldehyde, to preserve the RNA-protein interactions and freeze the transcriptional state at the time of fixation. Fixed cells are then lysed to release the cellular contents.
 - b. RNA Extraction: Total RNA extraction methods, such as phenol-chloroform extraction or column-based purification, are employed to isolate the RNA from the fixed cells. This step ensures the capture of both nascent and pre-existing RNA molecules.
 - c. Labeling and Isolation of Nascent RNA: Newly transcribed RNA molecules are labeled with a modified nucleotide analog, such as 5-ethynyluridine (EU) or 5-bromouridine (BrU). These modified nucleotides are incorporated into nascent RNA during transcription. Nascent RNA molecules are then specifically isolated and enriched using techniques like immunoprecipitation with antibodies against the modified nucleotides or biotin-based pulldown methods.
 - d. RNA Fragmentation and Library Preparation: The isolated nascent RNA is fragmented into smaller pieces to facilitate library preparation. The fragmented RNA undergoes reverse transcription to convert RNA into complementary DNA (cDNA). Next, sequencing adapters are ligated to the cDNA fragments, enabling them to be sequenced.
 - e. Sequencing: The prepared library is subjected to high-throughput sequencing using platforms such as Illumina sequencing machines. Sequencing generates short sequence reads representing the nascent RNA molecules.
 - f. Data Analysis: The generated sequencing data undergoes bioinformatics analysis to gain insights into transcriptional dynamics and gene regulation:
 - Read Trimming: The sequencing reads are processed to remove adapter sequences and low-quality bases.
 - Alignment to Reference Genome: The trimmed reads are aligned to a reference genome or transcriptome, allowing mapping of the nascent RNA molecules to specific genomic loci.
 - Quantification of Transcriptional Activity: The aligned reads are used to quantify the transcriptional activity at different genomic regions, such as promoters, enhancers, or gene bodies, by measuring read density or counting nascent RNA molecules.
 - Differential Transcription Analysis: Statistical methods can be applied to compare the nascent RNA levels between different conditions or experimental groups to identify differentially transcribed genes or regulatory regions.
 - g. Integration with other Omics Data: Nascent RNA-seq data can be integrated with other omics data, such as chromatin immunoprecipitation sequencing (ChIP-seq) or transcription factor binding data, to study the relationship between transcriptional regulation, chromatin states, and gene expression dynamics. Each type of transcriptomics technique has its unique strengths and applications, contributing to our understanding of gene expression, cellular processes, and the regulation of biological systems.

IV. The platforms for RNA Seq

- A. Illumina Sequencing:** Illumina is a leading company in the field of sequencing technologies. They offer several platforms for RNA-seq, including:
- HiSeq Series: Illumina HiSeq instruments, such as HiSeq 2500, HiSeq 3000/4000, and HiSeq X Series, provide high-throughput sequencing with excellent data quality and accuracy.
 - NovaSeq Series: The NovaSeq platforms, including NovaSeq 6000 and NovaSeq 5000, are designed for high-throughput sequencing at a larger scale, offering cost-effective and flexible options for RNA-seq experiments.
 - MiSeq: Illumina MiSeq is a benchtop sequencer suitable for smaller-scale RNA-seq projects. It provides rapid turnaround time and is often used for targeted RNA sequencing or smaller transcriptome studies. (Source: Illumina (<https://www.illumina.com>))
- B. Ion Torrent Sequencing:** Ion Torrent, developed by Thermo Fisher Scientific, is another widely used platform for RNA-seq. It is based on semiconductor sequencing technology, which detects changes in pH during nucleotide incorporation. The platforms offered by Ion Torrent include:
- Ion Proton: Ion Proton sequencers provide fast and scalable sequencing with a focus on speed and cost-effectiveness.
 - Ion GeneStudio: The Ion GeneStudio S5 and S5 Prime systems offer scalable and versatile solutions for RNA-seq, with the ability to process multiple samples simultaneously. (Source: Thermo Fisher Scientific (<https://www.thermofisher.com>)).

- C. **PacBio Sequencing:** Pacific Biosciences (PacBio) offers long-read sequencing technologies that are particularly useful for studying RNA isoforms, alternative splicing, and full-length transcript sequencing. The PacBio platforms used for RNA-seq include:
- Sequel System: The PacBio Sequel System enables long-read sequencing and provides high-quality data for comprehensive transcriptome analysis. (Source: Pacific Biosciences (<https://www.pacb.com>))
- D. **Oxford Nanopore Sequencing:** Oxford Nanopore Technologies specializes in nanopore-based sequencing, which allows long-read sequencing of RNA molecules. The platform offered by Oxford Nanopore for RNA-seq is:
- MinION: The MinION device is a portable and affordable sequencer that provides real-time RNA sequencing. It is suitable for fieldwork and rapid data generation. (Source: Oxford Nanopore Technologies (<https://nanoporetech.com>)). It's important to note that the field of sequencing technologies is rapidly evolving, and new platforms and advancements may arise over time.

