Gene 927 (2024) 148681

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Review

A comprehensive review of the applications of RNA sequencing in celiac disease research

Maryam Shoaran [a], Hani Sabaie [b], Mehrnaz Mostafavi [c], Maryam Rezazadeh [b,*]

[a] Pediatric Health Research Center, Tabriz University of Medical Sciences, Tabriz, Iran
[b] Department of Medical Genetics, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran
[c] Faculty of Allied Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ARTICLE INFO

Edited by Nicola Lopizzo

Keywords:
Autoimmune disease
Celiac disease
Next-generation sequencing
RNA sequencing
Transcriptomics

ABSTRACT

RNA sequencing (RNA-seq) has undergone substantial advancements in recent decades and has emerged as a vital technique for profiling the transcriptome. The transition from bulk sequencing to single-cell and spatial approaches has facilitated the achievement of higher precision at cell resolution. It provides valuable biological knowledge about individual immune cells and aids in the discovery of the molecular mechanisms that contribute to the development of autoimmune diseases. Celiac disease (CeD) is an autoimmune disorder characterized by a strong immune response to gluten consumption. RNA-seq has led to significantly advanced research in multiple fields, particularly in CeD research. It has been instrumental in studies involving comparative transcriptomics, nutritional genomics and wheat research, cancer research in the context of CeD, genetic and noncoding RNA-mediated epigenetic insights, disease monitoring and biomarker discovery, regulation of mitochondrial functions, therapeutic target identification and drug mechanism of action, dietary factors, immune cell profiling and the immune landscape. This review offers a comprehensive examination of recent RNA-seq technology research in the field of CeD, highlighting future challenges and opportunities for its application.

1. Introduction

Celiac disease (CeD) is a chronic inflammatory condition in the small intestine caused by the consumption of gluten proteins found in wheat, barley, and rye (Shewry, 2019). It causes symptoms such as abdominal pain, diarrhea, and malnutrition and can also spread outside the intestines (Ferguson et al., 1993; Green, 2005). The diagnosis is based on clinical, serological, and histopathological information from biopsies of the proximal small intestine (Al-Toma et al., 2019). Patients often have diminished or leveled intestinal villi, excessive growth of crypts, and elevated levels of lymphocytes in the lamina propria and epithelial layer (Sakula and Shiner, 1957; Ferguson and Murray, 1971; Marsh, 1988). A strict, lifelong gluten-free diet is generally effective in the management of CeD. CeD is an autoimmune disorder triggered by a strong immune response to gluten, characterized by the presence of disease-specific autoantibodies, the destruction of intestinal epithelial cells, and typical autoimmune genetics involving the dominant influence of human leukocyte antigen (HLA) genes (Sollid and Jabri, 2013; Sollid and Jabri, 2005). In CeD, CD4 T cells, which are responsive to gluten, interact with B cells and CD8 T cells (Jabri and Sollid, 2017). The participation of B cells, which relies on support from CD4 T cells, results in the generation of antibodies to Transglutaminase 2 (TG2). It is noteworthy that the involvement of B cells will create an amplification mechanism for the T cells. This amplification process will favor peptides

Abbreviations: BSA, bulked segregant analysis; BSR-seq, bulked segregant RNA-Seq; cDNA, complementary DNA; CeD, celiac disease; CVID, common variable immunodeficiency; DEGs, differentially expressed genes; DGEA, differential gene expression analysis; dRNA-seq, direct RNA-seq; EED, environmental enteric dysfunction; EoD, eosinophilic duodenitis; eQTL, expression quantitative trait loci; EVs, extracellular vesicles; GFD, gluten-free diet; GWAS, genome-wide association study; HLA, human leukocyte antigen; IE-CTL, intraepithelial cytotoxic T lymphocyte; IFNγ, interferon-gamma; LA, larazotide acetate; lncRNAs, long noncoding RNAs; miRNAs, microRNAs; ONT, Oxford Nanopore; PacBio, Pacific Biosciences; PBMCs, peripheral blood mononuclear cells; PCs, plasma cells; RNA-seq, RNA sequencing; RCDII, type II refractory celiac disease; SAM, severe acute malnutrition; scRNA-seq, single-cell RNA-seq; SNP, single nucleotide polymorphism; TG2, TRANSGLUTAMINASE 2; TIMP-GLIA, tolerance in poly(lactide-co-glycolide) nanoparticles encapsulating gliadin protein; TJs, tight junctions; UMI, unique molecular identifiers.

* Corresponding author.
E-mail address: rezazadehm@tbzmed.ac.ir (M. Rezazadeh).

https://doi.org/10.1016/j.gene.2024.148681
Received 2 February 2024; Received in revised form 6 June 2024; Accepted 10 June 2024
Available online 11 June 2024

that are effective substrates for TG2 as T-cell epitopes. Altogether, these findings indicate that the immune response against gluten by T cells needs to reach a specific threshold to cause disease. CD4+ T cells play a crucial role in the transformation of intraepithelial cytotoxic T lymphocytes (IE-CTLs) into effector T cells. However, the epithelium releases signals, such as IL-15 and nonclassical MHC class I molecules, that contribute to tissue destruction. These signals control tissue-resident CD8 T cells and tissue damage, enabling CTLs to specifically eliminate damaged tissue cells that have reduced their MHC class I molecules. The factors required to activate this program in CeD appear to be sustained by gluten exposure (Jabri and Sollid, 2017).

Since its inception in 2005 (Margulies et al., 2005), high-throughput sequencing has enabled us to gain a comprehensive understanding of molecular-level life processes and conduct in-depth investigations to unravel the genome and transcriptome. RNA sequencing (RNA-seq), particularly single-cell RNA-seq (scRNA-seq), is a crucial component of high-throughput sequencing. It offers valuable biological insights into individual immune cells and helps uncover the molecular mechanisms underlying the development of autoimmune diseases (Nagafuchi et al., 2022). Hence, RNA-seq could provide indispensable knowledge for the study and management of CeD.

2. Progress in RNA-seq technologies

RNA-seq is a widely used tool in biological research for differential gene expression analysis (DGEA) (Emrich et al., 2007; Lister et al., 2008). This method involves extracting RNA, enriching or depleting mRNA, synthesizing complementary DNA (cDNA), and preparing a sequencing library with adaptor ligation. The library is then sequenced using a high-throughput platform such as Illumina, resulting in 10–30 million reads per sample. The computational steps include aligning and/or assembling reads; quantifying, filtering, and normalizing the data; and performing statistical modeling to identify significant changes (Stark et al., 2019). Initial RNA-seq studies produced DGE data using samples of bulk tissue (Stark et al., 2019; Cloonan et al., 2008). RNA-seq has expanded the scope of mRNA splicing and gene expression control research by revealing noncoding RNAs and enhancer RNAs. (Wang et al., 2008; Djebali et al., 2012; Morris and Mattick, 2014). It offers a more comprehensive understanding of RNA biology and the transcriptome compared to microarray-based techniques (Stark et al., 2019). The majority of the progress in developing RNA-seq has focused on Illumina short-read sequencing instruments. However, recent advancements in long-read RNA-seq and direct RNA-seq (dRNA-seq) techniques have enabled users to address inquiries that cannot be resolved using Illumina short-read technologies (Byrne et al., 2017; Garalde et al., 2018; Smith et al., 2019). Table 1 provides an overview of the four main platforms used for RNA-seq. Illumina technology is currently the most widely used platform for short-read RNA-seq. It produces extensive and high-quality data that accurately measure the quantitative expression levels of the entire transcriptome (Leinonen et al., 2010).

3. Long-read cDNA RNA-seq and dRNA-seq

Long-read cDNA sequencing and dRNA-seq are becoming popular substitutes for short-read sequencing, allowing users to acquire enhanced data at the isoform level. Long-read technologies, such as those offered by Pacific Biosciences (PacBio) and Oxford Nanopore (ONT), enable the sequencing of entire individual RNA molecules as single molecules following their conversion to cDNA. This approach overcomes the limitations associated with short-read methods (Sharon et al., 2013; Cartolano et al., 2016; Oikonomopoulos et al., 2016). Long-read platforms offer the benefits of minimizing ambiguity in mapping sequence reads, detecting longer transcripts, and capturing a greater number of individual transcripts, resulting in a more comprehensive representation of isoform diversity (Engström et al., 2013). Nevertheless, long-read platforms presently experience reduced throughput and increased error rates in comparison to short-read platforms, which restricts their sensitivity and specificity for specific applications (Stark et al., 2019). ONT has developed a method called dRNA-seq (Jain et al., 2016; Jain et al., 2018), which enables the sequencing of RNA without the need for modification, cDNA synthesis, or PCR amplification (Garalde et al., 2018). This approach has the potential to overcome the limitations associated with the traditional method of sequencing mRNA. In this method, the library preparation from RNA includes a series of steps for ligating two adaptors. Initially, a duplex adaptor with an oligo (dT) overhang is attached and ligated to the RNA polyadenylation (poly (A)) tail, followed by an optional reverse-transcription process that enhances sequencing capacity. The second ligation step connects the preloaded sequencing adaptors containing the motor protein necessary for driving sequencing. Once prepared, the library is suitable for MinION sequencing, where RNA is sequenced directly from the 3′ poly(A) tail to the 5′ cap (Stark et al., 2019). Initial research has shown that dRNA-seq produces read lengths averaging around 1,000 bp and can achieve maximum lengths exceeding 10 kb (Garalde et al., 2018; Workman et al., 2019; Weirather et al., 2017). Long-read platforms possess the capability to identify isoforms that are not detected by short-read methods, especially in challenging-to-sequence yet medically significant regions, although they presently incur higher experimental expenses (Stark et al., 2019; Workman et al., 2019). The selection between short-read and long-read platforms is contingent upon the particular biological inquiry, as well as the computational resources at hand and the desired level of sensitivity and specificity for the experiment.