V. Applications of RNA-seq in crop Transcriptomics

- A. **Gene Expression Profiling:** RNA-seq enables the comprehensive profiling of gene expression in crops. By comparing transcriptomes across different tissues, developmental stages, or environmental conditions, researchers can gain insights into the genes and pathways involved in important agronomic traits(8,65). This information can aid in understanding crop development, response to stress, and the regulation of key metabolic processes.
- B. **Differential Gene Expression Analysis:** RNA-seq allows for the identification of differentially expressed genes (DEGs) between contrasting conditions. By comparing transcriptomes of, for example, disease-resistant and susceptible cultivars or plants under stress and control conditions, researchers can identify genes that play crucial roles in defense mechanisms or stress tolerance(20,21). This knowledge can guide breeding efforts to develop improved crop varieties with enhanced resistance and adaptability.
- C. **Alternative Splicing Analysis:** Alternative splicing is a post-transcriptional process that generates multiple mRNA isoforms from a single gene, thereby expanding the functional diversity of the transcriptome (53). RNA-Seq provides a powerful approach to studying alternative splicing events in crops. By identifying and quantifying different splice variants, researchers can unravel the complex regulatory mechanisms underlying traits such as flowering time, fruit development, and stress responses.
- D. **Non-Coding RNA Analysis:** RNA-seq enables the discovery and characterization of non-coding RNAs (ncRNAs) in crop plants. ncRNAs, such as microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), have been shown to play critical roles in regulating gene expression and various biological processes(57). By identifying and profiling these ncRNAs, researchers can gain insights into their involvement in crop development, stress responses, and other important agronomic traits.
- E. **Transcriptomic Annotation and Gene Discovery:** RNA-seq provides an opportunity to improve the annotation of crop genomes by identifying previously unannotated genes and transcripts. By aligning RNA-seq reads to the reference genome, novel genes, alternative transcription start sites, or untranslated regions (UTRs) can be discovered(24,35,48). This information aids in expanding our knowledge of the crop transcriptome and facilitates the identification of candidate genes for further functional characterization.
- F. **Pathway Analysis and Functional Annotation:** RNA-seq data can be leveraged to infer biological pathways and assign functional annotations to genes in crop plants. Researchers can unravel the molecular mechanisms underlying important agronomic traits by integrating transcriptomic data with existing pathway databases(35). This knowledge can assist in the development of crop improvement strategies, such as marker-assisted breeding or genetic engineering approaches.
- G. **Comparative Transcriptomics:** RNA-seq allows for comparative transcriptomic studies between different crop species or varieties. By comparing transcriptomes, researchers can identify conserved and divergent gene expression patterns, providing insights into the genetic basis of trait variations(22,66). Comparative transcriptomics can also shed light on the evolutionary relationships between different crop species, aiding in crop domestication studies and the identification of candidate genes for crop improvement.
- H. **Integration of transcriptomics data into breeding programs:** Transcriptomics-assisted breeding involves the integration of transcriptomics data into traditional breeding programs to enhance the efficiency and precision of crop improvement efforts. By incorporating transcriptomics data, which provides a comprehensive view of gene expression patterns, breeders can gain valuable insights into the genetic basis of agronomically important traits, identify potential gene candidates, and make informed decisions during the breeding process(22,67). Here is an elaboration on the integration of transcriptomics data into breeding programs:
- Trait Identification and Understanding: Transcriptomics data can aid in the identification and understanding of the genes and regulatory networks associated with specific traits of interest (65). By comparing the transcriptomes of different genotypes or contrasting phenotypes, breeders can identify differentially expressed genes and pathways related to desired traits, providing valuable information about the underlying molecular mechanisms.
 - Marker Development: Transcriptomics data can guide the development of molecular markers linked to target traits. By identifying differentially expressed genes associated with specific traits, breeders can select candidate genes as markers for marker-assisted breeding (MAB) programs. These markers can be used for efficient trait selection in breeding populations, reducing the time and resources required for traditional phenotypic evaluation.
 - Selection and Genomic Prediction: Transcriptomics data can be integrated with genomic prediction models to improve the accuracy of trait prediction and selection in breeding programs. By incorporating gene expression information, breeders can refine genomic prediction models and enhance the accuracy of predicting phenotypes based on genotypic information. This enables the early selection of superior genotypes with desirable trait performance.
 - Targeted Gene Manipulation: Transcriptomics data can aid in targeted gene manipulation through genetic engineering or gene editing techniques. By identifying key genes or pathways associated with target traits, breeders can use the

transcriptomics data to guide the selection of candidate genes for gene manipulation experiments. This targeted approach allows for more precise and efficient trait modification in crops.

- **Novel Trait Discovery:** Transcriptomics data can lead to the discovery of novel genes and pathways associated with desirable traits. By analyzing the transcriptomes of different germplasm resources or wild relatives, breeders can identify novel genetic resources and potential targets for trait improvement. This information can expand the genetic diversity available for breeding programs and provide new avenues for crop improvement. Case study regarding the crops improvement are given below (Table 1).

I. How Transcriptomics Data is Integrated?

Transcriptomics data is integrated into breeding programs through several steps:

- a. **Data Generation:** Transcriptomic data is generated using high-throughput sequencing technologies, such as RNA sequencing (RNA-Seq), to profile the gene expression patterns of different genotypes or phenotypically diverse populations.
- b. **Data Analysis:** Bioinformatics tools and pipelines are used to analyze the transcriptomics data, identify differentially expressed genes, and gain insights into the molecular mechanisms underlying specific traits.
- c. **Trait Identification:** Transcriptomics data aids in identifying differentially expressed genes or pathways associated with target traits, providing valuable information about the genetic basis of desired agronomic traits.
- d. **Marker Development:** Differentially expressed genes can serve as candidate markers for marker-assisted breeding (MAB) programs, allowing breeders to select individuals with the desired trait by genotyping rather than relying solely on phenotypic evaluation.
- e. **Selection and Genomic Prediction:** Transcriptomics data is integrated with genomic prediction models to enhance the accuracy of predicting phenotypes based on genotypic information, enabling breeders to select superior genotypes at early breeding stages.
- f. **Targeted Gene Manipulation:** Transcriptomics data guides the selection of candidate genes or pathways for targeted gene manipulation through genetic engineering or gene editing techniques, enabling breeders to introduce or modify specific traits in crops.

VI. Future perspectives and challenges

Emerging technologies and trends in crop transcriptomics: Future perspectives in crop transcriptomics are shaped by emerging technologies and trends that have the potential to revolutionize our understanding of gene expression regulation and its application in crop improvement. These developments hold great promise for enhancing crop productivity, resilience, and quality. Here, we discuss some of the exciting emerging technologies and trends in crop transcriptomics, as well as the challenges associated with their implementation in crop research:

- A. Single-Cell Transcriptomics in Crops:** Single-cell transcriptomics is gaining momentum in crop research, enabling the characterization of gene expression at the individual cell level. This technology provides insights into cellular heterogeneity, cell-type-specific gene expression patterns, and cellular dynamics during plant development and in response to environmental cues. Applying single-cell transcriptomics to crops will unravel the complexities of gene regulation within different cell types and developmental stages, leading to a deeper understanding of cellular processes and trait regulation in crop plants.
- B. Spatial Transcriptomics in Crop Tissues:** Spatial transcriptomics integrates spatial information with gene expression data, enabling the visualization and analysis of gene expression patterns within specific tissue sections in crops. By mapping gene expression in specific tissue compartments, such as roots, leaves, or floral organs, researchers can study tissue-specific gene regulation and spatial gene expression patterns. Spatial transcriptomics will provide valuable insights into tissue-specific processes, developmental gradients, and interactions among different cell types within crop tissues.
- C. Long-Read Sequencing in Crop Transcriptomics:** Long-read sequencing technologies, such as PacBio and Oxford Nanopore sequencing, offer the ability to sequence full-length transcripts, including isoforms and alternative splicing events, in crops. This technology provides more accurate and comprehensive transcriptomic data, enabling the identification of novel transcripts, alternative splicing events, and transcriptomic variations. Long-read sequencing will enhance our understanding of transcript diversity, isoform-specific gene regulation, and the functional significance of alternative splicing in crop plants.
- D. Integration of Multi-Omics Data in Crop Research:** Integrating transcriptomics data with other -omics technologies, such as genomics, proteomics, and metabolomics, is an emerging trend in crop research. This integrative approach allows for a comprehensive understanding of the functional interactions and regulatory networks underlying crop traits. By combining information from different omics layers, researchers can gain a holistic view of molecular processes, identify key regulatory factors, and unravel complex trait architectures in crops. This integration will enhance our ability to predict and manipulate complex traits in crop improvement.

VII. Challenges in implementing these emerging technologies and trends in crop transcriptomics include:

- A. Data Handling and Analysis:** The increasing volume and complexity of transcriptomics data generated from crops require robust bioinformatics tools and computational approaches for data storage, processing, analysis, and interpretation. Developing scalable and efficient algorithms, pipelines, and databases to handle big data in crop transcriptomics is a significant challenge.
- B. Data Integration and Cross-Omics Analysis:** Integrating transcriptomics data with other -omics datasets require advanced computational approaches to identify meaningful associations, decipher complex regulatory networks, and understand the functional relationships among different omics layers in crops. Developing robust methods for data integration, network analysis, and multi-omics modeling is crucial for exploiting the full potential of integrated omics approaches in crop research.

- C. **Crop-Specific Considerations:** Each crop has its own unique genomic characteristics, gene regulatory networks, and environmental interactions. Adapting and optimizing transcriptomics technologies and analysis methods to the specific needs and challenges of different crop species is essential. Crop-specific reference genomes, transcriptome annotations, and knowledge databases are critical for accurate interpretation and application of transcriptomics data in crop improvement.
- D. **Experimental Design and Validation:** Designing experiments that capture the relevant biological variation, replicate conditions, and address specific research questions is crucial for the generation of high-quality transcriptomics data in crops. Validating the findings from transcriptomics studies using independent methods, such as qPCR or functional assays, is essential for ensuring the reliability and biological relevance of transcriptomic markers and findings.

VIII. Overcoming Challenges in data analysis and Interpretation

Overcoming challenges in data analysis and interpretation is crucial for utilizing transcriptomics data effectively in crop improvement. Here, we discuss some strategies and approaches to address these challenges in the context of crop improvement:

- A. **Robust Bioinformatics Pipelines:** Developing robust bioinformatics pipelines is essential for handling and analyzing transcriptomics data in crops. These pipelines should include steps for quality control, read alignment, transcript quantification, differential gene expression analysis, and functional annotation. Employing standardized and reproducible analysis workflows ensures data consistency and facilitates comparison across different studies and crop species.
- B. **Advanced Statistical Methods:** Transcriptomics data analysis often involves complex statistical modeling due to the high-dimensional nature of the data and the presence of multiple confounding factors. Utilizing advanced statistical methods, such as linear models, mixed-effects models, or generalized linear models, can help account for experimental design, batch effects, and other sources of variation. Additionally, incorporating appropriate multiple testing correction methods, such as false discovery rate (FDR) control, helps identify differentially expressed genes with high confidence.
- C. **Integration with Other-Omics Data:** Integrating transcriptomics data with other-omics datasets, such as genomics, proteomics, or metabolomics, can provide a more comprehensive understanding of the biological processes underlying crop traits. This integration allows for the identification of key regulators, signaling pathways, and functional interactions. Advanced bioinformatics tools and algorithms, such as pathway and network analysis, can be employed to extract meaningful biological insights from integrated -omics data.
- D. **Functional Annotation and Gene Ontology Analysis:** Functional annotation of transcriptomics data is essential for assigning biological meaning to gene expression changes. Annotating genes with functional information, such as Gene Ontology (GO) terms or enrichment analysis, helps elucidate the biological processes, molecular functions, and cellular components associated with differentially expressed genes. Functional annotation facilitates the identification of key biological pathways and candidate genes involved in specific agronomic traits.
- E. **Cross-Validation and External Validation:** Cross-validation techniques can be employed to assess the robustness and reproducibility of transcriptomics results. Splitting the dataset into training and validation sets allows for internal validation of the transcriptomic markers and models generated. External validation can be achieved by comparing transcriptomics findings with independent experimental data or existing knowledge from literature or databases. This validation ensures the reliability and biological relevance of transcriptomic markers identified for crop improvement.
- F. **Integration of Prior Knowledge:** Incorporating prior knowledge, such as gene regulatory networks or functional annotations from related species, can enhance the interpretation of transcriptomics data in crops. Existing knowledge can provide insights into the biological relevance and potential functions of differentially expressed genes and pathways. Leveraging existing resources, databases, and computational tools facilitates the integration of prior knowledge into the analysis pipeline.
- G. **Collaborative Efforts and Data Sharing:** Collaboration among researchers, data sharing, and standardization of protocols are essential for overcoming challenges in data analysis and interpretation. Sharing raw data, metadata, and analysis pipelines promotes transparency, reproducibility, and cross-validation of results. Collaborative efforts enable the pooling of data across multiple studies, enhancing statistical power and providing a broader context for data interpretation.

IX. Conclusion:

Transcriptomics presents immense potential for sustainable agriculture by providing insights into gene expression patterns, regulatory networks, and molecular mechanisms underlying important crop traits. By unraveling the complexities of gene regulation, transcriptomics enables the development of stress-tolerant crop varieties, optimization of resource utilization, and enhancement of crop productivity. The integration of transcriptomics with other -omics technologies further enhance our understanding of complex biological processes. However, addressing challenges in data analysis, interpretation, and integration is crucial. Robust bioinformatics pipelines, advanced statistical methods, and collaborative efforts are necessary for accurate data analysis. Data sharing, standardization, and open-access resources promote progress and translation of transcriptomics research. Overall, transcriptomics offers opportunities to achieve sustainable agriculture, ensuring food security and environmental sustainability for future generations.

Table 1: Case studies showcasing RNA-seq for identifying differentially expressed genes

Plant Species	Transcriptomics	Research Goal	Sampling	Key Findings	Source
Soybean	mRNA Sequencing	Unravel molecular mechanisms of drought tolerance	Seedlings of drought-tolerant genotype Jindou 21 (JD) and drought-sensitive genotype Tianlong No.1 (N1)	Drought stress induces cell wall remodelling and upregulates stress-related protein genes, with JD playing a crucial role in multiple signal transduction pathways and stress-related transcription factors, thereby enhancing drought tolerance.	(35)
Rice	Total RNA Seq	Analyze the transcriptome of rice flag leaves biofortified with Zn, Se, and Zn-Se	Rice flag leaves biofortified with Zn, Se, and Zn-Se	3170 differentially expressed genes (DEGs) were identified. DEGs were related to metabolic pathways for micronutrient mobilization. Zn played a role in nitrogenous compound, carboxylic acid, organic acid, and amino acid biosynthesis while Se was involved in vitamin biosynthesis and ion homeostasis	(27)
Soybean	Comparative transcriptomics	Identify differential gene expression and mechanisms of nutrient differences	A7 and A35 soybean seeds	In the A7 variety, the DEGs were associated with cell cycle regulation and stress response, while in A35, they were involved in fatty acid and sugar metabolism. Differences between A7 and A35 were found in fatty acid metabolism and the synthesis of seed storage proteins (SSPs). Specific enzymes and transcription factors were identified as potential contributors to variations in oil and protein content.	(65)
Gramineae Crops	Comparative transcriptomics	Study the evolutionary patterns and selective forces of GH3 genes	Six representative Gramineae crops	The study identified different numbers of GH3 proteins: 17 in <i>Setaria italica</i> , 13 in <i>Zea mays</i> , 11 in <i>Sorghum bicolor</i> , 9 in <i>Hordeum vulgare</i> , 8 in <i>Brachypodium distachyon</i> , and 11 in <i>Oryza sativa</i> ssp. <i>japonica</i> .	(66)
Cotton	Small RNA Sequencing	Elucidate the molecular basis of miRNAs in response to flowering time regulation in cotton	Fifth true leaf stage	GhSPL4, which is targeted by GhmiR156, was found to be upregulated in EFCs. Through co-expression regulatory network analysis (WGCNA), several node genes, including GhSOC1, GhAP1, GhFD, GhCOL3, and GhAGL16, were identified in the auxin- and gibberellin-mediated flowering time regulatory networks in cotton.	(55)
Pepper	mRNA Seq	Study genome-wide gene expression under heat stress in pepper plants	Time-course transcriptome analysis	Comparative analysis between heat-treated and control plants revealed differentially expressed genes (DEGs) with the largest changes observed at 24 h.	(24)