Table 1
An overview of the four major sequencing platforms for RNA-seq.

Use	Platform	Company	Template preparation	Sequencing technology	Key applications
Short-read cDNA sequencing	Illumina	Illumina	Bridge PCR	Sequencing by synthesis	• Differential gene expression • Whole transcriptome analysis
	Ion Torrent	454 Life Sciences	Emulsion PCR	Sequencing by synthesis	• Small RNA-seq • Single-cell RNA-seq • Spatial RNA-seq • Nascent RNA and translatome analysis • RNA structure and RNA–protein interaction analysis
Long-read cDNA sequencing	Single-molecule real-time sequencing (SMRT)	Pacific Biosciences	PCR	Sequencing by synthesis	• Identification of isoforms • de novo transcriptome analysis
	Nanopore sequencing	Oxford Nanopore Technologies	Without PCR	Nanopore	• Fusion and complex transcripts analysis
Direct RNA sequencing	Nanopore sequencing	Oxford Nanopore Technologies	Without PCR	Nanopore	• Identification of isoforms • de novo transcriptome analysis • Fusion and complex transcripts analysis • Detecting ribonucleotide modifications

4. RNA-seq data analysis

RNA-Seq is a versatile technology that may be employed for DGE, variant calling, variant prioritization, quantitative gene expression analysis, and validation of variants of unknown significance in the genome. The availability of computational analysis tools for RNA-seq has significantly increased over the past decade (Stark et al., 2019; Conesa et al., 2016; Saeidian et al., 2020). The selection of a specific tool should be determined by the intended purpose and the level of precision required for its application. Fig. 1 and Table 2 displays the typical procedure for analyzing RNA-seq data and the bioinformatics tools that are commonly employed in RNA-seq data analysis.

5. Single-cell and spatial RNA-seq

Bulk RNA-seq has greatly enhanced our understanding of biology, yet it faces challenges in accurately identifying distinct cell types and maintaining spatial information; understanding the complexities of biological systems requires both of these challenges. Techniques that help users surpass bulk RNA are comparable to standard RNA-seq protocols, but they allow for the exploration of distinct inquiries. Single-cell sequencing has revealed previously unidentified cell types, such as ionocyte cells, in extensively researched diseases that may be relevant to the pathology of cystic fibrosis (Montoro et al., 2018). Spatially resolved RNA-seq offers a valuable approach to gain a deeper understanding of cell–cell interactions within solid tissues, including the expression of fetal marker genes in adult heart tissue (Asp et al., 2017).

The process of scRNA-seq entails extracting individual cells from a sample using diverse techniques, such as micropipetting, flow sorting, or microfluidic isolation (Fig. 2A). Subsequently, the individual cells are subjected to reverse transcription to generate cDNA, which is subsequently labeled with unique molecular identifiers (UMIs) for the purpose of preparing RNA-seq libraries and conducting sequencing. Discrete cell populations are identified using quality control, DGE analysis, and 2D visualization (Stark et al., 2019; Tang et al., 2009; Stegle et al., 2015; Svensson et al., 2018). Conversely, the workflow for spatial RNA-seq (spatialomics) entails the utilization of spatial encoding or in situ transcriptomics. Spatial encoding methods entail the use of frozen tissue sections applied to microarray slides with oligoarrays or densely packed beads coated with oligos (Fig. 2B). The mRNA molecules disperse across the surface of the slide and bind to oligo-dT cDNA synthesis primers that contain UMIs and spatial barcodes. The spatialomics data is subjected to computational analysis, which involves mapping sequence reads to their corresponding spatial coordinates. This is

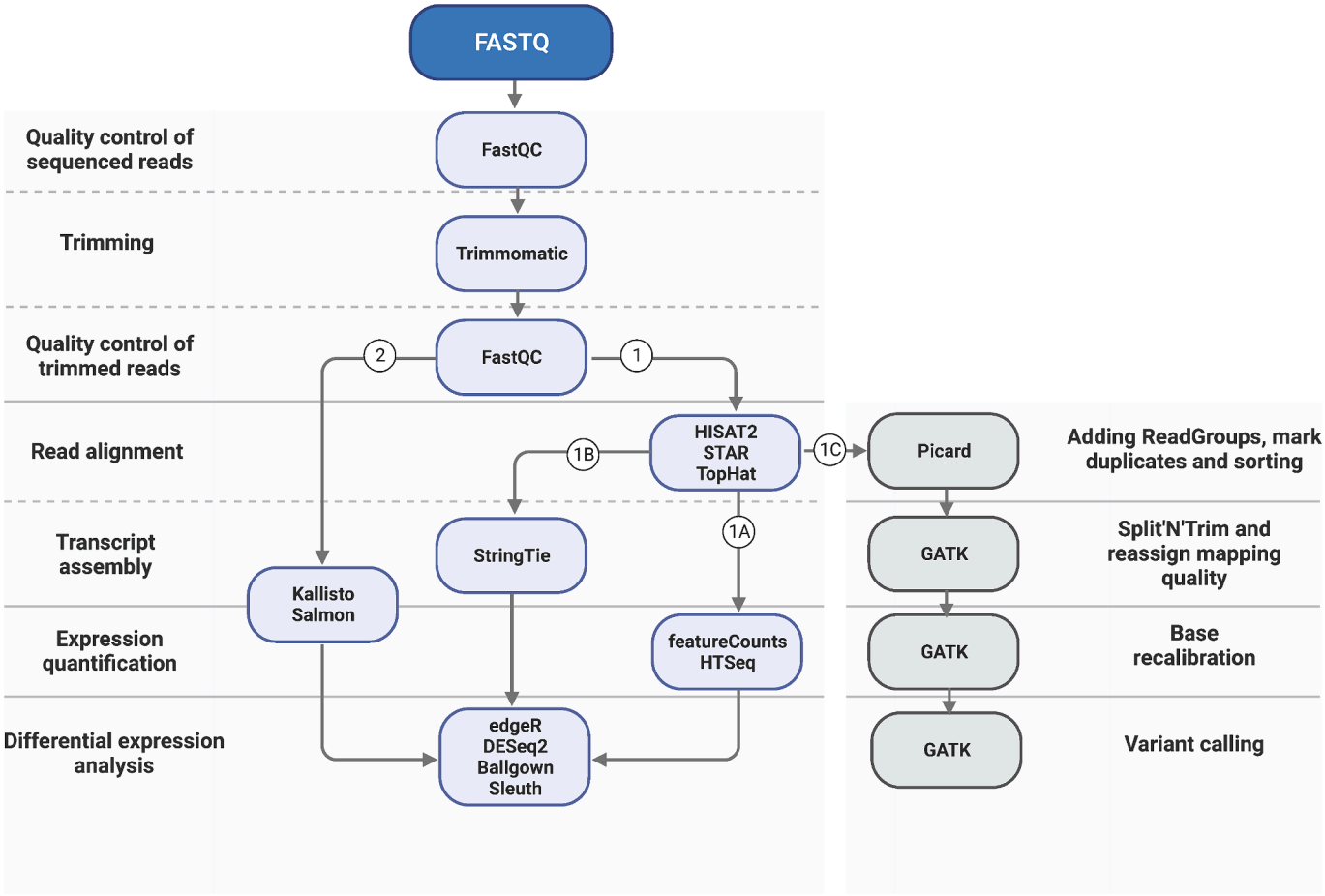

Fig. 1. Computational analysis of differential gene expression and variant calling. The analysis begins with preprocessing of the raw reads (FASTQ format). Four common workflows, namely, 1A, 1B, 1C, and 2, are shown. Workflow 1 involves the utilization of aligners like HISAT2, STAR, or TopHat to map reads to specific places in the genome using a reference genome. Subsequently, quantification tools like featureCounts or HTSeq are employed to assign reads to specific features (Workflow 1A). Alternatively, workflow 1B starts by aligning the reads, thereafter utilizing StringTie to construct a transcriptome model based on the alignment tools. Gene expression is ultimately represented through the utilization of tools such as edgeR, DESeq2, Ballgown, or Sleuth. These tools enable the generation of a comprehensive list of genes or transcripts that exhibit differential expression and can be further analyzed and interpreted. Workflow 2 utilizes advanced alignment-free tools, such as Kallisto and Salmon, to simultaneously construct a transcriptome and measure abundance. The results generated by these tools are often subjected to the same modeling process employed in workflow 1 to provide a compilation of genes or transcripts that exhibit differential expression. For variant calling via the GATK pipeline (workflow 1C), duplicate reads are identified and labeled, and then, variations are detected and subjected to filtering. Created with BioRender.com.