				More DEGs were identified in the root compared to the leaf. Analysis of 25 heat shock factor genes (HSFs) in pepper identified five HSFs that responded to heat stress. The study improves understanding of the genome-wide heat stress response in pepper plants.	
Tomato	RNA Seq	Investigate the implications of plant viruses in Belgian tomato farms	Comparison of growers' perception and visual inspection with high throughput sequencing technologies (HTS)	Differential expression analysis revealed more genes affected in the root compared to the leaf, with the largest changes observed at 24 hours. Five heat shock factor genes were identified in pepper, contributing to our understanding of heat tolerance mechanisms	(68)
Soybean	Comparative transcriptome	Unravel the molecular mechanisms under soybean drought tolerance	Jindou 21 (JD) and Tianlong No.1 (N1) seedlings	Drought-tolerant JD seedlings exhibited 6038 differentially expressed genes (DEGs) involved in signal transduction pathways, including hormone signaling, calcium signaling, and MAPK signaling, as well as cell wall remodeling and stress-related protein genes, indicating its higher drought tolerance compared to the drought-sensitive N1 seedlings in soybean.	(35)
Soybean	mRNA Transcriptomics	Elucidate the molecular mechanisms underlying soybean resistance to drought stress	Analysis of transcriptome and metabolome changes in soybean leaves at the seedling stage	Comprehensive analysis of data reveals metabolic regulation in response to drought stress. Identification of candidate genes (LOC100802571, LOC100814585, LOC100777350, LOC100787920, LOC100800547, LOC100785313) showing different expression trends between cultivars, potentially contributing to differences in drought resistance. Upregulation of Monohydroxy-trimethoxyflavone-O-(600-malonyl) glucoside and other flavonoids in response to drought stress. Identification of key candidate genes and metabolites involved in isoflavone biosynthesis and the TCA cycle, suggesting their important roles in soybean response to drought	(69)
Pennycress	Total RNA Seq	Identify targets for improving seed oil content in pennycress	Metabolomic and transcriptomic studies of developing embryos from 22 natural variants	Identified key processes crucial for enhancing oil content in pennycress include directing carbon partitioning towards the chloroplast, optimizing lipid metabolism, maximizing photosynthesis efficiency, and maintaining precise control over nitrogen availability.	(70)

Peach	Total RNA Seq	Investigate the molecular basis of low temperature conditioning (LTC) in alleviating peach fruit chilling injury in fruit maintained at 0 °C and LTC	Comparison of transcriptome, ethylene production, flesh softening, internal browning, and membrane lipids	Low temperature conditioning (LTC) in peaches results in increased ethylene production, accelerated flesh softening, reduced internal browning, enhanced fatty acid content and desaturation, and modulation of lipid composition and membrane stability.	(34)
Soybean	mRNA Seq	Identify potential differentially expressed genes (DEGs) in two wild soybean cultivars with different isoflavone compositions	RNA-seq technology	Identification of differentially metabolized genes and TF genes associated with isoflavone biosynthesis in different seed developmental stages. Six identified DEGs may be involved in the positive or negative regulation of isoflavone synthesis in soybean seeds.	(71)
Soybean	RNA Seq	Investigate of soybean root rot associated with continuous cropping by <i>Funneliformis mosseae</i>	Soil planted with soybean infected with <i>Fusarium oxysporum</i> and inoculated with <i>F. mosseae</i>	Transcriptomic analyses revealed the effect of <i>Funneliformis mosseae</i> on genes expression in <i>Fusarium oxysporum</i>	(72)
Soybean	RNA Seq	To investigate the molecular mechanisms of soybean resistance to anthracnose caused by <i>C. truncatum</i>	Pods from two soybean lines (ZC3 and ZC-2) infected with <i>C. truncatum</i>	Factors contributing to enhanced resistance of ZC-2 compared to ZC3 included signal transduction pathways, transcription factors (WRKY and bHLH), resistance genes (PTI1, RPP13, RGA2, RPS6, and ULP2B), pathogenesis-related genes (chitinase and lipid transfer protein), and terpenoid metabolism. Targeted metabolomic analysis revealed a more pronounced and rapid response of terpenoid metabolism to <i>C. truncatum</i> infection in ZC-2 compared to ZC3.	(73)
Soybean	RNA Seq	To investigate the role of lipid metabolism, specifically glycolysis, fatty acid biosynthesis, and membrane lipid biosynthesis, in soybean nodulation and rhizobia-soybean symbiosis.	Developing nodules and roots of soybean plants undergoing rhizobia-legume symbiosis.	Gene expression analysis revealed enhanced expression of metabolic genes involved in FA, membrane lipid, and 2-MAG biosynthesis and transport, mediated by the RAML-WRI-FatM-GPAT-STRL pathway. GmWRI1b overexpression increased nodulation by promoting glycolysis, FA biosynthesis, and membrane lipid biosynthesis. GmLEC2a overexpression led to reduced nodulation, associated with repression of glycolysis and FA and membrane lipid biosynthesis	(74)

Soybean	RNA Seq	To investigate the effects of <i>F. mosseae</i> treatment on soybean root rot caused by <i>F. oxysporum</i> and understand the molecular response of soybean roots to this interaction.	Roots of continuously cropped soybeans infected with <i>F. mosseae</i> and <i>F. oxysporum</i>	Transcriptome analysis identified 24,285 differentially expressed genes (DEGs), with upregulation of genes involved in the soybean defense response, including those encoding phenylalanine ammonia lyase (PAL), trans-cinnamate monooxygenase (CYP73A), cinnamyl-CoA reductase (CCR), chalcone isomerase (CHI), and coffee-coenzyme o-methyltransferase. The data revealed substantial changes in the abundance of intermediate metabolites and enzymes, particularly in the isoflavonoid biosynthesis pathway, suggesting its important role in the soybean root response to pathogen infection.	(75)
Soybean	RNA-sequencing	To investigate the transcriptomic responses of soybean roots to sucrose signalling as a response to nutrient deficiencies, specifically phosphorus (Pi) and iron (Fe) deficiencies.	Soybean Root	Gene ontology (GO) analysis revealed that sucrose-induced genes were mainly involved in signal transduction, particularly hormone, reactive oxygen species (ROS), and calcium signaling, as well as transcriptional regulation. Sucrose signaling also triggered crosstalk between biotic and abiotic stress responses, as indicated by GO enrichment analysis.	(76)
Brassica napus	Comparative transcriptomic	To identify genes involved in horizontal resistance to Sclerotinia stem rot in Brassica napus and investigate key pathways activated in response to the disease	Resistant and susceptible lines of Brassica napus	The resistant line exhibited activation of the biosynthesis of amino acids and secondary metabolites in response to the disease. Phenylpropanoid biosynthesis and arginine biosynthesis were specifically induced in the resistant line at different time points post inoculation. Glucosinolate biosynthesis, flavonoid biosynthesis, and the alanine, aspartate, and glutamate metabolism pathways were induced in the resistant line. Genes associated with ethylene, salicylic acid, and jasmonic acid were highly induced in the resistant line compared to the susceptible line, indicating their role in the defense response.	(77)
Soybean	RNA Seq	To understand the dynamic changes in metabolites and gene expression related to phosphorus absorption and utilization in soybean seedling roots under short-	Soybean seedling roots subjected to short- and long-term phosphorus deficiency stress	The expression of GmPHT1 and GmSPX triggered the phosphorus starvation signal pathway, leading to the upregulation of GmPS and GmPAP genes involved in the synthesis and secretion of organophosphorus hydrolase and organic acid in soybean roots. Phospholipid metabolism was significantly enhanced	(78)