Table 2
Description of bioinformatic tools that are commonly employed in RNA-seq data analysis.

Tool	Description	Availability	Link	Reference
FastQC	A quality control tool that assesses the quality of raw sequencing data, including per-base quality scores, GC content, and sequence duplication levels.	Open-source	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/	(Andrews, 2010)
Trimmomatic	Used for read trimming and adapter removal to improve data quality by removing low-quality bases and adapter sequences.	Open-source	http://www.usadellab.org/cms/?page=trimmomatic	(Bolger et al., 2014)
HISAT2	A fast and accurate aligner for mapping RNA-seq reads to a reference genome.	Open-source	https://daehwankimlab.github.io/hisat2/	(Kim et al., 2019)
STAR	A RNA-seq aligner that performs spliced alignment and is efficient for detecting novel splice junctions.	Open-source	https://github.com/alexdobin/STAR	(Dobin et al., 2013)
TopHat	An aligner that identifies splice junctions and aligns RNA-seq reads to a reference genome.	Open-source	https://ccb.jhu.edu/software/tophat/index.shtml	(Kim et al., 2013)
Kallisto	A pseudoalignment-based tool that quantifies transcript abundances without full alignment to the genome.	Open-source	https://pachterlab.github.io/kallisto/	(Bray et al., 2016)
Salmon	Salmon estimates transcript-level abundances using lightweight alignment-free methods.	Open-source	https://combine-lab.github.io/salmon/	(Patro et al., 2017)
StringTie	Assembles transcripts and estimates their abundances from RNA-seq data.	Open-source	https://ccb.jhu.edu/software/stringtie/	(Kovaka et al., 2019)
featureCounts	Counts the number of reads that map to genomic features (e.g., genes, exons) for differential expression analysis.	Open-source	https://subread.sourceforge.net/featureCounts.html	(Liao et al., 2013)
HTSeq	Assigns reads to genomic features for gene expression quantification.	Open-source	https://htseq.readthedocs.io/en/master/	(Anders et al., 2014)
edgeR	A statistical package for identifying differentially expressed genes between conditions.	Open-source	https://bioconductor.org/packages/release/bioc/html/edgeR.html	(Robinson et al., 2010)
DESeq2	Performs differential expression analysis by modeling count data and accounting for library size differences.	Open-source	https://bioconductor.org/packages/release/bioc/html/DESeq2.html	(Love et al., 2014)
Ballgown	A tool for analyzing differential expression at the transcript level.	Open-source	https://www.bioconductor.org/packages/release/bioc/html/ballgown.html	(Frazee et al., 2015)
Sleuth	Provides statistical methods for differential transcript expression analysis.	Open-source	https://pachterlab.github.io/sleuth/	(Pimentel et al., 2017)
Picard	A suite of tools for manipulating and analyzing high-throughput sequencing data.	Open-source	https://broadinstitute.github.io/picard/	Refer to the relevant link
GATK	The Genome Analysis Toolkit, primarily used for variant calling but can also be applied to RNA-seq data.	Open-source	https://gatk.broadinstitute.org/hc/en-us	(McKenna et al., 2010)

performed following the completion of DGE analysis. The purpose of this analysis is to enable the visualization of differential spatial expression (Chen et al., 2017; Rodriques et al., 2019; Crosetto et al., 2015; Moor and Itzkovitz, 2017; Lein et al., 2017; Ståhl et al., 2016).

The most commonly used platforms for single-cell and spatial RNA sequencing are summarized in Table 3.

Computational methods are also rapidly advancing, and recent guidelines have been established for scRNA-seq experimental design (Luecken and Theis, 2019). Single-cell analysis workflows comprise a collection of independently developed tools. To facilitate seamless data transfer between these tools, single-cell platforms have been created based on consistent data formats (Luecken and Theis, 2019). These platforms serve as a foundation for constructing analysis pipelines. Within command line platforms, Scater (McCarthy et al., 2017) and Seurat (Butler et al., 2018) integrate easily with the wide range of analysis tools available through the R Bioconductor project (Huber et al., 2015). Scater excels in quality control and preprocessing, while Seurat is widely recognized as the most popular and comprehensive platform offering a vast array of tools and tutorials (Luecken and Theis, 2019). Notably, computational approaches that combine single-cell with spatially resolved transcriptomics have emerged, leading to the development of tools like cell2location (Kleshchevnikov et al., 2022) designed to conduct joint analyses on multiple scRNA-seq and spatial transcriptomic datasets.

6. Applications of RNA-seq in CeD research

The availability of genomic data, such as RNA-seq data, has greatly increased as a result of the widespread use of high-throughput sequencing technologies. RNA-seq, a crucial component of next-generation sequencing, has led to significant advancements in research on a wide range of domains, particularly in CeD research (Fig. 3).

6.1. Comparative transcriptomics

RNA-seq has played a key role in revealing the unique transcriptome profiles in individuals with CeD and various conditions. The significant

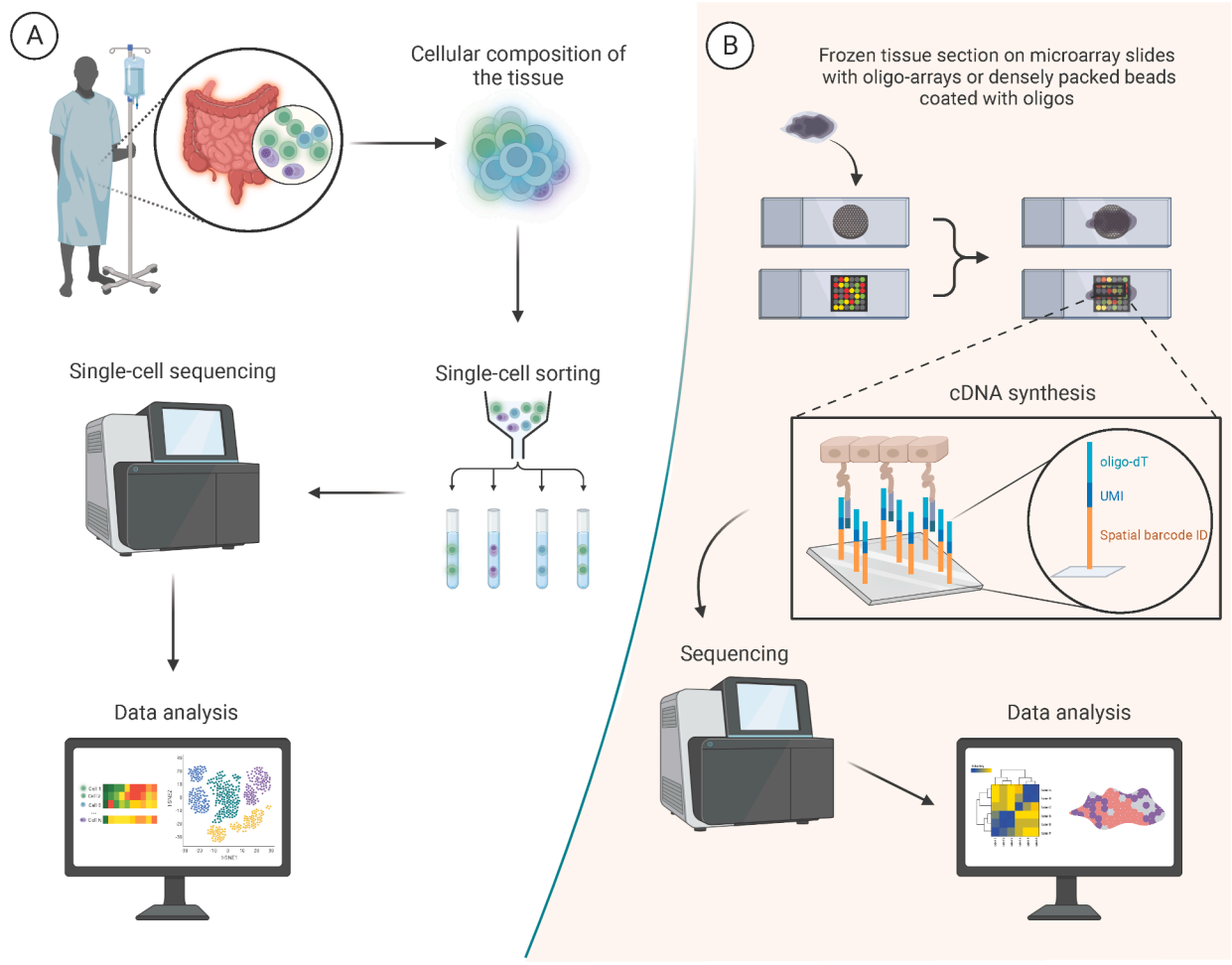

Fig. 2. The process of single-cell RNA sequencing (scRNA-seq) and spatial RNA-seq. (A) ScRNA-seq involves extracting individual cells from a sample using techniques like flow sorting. These cells are then reverse transcribed to generate cDNA, which is labeled with unique molecular identifiers (UMIs), for RNA-seq libraries and sequencing. Discrete cell populations are identified using quality control, differential gene expression analysis, and 2D visualization. (B) Spatial RNA-seq uses spatial encoding, using frozen tissue sections on microarray slides with oligoarrays or densely packed beads coated with oligos. The data is then analyzed computationally by mapping sequence reads to their corresponding spatial coordinates, enabling visualization of differential spatial expression. Created with BioRender.com.

Table 3
Summary of the most commonly used platforms for single-cell and spatial RNA sequencing.

Technique	Platforms	Reference
Single-cell RNA sequencing	Smart-Seq2	(Picelli et al., 2013)
	Fluidigm C1	(Xin et al., 2016)
	Drop-seq	(Macosko et al., 2015)
	inDrop-seq	(Klein et al., 2015)
	CEL-seq2	(Hashimshony et al., 2016)
	10x Genomics Chromium	(Zheng et al., 2017)
	Seq-Well	(Gierahn et al., 2017)
Spatial RNA sequencing	Seq-FISH	(Eng et al., 2019)
	Slide-seq	(Rodriques et al., 2019)
	10x Genomics Visium	(Vickovic et al., 2019)
	NanoString GeoMx	(Merritt et al., 2020)
	NanoString CosMx	(He et al., 2022)

findings from comparative transcriptomics using RNA-seq analysis are outlined in Table 4.

RNA-seq can elucidate the underlying mechanisms of duodenal inflammation in individuals with common variable immunodeficiency (CVID) and CeD by revealing their distinct transcriptome profiles. The RNA and protein profiles of CVID patients with duodenal inflammation and CeD exhibited notable differences. Although CVID, which is associated with duodenal inflammation, and CeD share histological similarities, they exhibit distinct RNA regulation (Kaarbø et al., 2023). In addition, researchers have utilized RNA-seq to identify the genes and pathways that are unique to environmental enteric dysfunction (EED), despite the notable molecular similarities shared with celiac disease. These include an increase in the expression of the *DUOX2* and *LCN2* genes, which are associated with the natural defense of the body against microbes. Conversely, there was a decrease in the expression of metallothioneins (MT family) and aldo–keto NADPH-dependent reduction genes (AKR1C family), which play a role in detoxifying environmental compounds (Haberman et al., 2021). Severe acute malnutrition (SAM) enteropathy is characterized by immune activation, which leads to hyperplastic enteropathy. This condition is associated with specific alterations in nutrient transport and xenobiotic metabolism. RNA-seq analysis revealed slight overlap between SAM and pediatric Crohn's disease but not as much overlap as between SAM and CeD (Chama et al., 2019). In another study, seven differentially expressed genes (DEGs) common to both CeD and SARS-CoV-2 infection were identified, namely, *KLRG1*, *NLRP3*, *IRAK3*, *MAFB*, *CD79B*, *MBP*, and *DDIT4* (Nashiry et al., 2021). These findings could help us understand the complex relationship between COVID-19 and CeD (Nashiry et al., 2021). RNA-seq also

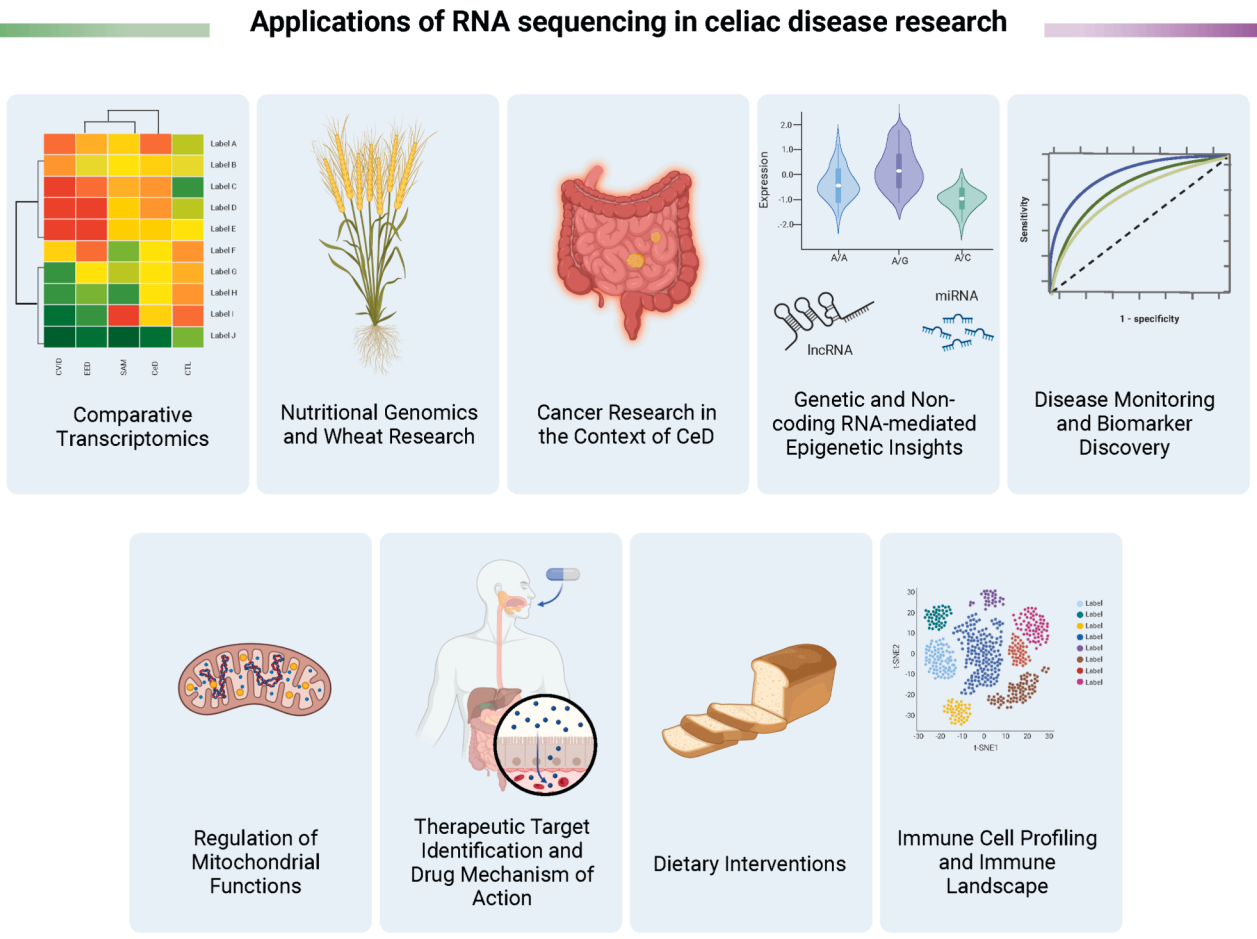

Fig. 3. Applications of RNA sequencing (RNA-seq) in celiac disease research. RNA-seq has made great contributions to CeD research, including studies on comparative transcriptomics, nutritional genomics and wheat research, cancer research in the context of CeD, genetic and noncoding RNA-mediated epigenetic insights, disease monitoring and biomarker discovery, regulation of mitochondrial functions, therapeutic target identification and drug mechanism of action, dietary factors, immune cell profiling and the immune landscape. Created with BioRender.com.

revealed distinct gene expression patterns in patients with eosinophilic duodenitis (EoD) and CeD. EoD and CeD suppressed metabolic processes and transporter activity in the duodenum, while CeD increased activity in pathways associated with type 1 immunity and extracellular matrix pathways (Shoda et al., 2023).

While RNA-seq has been widely used in comparative transcriptomics, it's important to consider the potential limitations associated with this approach. These include issues such as small sample size, individual heterogeneity, diverse treatments among subjects, use of older healthy controls for comparison, reliance on bulk biopsies rather than single-cell separation, and inadequate data related to gestational age, birth weight, and microbial information. It is imperative to address these concerns by conducting research that incorporates age-matched controls, explores advancements in single-cell separation technology, and collects more comprehensive datasets. Furthermore, repeating studies, relying on larger sample sizes, and implementing specific treatment regimens are also crucial steps recommended for future investigations.

6.2. *Nutritional genomics and wheat research*

Research using RNA sequencing has led to significant findings in nutritional genomics and wheat research. A method for sequencing alpha-gliadin transcripts, including the three main epitopes for CeD and their variations, was developed using 454 RNA amplicon sequencing. This method was used to evaluate developing grains on 61 distinct durum wheat cultivars and accessions, creating a prescreening tool for assessing the immunogenicity of CeD (Salentijn et al., 2013). Moreover, wheat RNA-seq analysis revealed complete transcripts of gliadins, the primary carriers of epitopes associated with CeD (Wang et al., 2017a). This study employed transcriptomic and proteomic techniques to investigate the expression and role of gliadin, yielding novel findings for genetic and genomic investigations and guiding future research endeavors (Wang et al., 2017a). In another study, bulked segregant RNA-Seq (BSR-seq) was used to identify the genetic defect responsible for a recessive, low-prolamin mutation in diploid barley, aiming to create a wheat variety with reduced gluten content (Moehs et al., 2019). BSR-seq is a very effective approach that combines RNA-seq with bulked segregant analysis (BSA). BSA is a technique that can be employed to detect markers associated with a particular gene or genomic area by utilizing two pools of DNA. Each pool, or bulk, comprises individuals who share the same trait or genomic region but differ at unrelated regions (Lin et al., 2021). BSR-seq is used for both profiling DGE and rapid gene/quantitative trait loci (QTL) mapping (Wang et al., 2017b; Hao et al., 2019). Additionally, RNA-seq identified protein compensation mechanisms in the E82 RNAi wheat line and its wild type during grain filling, revealing that inhibiting gliadin controls the production of seed storage proteins in the grain through a complex system of potential transcription factors (Marín-Sanz and Barro, 2022).