		and long-term phosphorus deficiency stress.		after 15 days of phosphorus deficiency stress, accompanied by the upregulation of GmSQD, a crucial enzyme in lipid biosynthesis.	
Soybean	RNA-Seq	To understand the molecular mechanisms governing resistance to <i>Corynespora cassicola</i> infection in soybean and identify differentially expressed genes (DEGs) involved in the regulatory network between soybean and <i>C. cassicola</i> .	Two known resistant genotypes and two susceptible genotypes of soybean under infected and control conditions	The identified DEGs were associated with secondary metabolites, immune response, defense response, phenylpropanoid, and flavonoid/isoflavonoid pathways in all genotypes. The two resistant genotypes showed additional upregulated DEGs related to flavonoids, jasmonic acid, salicylic acid, and brassinosteroids, indicating an enhanced defense network. Differentially expressed transcription factors, immune receptors, and defense genes with leucine-rich repeat domain, dirigent proteins, and cysteine (C)-rich receptor-like kinases were identified.	(79)
Soybean	Single nematode transcriptomic profiling combined with long-read sequencing	To understand the mechanisms responsible for resistance breakdown in soybean cyst nematode (SCN) on soybean cultivars derived from the same source of resistance (PI 88788).	Soybean plants infected with SCN, specifically focusing on the late infection stage	Among the identified effector candidates, the novel gene Hg-CPZ-1 and a pioneer effector transcript generated through alternative splicing of the non-effector gene Hetgly21698 were found. Limited evidence was found to directly link alternative splicing in effectors to the breakdown of resistance. However, the analysis revealed a distinct pattern of effector upregulation in response to PI 88788 resistance, suggesting a possible adaptation process by SCN to host resistance.	(80)
Soybean	Comparative transcriptomics	To investigate the molecular mechanism underlying cytoplasmic male sterility (CMS) in soybean, a critical component of heterosis-based breeding.	Experiment conducted on pollen from the CMS line W931A and its maintainer line, W931B,	The findings suggest that the development of pollen in W931A is likely regulated through the suppression of identified DEGs and DEPs. This study enhances our understanding of the molecular mechanism underlying CMS in soybean, providing valuable insights for soybean fertility research and the utilization of heterosis for soybean improvement.	(81)
Soybean	mRNA Seq	To investigate the gene expression networks and molecular regulatory mechanisms controlling soybean seed oil and protein contents during seed development.	Analyses were performed on soybean seeds during development in two soybean varieties with contrasting protein and oil contents.	A total of 41,036 genes and 392 metabolites were identified. DEGs were enriched in pathways related to phenylpropanoid biosynthesis, glycerolipid metabolism, carbon metabolism, plant hormone signal transduction, linoleic acid metabolism, and amino acid and secondary metabolite biosynthesis. K-means analysis divided the DEGs into 12 distinct clusters.	(82)

				Candidate gene sets regulating protein and oil biosynthesis in soybean seeds were identified.	
Soybean	mRNA Seq	To identify candidate genes and related pathways involved in the regulation of unsaturated fatty acid (UFA) contents during soybean seed development.	RNA-seq analysis was performed on soybean seeds at three developmental stages in two soybean lines with different UFA profiles	A series of genes and pathways related to fatty acid metabolism were identified through Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. The crucial period in the formation of UFA profiles was determined to be 40 days after flowering (DAF). Weighted gene co-expression network analysis revealed three modules containing six genes highly associated with the contents of oleic acid and linoleic acid.	(83)
Soybean	Total RNA Seq	To investigate the molecular mechanisms of soybean resistance to anthracnose (caused by <i>Colletotrichum truncatum</i>) and identify factors contributing to enhanced resistance.	Analysis of pods from two soybean lines,	ZC-2 exhibited enhanced resistance to anthracnose compared to ZC3. Factors contributing to enhanced resistance included signal transduction pathways (jasmonic acid, auxin, mitogen-activated protein kinase, and Ca ²⁺ signaling), transcription factors (WRKY and bHLH), resistance genes (PTI1, RPP13, RGA2, RPS6, and ULP2B), pathogenesis-related genes (chitinase and lipid transfer protein), and terpenoid metabolism. Targeted metabolomic analysis showed that terpenoid metabolism responded more promptly and intensely to <i>C. truncatum</i> infection in ZC-2 than in ZC3.	(73)
Soybean	Total RNA Seq	Investigate the defense mechanism of soybean response to <i>Peronospora manshurica</i> (Pm) infection	Analysis of WRKY transcription factors (TFs) in SDM-HR and SDM-HS genotypes	Differential expression of 22 WRKY TFs in SDM-high resistant (HR) and SDM-high susceptible (HS) genotype identification of 16 WRKY TFs specifically responding to fungal induction. Characterization of GmWRKY31 as a TF that binds to the cis-acting W-box element in the promoter region of the GmSAGT1 gene Higher transcriptional expression of GmSAGT1 associated with enhanced SDM- resistance speculation that GmWRKY31 regulates GmSAGT1 gene expression and is involved in SA-mediated immune responses in soybean.	(84)