The PacBio third-generation DNA sequence analysis technology provides a promising solution for overcoming the challenges associated with second-generation sequencing of gliadin genes (Zhang et al., 2014). It offers precise detection of transcribed mRNAs from these genes along with retrieval of complete sequence information. By integrating third-

Table 4
Key findings from comparative transcriptomics using RNA-seq analysis.

Conditions compared	Main findings from RNA-seq	Insights provided	Reference
CVID vs CeD	Revealed distinct transcriptome profiles and differences in RNA regulation despite histological similarities	Understanding molecular differences between conditions with similar presentations	(Kaarbø et al., 2023)
EED vs CeD	Identified genes/ pathways unique to EED like DUOX2/LCN2 upregulation and metallothionein downregulation	Identifying condition-specific molecular signatures and pathways	(Haberman et al., 2021)
SAM vs Crohn's disease vs CeD	Showed overlap between conditions. Identified alterations in nutrient transport and xenobiotic metabolism in SAM	Elucidating relationships and commonalities between SAM, Crohn's disease, and CeD	(Chama et al., 2019)
SARS-CoV-2 infection vs CeD	Identified seven differentially expressed genes common to both, providing clues about their interaction	Providing molecular insights into link between COVID-19 and CeD	(Nashiry et al., 2021)
EoD vs CeD	Revealed distinct expression patterns between conditions in metabolic processes and pathways involved	Understanding differences in molecular pathophysiology of EoD and CeD	(Shoda et al., 2023)

CVID, common variable immunodeficiency; CeD, celiac disease; EED, environmental enteric dysfunction; SAM, severe acute malnutrition; EoD, eosinophilic duodenitis.

generation RNA sequencing with proteomic analysis and advanced bioinformatics tools, it becomes possible to gain deeper insights into complex gliadins as well as the resulting CeD epitopes present in bread wheat. This integrated approach facilitates more thorough investigations into the role of gliadins while also supporting efforts towards developing healthier varieties of wheat (Altenbach et al., 2010).

6.3. Cancer research in the context of CeD

Through the use of RNA-seq, two distinct molecular subtypes have been consistently observed in a series of small bowel carcinomas associated with CeDs. These subtypes are characterized by microsatellite instability, immune and mesenchymal characteristics. The microsatellite instability-immune subtype is likely associated with less aggressive tumor behavior, while the mesenchymal subtype is associated with more unfavorable tumor behavior (Rizzo et al., 2020). Moreover, functional analyses revealed biological function categories related to cancer and the inflammatory response (Rizzo et al., 2020). Patients with CeD may show an increase in the frequency of abnormal cells; called intraepithelial intracellular CD (iCD)3+ surface CD (sCD)3-CD7+ CD56−, in their duodenum; this disease is known as type II refractory celiac disease (RCDII) and is classified as low-grade lymphoma (Cellier et al., 1998). Using high-dimensional single-cell and spatial technologies, it has been shown that the RCDII exhibits significant variation in the abnormal cell population, both within and between tumors (Dieckman et al., 2022). This highlights the importance of developing individualized immune profiles for personalized treatment. Furthermore, in another study, RNA-seq was used to identify the gene drivers responsible for lymphomagenesis associated with CeD. This discovery revealed potential targets for treatment (Cording et al., 2022).

RNA-seq has provided valuable insights into cancer research in the context of CeD, but it is crucial to acknowledge the constraints of this approach. The complexity and diversity of data obtained from RNA-seq analysis can pose challenges in interpretation, potentially leading to biases in results. Additionally, RNA-seq may not fully capture the intricacies of biological processes linked to cancer and inflammatory responses. Additional omics methods and functional studies may be necessary for a more comprehensive understanding of disease mechanisms. To enhance its applicability in fields such as epigenetics and proteomics, integrating information from various omics studies is essential (Hong et al., 2020). For instance, an assay utilizing hyperactive Tn5 transposase named Transposase-Accessible Chromatin using sequencing (ATAC-seq) was created to identify accessible chromatin regions and binding sites for transcription factors (Buenrostro et al., 2013). Integrating ATAC-seq with RNA-seq enables us to clarify transcription factor-targeted genes along with their transcripts (Hong et al., 2020). Moreover, progress has been achieved in chromatin conformation capture analysis technology (Simonis et al., 2007) and its variations like Capture Hi-C (Mifsud et al., 2015) for identifying chromatin structure as well as unknown interacting regions. When combined with genomic methylation and chromatin accessibility analysis, RNA-seq discloses the influence of epigenetic alterations on tumor heterogeneity providing insight into personalized treatment plans for cancer patients. Lately, the innovative technologies Cellular Indexing of Transcriptomes and Epitopes by Sequencing (CITE-Seq) and CRISPR droplet sequencing (CROP-Seq) integrate single-cell RNA sequencing with cell phenotype research alongside CRISPR screening showing significant potential for identifying tumor cell epitope protein indices while also screening drug-resistant tumor cells (Zhang et al., 2021).

6.4. Genetic and noncoding RNA-mediated epigenetic insights

Interferon-gamma (IFNγ)-induced genes of unknown expression were found to be highly concentrated in autoimmune susceptibility loci, specifically those associated with CeD (Molinie et al., 2014). Using RNA-seq data, the discovery of a complex transcriptional connection between CeD susceptibility genes and IFNγ has provided new insights into the mechanisms and pathways of CeD (Kumar et al., 2015). Moreover, it has been shown that CeD is associated with 118 prioritized genes and 172 combinations of genes that are affected by these genes in trans. Systematic prioritization of candidate genes at disease loci has identified *TRAFD1* as a key regulator of IFNγ signaling in CeD (van der Graaf et al., 2020). The analysis of gene expression profiles in cell types relevant to the disease revealed tissue-specific biological functions. RNA-seq of CD326+ epithelial cells from patients with CeD revealed distinct expression patterns of 1194 genes in comparison to those of individuals without CeD. Genes related to the plasma membrane, extracellular exosomes, and the extracellular region were found to be significantly more abundant in the epithelial cells of patients with CeD (Sowińska et al., 2020). This discovery provides valuable knowledge about the molecular mechanisms underlying the impairment of the epithelial barrier in CeD. Additional research identified 112 expression quantitative trait loci (eQTL) genes within 32 out of 42 non-HLA CeD loci (Withoff et al., 2016). Approximately 10 % of these genes were classified as regulatory RNAs. Examination of publicly available expression data revealed that numerous *cis*-eQTL genes exhibited specificity for particular cell types. Cis-eQTL analysis gives priority to genes that are commonly regarded as 'causal' disease genes (Withoff et al., 2016). However, the interpretation of these results is challenging due to the influence of tissue composition and the context-specific effects of eQTLs. Notably, among the 112 *cis*-eQTL genes, 13 are classified as long noncoding RNAs (lncRNAs), which emphasizes their significance in the development of CeD pathology (Withoff et al., 2016).

Whole-genome RNA sequencing was utilized to identify differentially expressed lncRNAs (including those that are known, annotated, or novel) in leukocytes from patients with CeD. In addition, novel lncRNA regions, such as *IFNG-R-49*, were discovered in close proximity to transcriptional enhancers in leukocytes, particularly superenhancers, as opposed to regular enhancers (Aune et al., 2017). The genome contains both previously identified and newly discovered lncRNA loci that are located near genetic variants associated with autoimmune diseases, such

as CeD (Aune et al., 2017). Genetic and environmental factors can modify the function of enhancers, including the expression of lncRNAs. This can lead to changes in cellular characteristics that contribute to the risk and development of diseases (Aune et al., 2017). The Epigenome Roadmap project detected RNA-seq signals in the genomic region of *Carlr*, suggesting the presence of a human lncRNA that is similar to mouse *Carlr* (Castellanos-Rubio et al., 2017). Elevated levels of the *Carlr* transcript were observed in the cytoplasm of human CeD patient samples, along with increased expression of genes associated with the NF-kB pathway (Castellanos-Rubio et al., 2017). Upon re-examination of the raw transcriptomic data, it was discovered that there were dysregulated long lncRNAs in the CeD cohort as well as altered expression of *GATA6-AS1* (Sosnovski et al., 2023). Moreover, researchers have discovered and examined *lnc13*, a lncRNA associated with CeD. The small intestinal biopsy samples of patients with CeD exhibited downregulation of *lnc13* (Castellanos-Rubio et al., 2016). *Lnc13* also influences the expression of genes associated with inflammation in CeD (Castellanos-Rubio et al., 2016). The following transcripts were found to be significantly enriched in noncoding regions associated with CeD-associated single nucleotide polymorphisms (SNPs): *RP11-98D18.3*, *AL450992.2*, *RP11-509E16.1*, *LINC00877*, *LUCAT1*, *RP11-430C7.4*, *CHRM3-AS2*, *AC007278.3*, *AC008063.2*, *LINC00861*, *CTC-378H22.2*, *RP3-395M20.8*, and *RP11-234B24.4* (Tuomela et al., 2016). In addition, expression profiling and SNP data indicated that Th17 cells, a subset of CD4+ cells that secrete IL17, are involved in autoimmune diseases like CeD (Tuomela et al., 2016). A lack of Th17 cells increases susceptibility to infections (Korn et al., 2009). Another investigation revealed that 42 SNPs linked to autoimmune diseases directly affect 53 noncoding RNA genes, including lncRNAs and microRNAs (miRNAs) (Ricaño-Ponce et al., 2016). The noncoding genome plays a role in autoimmune diseases, and SNPs are linked to functional regulatory elements. RNA-seq data provide a comprehensive understanding of the transcriptome, highlighting the superiority of RNA sequencing over conventional microarrays (Ricaño-Ponce et al., 2016).