REFERENCES

- [1] Qaim M. Role of New Plant Breeding Technologies for Food Security and Sustainable Agricultural Development. *Appl Econ Perspect Policy*. 2020 Jun 1;42(2):129–50.
- [2] Tyagi P, Singh D, Mathur S, Singh A, Ranjan R. Upcoming progress of transcriptomics studies on plants: An overview. *Front Plant Sci*. 2022;13(December):1–15.
- [3] Segelbacher G, Bosse M, Burger P, Galbusera P, Godoy JA, Helsen P, et al. New developments in the field of genomic technologies and their relevance to conservation management. Vol. 23, *Conservation Genetics*. Springer Science and Business Media B.V.; 2022. p. 217–42.
- [4] Pazhamala LT, Kudapa H, Weckwerth W, Millar AH, Varshney RK. *Systems biology for crop improvement*. Vol. 14, *Plant Genome*. John Wiley and Sons Inc; 2021.
- [5] Razzaq A, Wishart DS, Wani SH, Hameed MK, Mubin M, Saleem F. *Advances in Metabolomics-Driven Diagnostic Breeding and Crop Improvement*. *Metabolites*. 2022 Jun 1;12(6).
- [6] Werkissa Y. Application of Genetically Modified Organism (GMO) crop technology and its implications in modern agriculture. *International Journal of Agricultural Science and Food Technology*. 2022 Jan 8;8(1):014–20.
- [7] Das T, Anand U, Pal T, Mandal S, Kumar M, Radha, et al. Exploring the potential of CRISPR/Cas genome editing for vegetable crop improvement: An overview of challenges and approaches. Vol. 120, *Biotechnology and Bioengineering*. John Wiley and Sons Inc; 2023. p. 1215–28.
- [8] Khan S, Anwar S, Yu S, Sun M, Yang Z, Gao ZQ. Development of drought-tolerant transgenic wheat: Achievements and limitations. *Int J Mol Sci*. 2019 Jul 1;20(13).
- [9] Varshney RK, Bohra A, Yu J, Graner A, Zhang Q, Sorrells ME. *Designing Future Crops: Genomics-Assisted Breeding Comes of Age*. Vol. 26, *Trends in Plant Science*. Elsevier Ltd; 2021. p. 631–49.
- [10] Ashraf MF, Hou D, Hussain Q, Imran M, Pei J, Ali M, et al. Entailing the Next-Generation Sequencing and Metabolome for Sustainable Agriculture by Improving Plant Tolerance. Vol. 23, *International Journal of Molecular Sciences*. MDPI; 2022.
- [11] Yang Y, Saand MA, Huang L, Abdelaal WB, Zhang J, Wu Y, et al. Applications of Multi-Omics Technologies for Crop Improvement. Vol. 12, *Frontiers in Plant Science*. *Frontiers Media S.A.*; 2021.
- [12] Temple C, Blouin AG, Tindale S, Steyer S, Marechal K, Massart S. High throughput sequencing technologies complemented by growers' perceptions highlight the impact of tomato virome in diversified vegetable farms and a lack of awareness of emerging virus threats. *Front Sustain Food Syst*. 2023 Jun 16;7.
- [13] Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T. Transcriptomics technologies. *PLoS Comput Biol*. 2017 May 1;13(5).
- [14] Mikołajczak K, Kuczyńska A, Ogródowicz P, Kielbowicz-Matuk A, Ćwiek-Kupczyńska H, Daszkowska-Golec A, et al. High-throughput sequencing data revealed genotype-specific changes evoked by heat stress in crown tissue of barley *sdw1* near-isogenic lines. *BMC Genomics*. 2022 Dec 1;23(1).
- [15] Villamor DEV, Keller KE, Martin RR, Tzanetakis IE. Comparison of High Throughput Sequencing to Standard Protocols for Virus Detection in Berry Crops. *Plant Dis*. 2022 Feb 1;106(2):518–25.
- [16] Yang Y, Saand MA, Huang L, Abdelaal WB, Zhang J, Wu Y, et al. Applications of Multi-Omics Technologies for Crop Improvement. Vol. 12, *Frontiers in Plant Science*. *Frontiers Media S.A.*; 2021.
- [17] Singh BK, Venkadesan S, Ramkumar MK, Shanmugavadivel PS, Dutta B, Prakash C, et al. Meta-Analysis of Microarray Data and Their Utility in Dissecting the Mapped QTLs for Heat Acclimation in Rice. *Plants*. 2023 Apr 1;12(8).
- [18] Liu T, Chen T, Kan J, Yao Y, Guo D, Yang Y, et al. The GhMYB36 transcription factor confers resistance to biotic and abiotic stress by enhancing PR1 gene expression in plants. *Plant Biotechnol J*. 2022 Apr 1;20(4):722–35.
- [19] Freitas FCP, Depintor TS, Agostini LT, Luna-Lucena D, Nunes FMF, Bitondi MMG, et al. Evaluation of reference genes for gene expression analysis by real-time quantitative PCR (qPCR) in three stingless bee species (Hymenoptera: Apidae: Meliponini). *Sci Rep*. 2019 Dec 1;9(1).
- [20] Qi W, Chen J, Han Y, Li Z, Su X, Yeo FKS. Editorial: Omics-driven crop improvement for stress tolerance. Vol. 14, *Frontiers in Plant Science*. *Frontiers Media S.A.*; 2023.
- [21] González-Morales S, Solís-Gaona S, Valdés-Caballero MV, Juárez-Maldonado A, Loredó-Treviño A, Benavides-Mendoza A. Transcriptomics of Biostimulation of Plants Under Abiotic Stress. Vol. 12, *Frontiers in Genetics*. *Frontiers Media S.A.*; 2021.
- [22] Baldoni E. Improving drought tolerance: Can comparative transcriptomics support strategic rice breeding? *Plant Stress*. 2022 Jan 1;3.
- [23] Mikołajczak K, Kuczyńska A, Ogródowicz P, Kielbowicz-Matuk A, Ćwiek-Kupczyńska H, Daszkowska-Golec A, et al. High-throughput sequencing data revealed genotype-specific changes evoked by heat stress in crown tissue of barley *sdw1* near-isogenic lines. *BMC Genomics*. 2022 Dec 1;23(1).
- [24] Tang B, Li X, Zhang X, Yin Q, Xie LL, Zou X, et al. Transcriptome data reveal gene clusters and key genes in pepper response to heat shock. *Front Plant Sci*. 2022 Sep 21;13.
- [25] Bacher H, Zhu F, Gao T, Liu K, Dhaff BK, Awada T, et al. Wild emmer introgression alters root-to-shoot growth dynamics in durum wheat in response to water stress. *Plant Physiol*. 2021 Nov 1;187(3):1149–62.
- [26] Afzal M, Hindawi SES, Alghamdi SS, Migdadi HH, Khan MA, Hasnain MU, et al. Potential Breeding Strategies for Improving Salt Tolerance in Crop Plants. Vol. 42, *Journal of Plant Growth Regulation*. Springer; 2023. p. 3365–87.
- [27] Roda FA, Marques I, Batista-Santos P, Esquivel MG, Ndayiragije A, Lidon FC, et al. Rice Biofortification With Zinc and Selenium: A Transcriptomic Approach to Understand Mineral Accumulation in Flag Leaves. *Front Genet*. 2020 Jul 7;11.
- [28] Cruet-Burgos C, Rhodes DH. Unraveling transcriptomics of sorghum grain carotenoids: a step forward for biofortification. *BMC Genomics*. 2023 Dec 1;24(1).
- [29] Dueñas CJ, Slamet-Loedin I, Macovei A. Transcriptomics view over the germination landscape in biofortified rice. Vol. 12, *Genes*. MDPI; 2021.
- [30] Singh D, Prasanna R. Potential of microbes in the biofortification of Zn and Fe in dietary food grains. A review. Vol. 40, *Agronomy for Sustainable Development*. Springer; 2020.
- [31] Pantha P, Chalivendra S, Oh DH, Elder BD, Dassanayake M. A tale of two transcriptomic responses in agricultural pests via host defenses and viral replication. *Int J Mol Sci*. 2021 Apr 1;22(7).
- [32] Schumpp O. High-throughput sequencing for the diagnostic of plant pathologies and identification of pests: recommendations and challenges. *Peer Community In Infections*. 2022 Oct 7;
- [33] Gupta A, Sharma T, Singh SP, Bhardwaj A, Srivastava D, Kumar R. Prospects of microgreens as budding living functional food: Breeding and biofortification through OMICS and other approaches for nutritional security. Vol. 14, *Frontiers in Genetics*. *Frontiers Media S.A.*; 2023.
- [34] Wang K, Yin XR, Zhang B, Grierson D, Xu CJ, Chen KS. Transcriptomic and metabolic analyses provide new insights into chilling injury in peach fruit. *Plant Cell Environ*. 2017 Aug 1;40(8):1531–51.
- [35] Xuan H, Huang Y, Zhou L, Deng S, Wang C, Xu J, et al. Key Soybean Seedlings Drought-Responsive Genes and Pathways Revealed by Comparative Transcriptome Analyses of Two Cultivars. *Int J Mol Sci*. 2022 Mar 1;23(5).
- [36] Fu A, Wang Q, Mu J, Ma L, Wen C, Zhao X, et al. Combined genomic, transcriptomic, and metabolomic analyses provide insights into chayote (*Sechium edule*) evolution and fruit development. *Hortic Res*. 2021 Dec 1;8(1).
- [37] García-Gómez BE, Salazar JA, Nicolás-Almansa M, Razi M, Rubio M, Ruiz D, et al. Molecular bases of fruit quality in prunus species: An integrated genomic, transcriptomic, and metabolic review with a breeding perspective. Vol. 22, *International Journal of Molecular Sciences*. MDPI AG; 2021. p. 1–38.
- [38] He Y, Ye G, Chi S, Tan B, Dong X, Yang Q, et al. Integrative Transcriptomic and Small RNA Sequencing Reveals Immune-Related miRNA–mRNA Regulation Network for Soybean Meal-Induced Enteritis in Hybrid Grouper, *Epinephelus fuscoguttatus* ♀ × *Epinephelus lanceolatus* ♂. *Front Immunol*. 2020 Aug 6;11.
- [39] El-Sawah AM, El-Keblawy A, Ali DFI, Ibrahim HM, El-Sheikh MA, Sharma A, et al. Arbuscular mycorrhizal fungi and plant growth-promoting rhizobacteria enhance soil key enzymes, plant growth, seed yield, and qualitative attributes of guar. *Agriculture (Switzerland)*. 2021 Mar 1;11(3):1–19.
- [40] Kumawat KC, Sharma B, Nagpal S, Kumar A, Tiwari S, Nair RM. Plant growth-promoting rhizobacteria: Salt stress alleviators to improve crop productivity for sustainable agriculture development. Vol. 13, *Frontiers in Plant Science*. *Frontiers Media S.A.*; 2023.