A study using prediction tools and genome-wide association study (GWAS) data identified 34 SNPs that affect miRNA binding sites in genes related to CeD (de Almeida et al., 2018). The DEG-miRNA interaction networks were also discovered using RNA-seq data, enhancing our understanding of the molecular and cellular interactions between CeD and SARS-CoV-2 infection, aiding in therapeutic targeting and strategies (Nashiry et al., 2021). A study revealed that miRNAs linked to CeD can control gene expression in the small intestine, impacting immune and metabolic processes (Tan et al., 2021b). Certain miRNAs, such as miR-155-5p, disrupt genes influenced by genomic variants, contributing to the abnormal CeD transcriptome, such as that of *STAT1* (Tan et al., 2021b). Furthermore, RNA-seq revealed numerous circulating miRNAs that serve as early indicators of the development of CeD. A more detailed analysis of these miRNAs is provided in the subsequent section.

Although RNA-seq has shown promise in evaluating genetic and noncoding RNA-mediated epigenetic information, several limitations need to be addressed. These include the difficulty in assigning functions to certain genes, the need for further validation and exploration, and the reliance on predicted interactions. Future research should focus on conducting functional studies, improving the sensitivity and dynamic range, addressing technical variability, and validating findings in larger cohorts.

6.5. Disease monitoring and biomarker discovery

RNA-seq analysis of gluten-induced transcriptomic alterations, persistent disease patterns in gluten-free patients, and gene transcript levels associated with mucosal injury in CeD patients was used to identify biomarkers and therapeutic targets for CeD (Dotsenko et al., 2021). Another study identified specific gene expression patterns, including those of *CDH18*, *CXCL9*, *CXCL10*, *GBP5*, *IFI27*, *IFNG*, and *UBD*, in the small intestines of CeD patients, suggesting that specific biological pathways are involved in disease development (Banerjee et al., 2023; Bragde et al., 2018). RNA-seq has also identified pathways and biomarkers associated with active disease and mucosal recovery, with CeD patients showing distinct innate immunity genes (*IL12R*, *ITGAM*, and *IGSF4*) (Leonard et al., 2019). RNA-seq analysis of organoids revealed distinct gene expression variations associated with the development of CeD. The utilization of RNA-seq allowed for the identification of specific gene expression patterns and the discovery of potential biomarkers (Freire et al., 2019). The mucosa of individuals with active CeD exhibited a newly identified immune response and cell adhesion genes, such as *IL37*, *CCL25*, *MUC6*, *CLDN18*, and *CCL24* (Freire et al., 2019). Interestingly, an RNA-seq study revealed that stabilized whole blood is not an appropriate sample for clinical diagnosis of CeD based on individual genes. Nevertheless, the possibility of utilizing a gene expression panel focused on specific pathways for diagnostic purposes is worth exploring, although additional research is necessary (Bragde et al., 2020).

Noncoding RNAs could also serve as early indicators of CeD. RNA-seq was used to detect miRNAs in the bloodstream and investigate their potential as biomarkers for CeD research. Prior to the diagnosis of CeD, the expression of certain miRNAs, namely, *miR-21-3p*, *miR-374a-5p*, *miR-144-3p*, *miR-500a-3p*, *miR-486-3p*, *let-7d-3p*, *let-7e-5p*, and *miR-3605-3p*, exhibited alterations. A diet free of gluten restored the expression of miRNAs, including *miR-150-5p*, *miR-150-3p*, *miR-1246*, *miR-342-3p*, *miR-375-3p*, and *let-7a-5p*, to normal levels (Tan et al., 2021a). Moreover, the serum miRNAs of individuals with CeD and those who were fed a gluten-free diet (GFD) were subjected to sequencing. The adoption of a GFD successfully normalized the dysregulated levels of miRNAs (*miR-192-5p*, *miR-215-5p*, and *miR-125b-5p*), which have the potential to serve as biomarkers for this disease (Felli et al., 2022). The fecal samples of CeD patients showed changes in miRNA and microbial profiles as a result of following a gluten-free diet. There was a correlation between the duration of GFD and the levels of *miR-4533-3p* and *miR-2681-3p* (Francavilla et al., 2023). The study indicated that individuals with CeD exhibit distinct molecular patterns in their stool, which could serve as biomarkers for identifying the efficacy of a GFD (Francavilla et al., 2023). In relation to lncRNAs, *GATA6-AS1* shows promise as a biomarker and target for enhancing epithelial function and promoting mucosal healing through the use of RNA-seq (Sosnovski et al., 2023).

Disease monitoring and biomarker discovery studies may have certain inherent limitations, including restricted sample sizes; inadequate investigations of environmental factors and the gut microbiota; and the need for larger cohorts, validation, and functional studies. Subsequent investigations should aim to overcome these limitations by conducting studies on a larger scale, utilizing independent cohorts, incorporating functional studies, and taking into account potential confounding factors.

6.6. Regulation of mitochondrial functions

Recent studies have shed light on the intricate role of *GATA6-AS1* in regulating mitochondrial functions within epithelial cells (Sosnovski et al., 2023). The *GATA6-AS1* lncRNA interactome exhibited significant enrichment of mitochondrial proteins, such as TGM2. TGM2 serves as an autoantigen in CeD and is triggered in individuals with ulcerative colitis, Crohn's disease, or CeD. Further research has delved into the impacts of suppressing *GATA6-AS1* on the activation of TGM2, leading to a cascade of effects on mitochondrial membrane potential, respiration, and metabolites crucial for aerobic respiration, all of which are closely linked to mucosal inflammation (Sosnovski et al., 2023). This connection underscores the potential significance of *GATA6-AS1* as a target for modulating mitochondrial functions in epithelial cells, particularly through its influence on TGM2 levels. Additionally, in a previous pilot study, biopsies were taken from the distal duodenum of one CeD patient and one control. Approximately 8,000 cells were analyzed from each

sample. The results revealed that the cells grouped into 10 different clusters, with a subset of cells from the CeD patient displaying high mitochondrial content, which could potentially be a significant molecular feature of the disease. However, it's important to note that this study is in its early stages and these findings are very preliminary (De Leo et al., 2023).

These findings open up new avenues for understanding the intricate interplay between mitochondrial function and the pathophysiology of CeD, pointing towards a potential link between mitochondrial gene expression and the underlying mechanisms of the disease. Larger-scale validation studies, functional studies, and investigations of regulatory networks involving lncRNAs could provide further insights into the pathogenesis of CeD.