- [41] Eposito A, Colantuono C, Ruggieri V, Chiusano ML. Bioinformatics for agriculture in the next-generation sequencing era. Vol. 3, Chemical and Biological Technologies in Agriculture. Springer International Publishing; 2016.
- [42] Sharmin RA, Bhuiyan MR, Lv W, Yu Z, Chang F, Kong J, et al. RNA-Seq based transcriptomic analysis revealed genes associated with seed-flooding tolerance in wild soybean (*Glycine soja* Sieb. & Zucc.). *Environ Exp Bot*. 2020 Mar 1;171.
- [43] Liu A, Xiao Z, Li MW, Wong FL, Yung WS, Ku YS, et al. Transcriptomic reprogramming in soybean seedlings under salt stress. *Plant Cell Environ*. 2019 Jan 1;42(1):98–114.
- [44] Zhang L, Shang J, Wang W, Du J, Li K, Wu X, et al. Comparison of Transcriptome Differences in Soybean Response to Soybean Mosaic Virus under Normal Light and in the Shade. *Viruses*. 2019 Sep 1;11(9).
- [45] Benny J, Marchese A, Giovino A, Marra FP, Perrone A, Caruso T, et al. Gaining insight into exclusive and common transcriptomic features linked to drought and salinity responses across fruit tree crops. *Plants*. 2020 Sep 1;9(9):1–20.
- [46] Riccini A, Picarella ME, De Angelis F, Mazzucato A. Bulk RNA-Seq analysis to dissect the regulation of stigma position in tomato. *Plant Mol Biol*. 2021 Feb 1;105(3):263–85.
- [47] Li X, Wang CY. From bulk, single-cell to spatial RNA sequencing. Vol. 13, International Journal of Oral Science. Springer Nature; 2021.
- [48] Karim MM, Dakouri A, Zhang Y, Chen Q, Peng G, Strelkov SE, et al. Two clubroot-resistance genes, RCR3 and RCR9wa, mapped in brassica RAPA using bulk segregant rna sequencing. *Int J Mol Sci*. 2020 Jul 2;21(14):1–15.
- [49] Sunaga-Franze DY, Muino JM, Braeuning C, Zong M, Smaczniak C, Yan W, et al. Single-nuclei RNA-sequencing of plants. Available from: <https://doi.org/10.1101/2020.11.14.382812>
- [50] Sunaga-Franze DY, Muino JM, Braeuning C, Xu X, Zong M, Smaczniak C, et al. Single-nucleus RNA sequencing of plant tissues using a nanowell-based system. *Plant Journal*. 2021 Nov 1;108(3):859–69.
- [51] Bawa G, Liu Z, Yu X, Qin A, Sun X. Single-Cell RNA Sequencing for Plant Research: Insights and Possible Benefits. Vol. 23, International Journal of Molecular Sciences. MDPI; 2022.
- [52] Xanthopoulou A, Montero-Pau J, Picó B, Boumpas P, Tsaliki E, Paris HS, et al. A comprehensive RNA-Seq-based gene expression atlas of the summer squash (*Cucurbita pepo*) provides insights into fruit morphology and ripening mechanisms. *BMC Genomics*. 2021 Dec 1;22(1).
- [53] Tang C, Xu Q, Zhao J, Yue M, Wang J, Wang X, et al. A rust fungus effector directly binds plant pre-mRNA splice site to reprogram alternative splicing and suppress host immunity. *Plant Biotechnol J*. 2022 Jun 1;20(6):1167–81.
- [54] Yan W, Cao S, Wu Y, Ye Z, Zhang C, Yao G, et al. Integrated Analysis of Physiological, mRNA Sequencing, and miRNA Sequencing Data Reveals a Specific Mechanism for the Response to Continuous Cropping Obstacles in *Pogostemon cablin* Roots. *Front Plant Sci*. 2022 Apr 1;13.
- [55] Zhou Y, Myat AA, Liang C, Meng Z, Guo S, Wei Y, et al. Insights Into MicroRNA-Mediated Regulation of Flowering Time in Cotton Through Small RNA Sequencing. *Front Plant Sci*. 2022 Mar 31;13.
- [56] Raabe CA, Tang TH, Brosius J, Rozhdestvensky TS. Biases in small RNA deep sequencing data. *Nucleic Acids Res*. 2014 Feb;42(3):1414–26.
- [57] Benesova S, Kubista M, Valihrah L. Small rna-sequencing: Approaches and considerations for mirna analysis. Vol. 11, Diagnostics. MDPI; 2021.
- [58] Martinez-Seidel F, Beine-Golovchuk O, Hsieh YC, Kopka J. Systematic review of plant ribosome heterogeneity and specialization. Vol. 11, Frontiers in Plant Science. Frontiers Media S.A.; 2020.
- [59] Wu HYL, Song G, Walley JW, Hsu PY. The tomato translational landscape revealed by transcriptome assembly and ribosome profiling. *Plant Physiol*. 2019 Sep 1;181(1):367–80.
- [60] Ma Q, Wang Y, Li S, Wen J, Zhu L, Yan K, et al. Ribosome footprint profiling enables elucidating the systemic regulation of fatty acid accumulation in *Acer truncatum*. *BMC Biol*. 2023 Dec 1;21(1).
- [61] Walker DC, Lozier ZR, Bi R, Kanodia P, Miller WA, Liu P. Variational inference for detecting differential translation in ribosome profiling studies. *Front Genet*. 2023 Jun 23;14.
- [62] Fujita T, Kurihara Y, Iwasaki S. The Plant Translatome Surveyed by Ribosome Profiling. Vol. 60, Plant and Cell Physiology. Oxford University Press; 2019. p. 1917–26.
- [63] Wang Y, Zhang H, Lu J. Recent advances in ribosome profiling for deciphering translational regulation. Vol. 176, Methods. Academic Press Inc.; 2020. p. 46–54.