6.7. Therapeutic target identification and drug mechanism of action

Larazotide acetate (LA) is a synthetic peptide consisting of eight amino acids. It is recognized for its ability to regulate tight junctions (TJs) by closing open or "leaky" interepithelial junctions (Slifer et al., 2021). LA has been found to be effective and safe in reducing gastrointestinal symptoms in patients who are undergoing a gluten challenge (Hoilat et al., 2022). While it may not provide a complete cure for CeD, it could potentially be used alongside the standard gluten-free diet as a supportive rather than alternative treatment option (Hoilat et al., 2022). Using RNA-seq, Jin et al. identified crucial signaling pathways that exhibited differential expression in cells treated with LA (Jin et al., 2020). These pathways influence the orientation of cells, progression of the cell cycle, binding of Ras/Rho GTPases, activity of protein serine/threonine kinases, and area of epithelial junctions. These findings provide insight into the molecular mechanism by which LA protects the integrity of TJs and its potential application in significant gastrointestinal diseases. Dotsenko et al. employed RNA-seq to assess the efficacy of the TG2 inhibitor ZED1227 in safeguarding against gluten-induced intestinal harm in individuals with CeD (Dotsenko et al., 2022). ZED1227 effectively inhibited gluten-induced transcriptional alterations, reversed mucosal abnormalities and inflammation, and standardized cell differentiation-related gene patterns (Dotsenko et al., 2022). RNA-seq was used to evaluate the effectiveness of TG2 inhibition in treating CeD, demonstrating its potential for treating this disease. Another study employed RNA-seq to investigate the synergistic activation of intestinal intraepithelial cytotoxic T cells by IL15 and IL21, which resulted in enhanced transcriptional activity, proliferation, and cytolytic activity (Ciszewski et al., 2020). RNA-seq revealed the molecular mechanisms of IL15 and IL21 in CeD and the potential of BNZ-2 therapy. BNZ-2 specifically suppressed these effects without affecting the effects of IL2 (Ciszewski et al., 2020). Freitag et al. studied the induction of tolerance in poly(lactide-co-glycolide) nanoparticles encapsulating gliadin protein (TIMP-GLIA) using RNA-seq and identified differentially expressed genes in spleen cells (Freitag et al., 2020). A total of 77 genes exhibited differential expression, 15 of which exhibited statistically significant differences between TIMP-GLIA and control mice. The pathways involved include antigen presentation, B-cell activation and differentiation, peptide loading onto MHC II molecules, and T-cell secretion of cytokines (Freitag et al., 2020). RNA-seq was used to analyze alterations in signature genes and pathways following resveratrol administration in individuals with CeD (Yu et al., 2022). Resveratrol, a polyphenol present in medicinal plants, grapes, and red wine (Chimento et al., 2019), has demonstrated potential in the prevention and treatment of chronic inflammatory diseases by eliminating free radicals and controlling the activity of different enzymes, such as COX and iNOS (Luca et al., 2020). The analysis of RNA-seq data revealed that resveratrol downregulated six genes (*Orm1*; *Nr0b2*, *Fbxo27*, *Fgf15*, *Fabp4*, and *Amy1*) and upregulated three genes (*Ubd*, *Lat*, and *Aire*) (Yu et al., 2022). The DEGs were classified, and subsequent GO enrichment analysis revealed their involvement in biological processes associated with reactions involving oxygenated compounds, phylogeny, and organic nitrogen compounds. According to KEGG enrichment analysis, the DEGs were primarily linked to pathways such as PPAR signaling, nitrogen metabolism, AMPK signaling, and FoxO signaling (Yu et al., 2022). RNA-seq was used to elucidate the specific molecular mechanisms through which resveratrol mitigates the symptoms of CeD by modulating specific genes and pathways.

The application of RNA-seq in therapeutic target identification and drug mechanism of action may be constrained by the need for gene and pathway validation, the intricacy of functional implications, limitations in data analysis, and the necessity of investigating dynamic changes over time. To overcome these limitations, one could employ validation, functional studies, advanced data analysis, and longitudinal studies. Additional investigations are needed to validate the results in human subjects, as most related studies have focused primarily on cell lines and individual cells. By utilizing more extensive in vivo models and conducting direct studies on patients, researchers can obtain more pertinent insights with clinical significance.

6.8. Applications of RNA-seq in relation to dietary interventions

RNA-seq has been used in CeD research to detect early changes in gene expression caused by gluten in duodenal biopsies; patients who adhere to a strict GFD still exhibit molecular disease patterns, and a relationship between gene transcript levels and gluten-induced damage to the intestinal lining has been identified (Dotsenko et al., 2021). Moreover, substantial alterations in the transcriptome of immune cells within CeD lesions were assessed by utilizing single-cell transcriptomic analysis. Compared with control patients, CeD patients on a GFD exhibited only partial restoration of their immune landscape (Atlasy et al., 2022). Another study revealed that a GFD can potentially rectify imbalanced levels of miRNA and restore miRNA expression to a state similar to that of unaffected individuals, thereby restoring a normal intestinal phenotype (Felli et al., 2022). Furthermore, after participants began a GFD, the expression of six of these miRNAs returned to normal levels, indicating that these miRNAs may serve as biomarkers for assessing the response to a GFD (Tan et al., 2021a). Several miRNAs exhibited abnormal regulation in treated CeD patients, with certain levels showing a correlation with adherence to a GFD. Over the course of the diet, the miRNA expression levels of CeD patients treated with some genes were similar to those of healthy controls (Francavilla et al., 2023). An RNA-seq study on gluten exposure in individuals with CeD revealed that gluten exposure significantly increased the activation of both innate and adaptive immune response pathways (Yohannes et al., 2020). Patients with CeD, regardless of the treatment they received, exhibited consistent modifications in pathways related to tight junctions, olfactory transduction, unsaturated fatty acid metabolism, amino acid metabolism, and microbial infection. An analysis was also conducted to predict upstream regulators using genes that were differentially expressed. The analysis revealed that regulators near loci associated with CeD, particularly *SMARCA4* and *CSF2*, were consistently activated (Yohannes et al., 2020).

Although RNA-seq has shown promise in dietary interventions related to CeD, the miRNAs identified as biomarkers for CeD need further validation in independent studies, and the sensitivity and specificity of individual miRNA markers should be assessed in larger sample sets. The small sample size and unavailability of samples for some patients and healthy controls are also potential limitations. To address these limitations, larger sample sizes should be considered, and efforts should be made to ensure the availability of samples from all participants. Deeper sequencing methods and complementary techniques such as quantitative PCR could increase the sensitivity and provide a more robust assessment of gene expression levels. Future research could also focus on conducting larger-scale studies with a more diverse sample of CeD patients, investigating the functional activity of identified markers, exploring the roles of other small noncoding RNAs, and identifying additional markers for GFD monitoring over time.

6.9. Immune cell profiling and the immune landscape

RNA-seq, specifically scRNA-seq, has the ability to offer a comprehensive understanding of the immune landscape and cellular diversity in CeD patients. Researchers have employed unbiased scRNA-seq to investigate the diversity of CD45+ immune cells in the human small intestine (Atlasy et al., 2022). The study revealed modified myeloid cell transcriptomes in active celiac lesions, with significant changes in CD4+ and CD8+ T-cell transcriptomes. A diminished native intraepithelial lymphocyte population was also detected in individuals with CeD. The immune profile of those with CeD who adhered to a GFD was only partially regenerated compared to that of control samples (Atlasy et al., 2022). This study underscores the importance of understanding the key immune system components in CeD due to its complex disease, as these components impact both the digestive tract and systemic symptoms.

Researchers have identified gluten-specific CD4+ T cells in CeD patients, which exhibit distinctive behaviors due to gluten exposure (Christophersen et al., 2021). Using RNA-seq data and mass cytometry, they identified markers for identifying pathogenic T cells in CeD. These markers are crucial for monitoring cells and developing drugs (Christophersen et al., 2021). The unique nonproliferative characteristics of gluten-specific CD4+ T cells in CeD have also been recognized using RNA-seq (Christophersen et al., 2019). Like activated effector memory T cells, gluten-specific CD4+ T cells also exhibit unique transcriptomic profiles (Yao et al., 2021). They exhibit varying levels of gene expression related to T-cell receptor signaling, translation, cell death, fatty acid transport, and redox potentials (Yao et al., 2021). In a previous study. During the examination of the transcriptome of CD4+ T cells, *IFNy* exhibited the highest level of upregulation in CD patients compared to controls (Quinn et al., 2015). A study also revealed a robust correlation between CeD and a cluster of genes regulated by *BACH2*; confirming the involvement of BACH2 in T-cell differentiation and the prevention of autoimmune diseases (Quinn et al., 2015). In the analysis of individual RNA molecules in peripheral blood mononuclear cells (PBMCs) obtained from children with CeD, researchers found consistent cell proportions over time and under various health conditions (Ramírez-Sánchez et al., 2022). However, there were variations in the expression of genes (e.g., *CD52*, *SELL*, *S100A4*, *NFKB2*, and *NFKBIA*) within CD4+ T cells between CeD patient samples before and after disease onset, as well as between patients and control subjects. This study also identified biomarkers that can be used to diagnose CeD, particularly in the NK cell compartment (such as *GZMA*, *GZMM*, *PRF1*, *TXNIP*, and *TAGAP*), providing alternative methods for diagnosing CeD without the need for a biopsy. This study revealed the importance of TNF pathways in the development and progression of CeD (Ramírez-Sánchez et al., 2022), highlighting the potential of scRNA-seq in understanding this process.

A previous study revealed the presence of plasma cells (PCs) in the small intestinal lesions of patients with active CeD through RNA-seq analysis (Snir et al., 2019). These PCs include genes that encode inflammatory mediators, receptors, costimulatory molecules, and HLA class II molecules. These cells generate inflammatory cytokines and chemokines, exhibit responsiveness to environmental stimuli, and interact with other immune cells, specifically CD4+ T cells (Snir et al., 2019). Transcriptome analysis of CeD revealed that the autoantigen TG2 has a negligible effect on the transcription of autoreactive B cells (Du Pré et al., 2020). These results support a model of CeD in which gluten-reactive T cells assist autoreactive TG2-specific B cells through the formation of gluten-TG2 complexes. This model also highlights a general mechanism of autoimmunity in which autoantibodies are produced by uninformed B cells when assisted by T cells (Du Pré et al., 2020). RNA-seq analysis has also revealed the role of longevity, clonal relationships, and transcriptional programs in CeD-specific PCs. One study revealed antigen-dependent V-gene selection and stereotypic antibodies (Lindeman et al., 2021), and variations in immunoglobulin genes impact the production of antibodies specific for CeD. The study identified transcriptional variations in short-lived and long-lived PCs specific to CeD and found a significant buildup of short-lived CD19+ CD45+ cells in disease-specific gut plasma cells (Lindeman et al., 2021).