- [64] Jian H, Wen S, Liu R, Zhang W, Li Z, Chen W, et al. Dynamic Translational Landscape Revealed by Genome-Wide Ribosome Profiling under Drought and Heat Stress in Potato. *Plants*. 2023 Jun 6;12(12):2232.
- [65] Peng L, Qian L, Wang M, Liu W, Song X, Cheng H, et al. Comparative transcriptome analysis during seeds development between two soybean cultivars. *PeerJ*. 2021 Mar 1;9.
- [66] Kong W, Zhang Y, Deng X, Li S, Zhang C, Li Y. Comparative Genomic and Transcriptomic Analysis Suggests the Evolutionary Dynamic of GH3 Genes in Gramineae Crops. *Front Plant Sci*. 2019 Oct 15;10.
- [67] Alves de Freitas Guedes F, Menezes-Silva PE, DaMatta FM, Alves-Ferreira M. Using transcriptomics to assess plant stress memory. *Theor Exp Plant Physiol*. 2019 Mar 15;31(1):47–58.
- [68] Choi J, Osatuke AC, Erich G, Stevens K, Hwang MS, Al Rwahnih M, et al. High-Throughput Sequencing Reveals Tobacco and Tomato Ringspot Viruses in Pawpaw. *Plants*. 2022 Dec 1;11(24).
- [69] Wang X, Song S, Wang X, Liu J, Dong S. Transcriptomic and Metabolomic Analysis of Seedling-Stage Soybean Responses to PEG-Simulated Drought Stress. *Int J Mol Sci*. 2022 Jun 1;23(12).
- [70] Arias CL, García Navarrete LT, Mukundi E, Swanson T, Yang F, Hernandez J, et al. Metabolic and transcriptomic study of pennycress natural variation identifies targets for oil improvement. *Plant Biotechnol J*. 2023;
- [71] Zhang B, Zhao K, Ren H, Lamlo SF, Liu X, Wang X, et al. Comparative Study of Isoflavone Synthesis Genes in Two Wild Soybean Varieties Using Transcriptomic Analysis. *Agriculture*. 2023 May 30;13(6):1164.
- [72] Zhang X, Bai L, Guo N, Cai B. Transcriptomic analyses revealed the effect of *Funneliformis mosseae* on genes expression in *Fusarium oxysporum*. *PLoS One*. 2020 Jul 1;15(7 July).
- [73] Zhu L, Yang Q, Yu X, Fu X, Jin H, Yuan F. Transcriptomic and Metabolomic Analyses Reveal a Potential Mechanism to Improve Soybean Resistance to Anthracnose. *Front Plant Sci*. 2022 Apr 27;13.
- [74] Zhang G, Ahmad MZ, Chen B, Manan S, Zhang Y, Jin H, et al. Lipidomic and transcriptomic profiling of developing nodules reveals the essential roles of active glycolysis and fatty acid and membrane lipid biosynthesis in soybean nodulation. *Plant Journal*. 2020 Aug 1;103(4):1351–71.
- [75] Lu CC, Guo N, Yang C, Sun HB, Cai BY. Transcriptome and metabolite profiling reveals the effects of *Funneliformis mosseae* on the roots of continuously cropped soybeans. *BMC Plant Biol*. 2020 Dec 1;20(1).
- [76] Nidumolu LCM, Lorilla KM, Chakravarty I, Uhde-Stone C. Soybean Root Transcriptomics: Insights into Sucrose Signaling at the Crossroads of Nutrient Deficiency and Biotic Stress Responses. *Plants*. 2023 Jun 1;12(11).
- [77] Qasim MU, Zhao Q, Shahid M, Cai G, Huang H, Li G, et al. Overlapping pathways involved in resistance against *Sclerotinia stem rot* in *Brassica napus* revealed through transcriptomic and metabolomic profiling. *Plant Growth Regul*. 2023 Apr 5;
- [78] Li M, Zhou J, Liu Q, Mao L, Li H, Li S, et al. Dynamic variation of nutrient absorption, metabolomic and transcriptomic indexes of soybean (*Glycine max*) seedlings under phosphorus deficiency. *AoB Plants*. 2023 Feb 1;15(2).
- [79] Patel S, Patel J, Silliman K, Hall N, Bowen K, Koebnick J. Comparative Transcriptome Profiling Unfolds a Complex Defense and Secondary Metabolite Networks Imparting *Corynespora cassiicola* Resistance in Soybean (*Glycine max* (L.) Merrill). *Int J Mol Sci*. 2023 Jun 23;24(13):10563.
- [80] Ste-Croix DT, Bélanger RR, Mimee B. Single Nematode Transcriptomic Analysis, Using Long-Read Technology, Reveals Two Novel Virulence Gene Candidates in the Soybean Cyst Nematode, *Heterodera glycines*. *Int J Mol Sci*. 2023 Jun 1;24(11).
- [81] Wang D, Wang Y, Zhang L, Yang Y, Wu Q, Hu G, et al. Integrated transcriptomic and proteomic analysis of a cytoplasmic male sterility line and associated maintainer line in soybean. *Front Plant Sci*. 2023 Feb 2;14.
- [82] Xu W, Wang Q, Zhang W, Zhang H, Liu X, Song Q, et al. Using transcriptomic and metabolomic data to investigate the molecular mechanisms that determine protein and oil contents during seed development in soybean. *Front Plant Sci*. 2022 Sep 29;13.
- [83] Liu J, Dong L, Duan R, Hu L, Zhao Y, Zhang L, et al. Transcriptomic Analysis Reveals the Regulatory Networks and Hub Genes Controlling the Unsaturated Fatty Acid Contents of Developing Seed in Soybean. *Front Plant Sci*. 2022 May 12;13.
- [84] Dong H, Tan J, Li M, Yu Y, Jia S, Zhang C, et al. Transcriptome analysis of soybean WRKY TFs in response to *Peronospora manshurica* infection. *Genomics*. 2019 Dec 1;111(6):1412–22.