RNA-seq has been used to study the transcriptional activity of IE-CTLs in response to the cytokines IL15 and IL21 (Ciszewski et al., 2020). It has also been used to identify DEGs, construct immune infiltration networks, and verify the occurrence of pyroptosis. ScRNA-seq analysis has expanded our understanding of CeD through the exploration of transcriptome profiles in duodenal tissue. γδT cells with high *IFN-γ* expression have been identified as relevant cells associated with pyroptosis (Chen et al., 2023). Another study also revealed differences in the TCR repertoire between individuals in good health and those with CeD, with specific V-region genes being more commonly used in CeD (Patrick et al., 2021). Two distinct clusters of CD8+ T cells exhibited increased activity in individuals with CeD, as indicated by an activated, cytotoxic transcriptional profile and high expression of immune checkpoint molecules and transcription factors (Patrick et al., 2021). Interestingly, RNA-seq analysis revealed KIR+ CD8+ T cells as targeted eliminators of gliadin-specific CD4+ T cells (Li et al., 2022). A higher frequency was associated with more severe disease and differential expression of 963 genes in patients with higher frequencies (Li et al., 2022).

The studies highlighted limitations in scRNA-seq data, such as focusing on specific markers and phenotypic changes without considering functional implications; requiring integration with other omics data, such as single-cell proteomics or spatial transcriptomics; excluding noncoding RNAs; not considering different T-cell subpopulations and time points; and potentially confounding effects of autoantigen exposure. Future research should address these limitations to improve the understanding and clinical significance of the findings.

7. Conclusions

High-throughput RNA-seq technology has revolutionized the study of the transcriptome, providing valuable insights into new research domains. However, there are still limitations in its application in CeD research. Future research should focus on larger sample sizes, validation and functional studies, user-friendly analysis tools, improved sensitivity and dynamic ranges, and validation of findings in larger cohorts. To improve the quality of RNA-seq experiments, it is crucial to incorporate an adequate number of biological replicates (Lamarre et al., 2018); evaluate factors such as effect size, maximum sample size, false-positive and false-negative results, and within-group variation (Busby et al., 2013; Wu et al., 2015); and consider sequencing read depth and single- or paired-end sequencing reads (Stark et al., 2019). Future research should focus on tissue composition, cell type expression, and multiomic approaches to better understand CeD pathogenesis and identify new therapeutic targets. The integration of spatial transcriptomics data with scRNA-seq data could reveal intricate immune cell networks that interact with epithelial and stromal cells in CeD pathology.

Ethics approval and consent to participate

Not applicable.

Funding

Not applicable.

Authors' contributions

MS conceptualized the work, drafted the manuscript, and substantively revised it; HS designed and visualized the figures, drafted the manuscript, and substantively revised it; MM conducted data collection, drafted the manuscript, and substantively revised it; MR supervised the study and substantively revised the manuscript. All the authors read and approved the final manuscript.

CRediT authorship contribution statement

Maryam Shoaran: Writing – review & editing, Supervision. **Hani Sabaie:** Writing – review & editing, Writing – original draft, Visualization. **Mehrnaz Mostafavi:** Writing – review & editing, Writing – original draft, Investigation. **Maryam Rezazadeh:** Writing – review & editing, Supervision.

Consent for publication

Not applicable.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Quillbot in order to improve the clarity, grammar, and readability of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

We would like to thank the Clinical Research Development Unit of Zahra Mardani Azari Children's Hospital, Tabriz University of Medical Sciences, for their support.

References

Altenbach, S.B., Vensel, W.H., DuPont, F.M., 2010. Integration of transcriptomic and proteomic data from a single wheat cultivar provides new tools for understanding the roles of individual alpha gliadin proteins in flour quality and celiac disease. J. Cereal Sci. 52 (2), 143–151.
Al-Toma, A., et al., 2019. European Society for the Study of Coeliac Disease (ESsCD) guideline for coeliac disease and other gluten-related disorders. United Eur. Gastroenterol. J. 7 (5), 583–613.
Anders, S., Pyl, P.T., Huber, W., 2014. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31 (2), 166–169.
Andrews, S., 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. United Kingdom, Cambridge.
Asp, M., et al., 2017. Spatial detection of fetal marker genes expressed at low level in adult human heart tissue. Sci. Rep. 7 (1), 12941.
Atlasy, N., et al., 2022. Single cell transcriptomic analysis of the immune cell compartment in the human small intestine and in Celiac disease. Nat. Commun. 13 (1).
Aune, T.M., et al., 2017. Expression of long non-coding RNAs in autoimmunity and linkage to enhancer function and autoimmune disease risk genetic variants. J. Autoimmun. 81, 99–109.
Banerjee, P., et al., 2023. A duodenal mucosa transcriptome study identified reduced expression of a novel gene CDH18 in celiac disease. Dig. Liver Dis.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30 (15), 2114–2120.
Bragde, H., et al., 2018. Celiac disease biomarkers identified by transcriptome analysis of small intestinal biopsies. Cell. Mol. Life Sci. 75 (23), 4385–4401.
Bragde, H.G., et al., 2020. Characterisation of gene and pathway expression in stabilised blood from children with coeliac disease. BMJ Open Gastroenterol. 7 (1).
Bray, N.L., et al., 2016. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34 (5), 525–527.
Buenrostro, J.D., et al., 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods 10 (12), 1213–1218.
Busby, M.A., et al., 2013. Scotty: a web tool for designing RNA-Seq experiments to measure differential gene expression. Bioinformatics 29 (5), 656–657.
Butler, A., et al., 2018. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. 36 (5), 411–420.
Byrne, A., et al., 2017. Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. Nat. Commun. 8 (1), 16027.
Cartolano, M., et al., 2016. cDNA library enrichment of full length transcripts for SMRT long read sequencing. PLoS One 11 (6), e0157779.
Castellanos-Rubio, A., et al., 2016. A long noncoding RNA associated with susceptibility to celiac disease. Science 352 (6281), 91–95.
Castellanos-Rubio, A., et al., 2017. Cytoplasmic form of carlr lncRNA facilitates inflammatory gene expression upon NF-kB activation. J. Immunol. 199 (2), 581–588.
Cellier, C., et al., 1998. Abnormal intestinal intraepithelial lymphocytes in refractory sprue. Gastroenterology 114 (3), 471–481.
Chama, M., et al., 2019. Transcriptomic analysis of enteropathy in Zambian children with severe acute malnutrition. EBioMedicine 45, 456–463.
Chen, J., et al., 2017. Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. Nat. Protoc. 12 (3), 566–580.
Chen, S., et al., 2023. Transcriptome profile and immune infiltrated landscape revealed a novel role of γδT cells in mediating pyroptosis in celiac disease. J. Transl. Med. 21 (1).
Chimento, A., et al., 2019. Progress to improve oral bioavailability and beneficial effects of resveratrol. Int. J. Mol. Sci. 20 (6).
Christophersen, A., et al., 2019. Distinct phenotype of CD4+ T cells driving celiac disease identified in multiple autoimmune conditions. Nat. Med. 25 (5), 734–737.
Christophersen, A., et al., 2021. Pathogenic T cells in celiac disease change phenotype on gluten challenge: implications for T-cell-directed therapies. Adv. Sci. (Weinheim, Baden-Wurttemberg, Germany) 8, e2102778.
Ciszewski, C., et al., 2020. Identification of a γc receptor antagonist that prevents reprogramming of human tissue-resident cytotoxic T cells by IL15 and IL21. Gastroenterology 158 (3), 625–637.e13.
Cloonan, N., et al., 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat. Methods 5 (7), 613–619.
Conesa, A., et al., 2016. A survey of best practices for RNA-seq data analysis. Genome Biol. 17 (1), 1–19.
Cording, S., et al., 2022. Oncogenetic landscape of lymphomagenesis in coeliac disease. Gut 71 (3), 497–508.
Crosetto, N., Bienko, M., Van Oudenaarden, A., 2015. Spatially resolved transcriptomics and beyond. Nat. Rev. Genet. 16 (1), 57–66.
de Almeida, R.C., et al., 2018. Integrative analysis identifies genetic variants associated with autoimmune diseases affecting putative microRNA binding sites. Front. Genet. 9 (Apr).
De Leo, L., et al., 2023. Understanding celiac disease at single cell resolution. J. Pediatr. Gastroenterol. Nutr. 76, 1.
Dieckman, T., et al., 2022. Single-cell analysis of refractory celiac disease demonstrates inter- and intra-patient aberrant cell heterogeneity. Cell. Mol. Gastroenterol. Hepatol. 14 (1), 173–192.
Djebali, S., et al., 2012. Landscape of transcription in human cells. Nature 489 (7414), 101–108.
Dobin, A., et al., 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29 (1), 15–21.
Dotsenko, V., et al., 2021. Genome-wide transcriptomic analysis of intestinal mucosa in celiac disease patients on a gluten-free diet and postgluten challenge. Cell. Mol. Gastroenterol. Hepatol. 11 (1), 13–32.
Dotsenko, V., et al., 2022. Transglutaminase 2 (TG2) inhibitor protects gluten-induced intestinal damage in celiac disease: transcriptomic analysis of a randomized gluten challenge study. Gastroenterology 162 (7), S-205.
Du Pré, M.F., et al., 2020. B cell tolerance and antibody production to the celiac disease autoantigen transglutaminase 2. J. Exp. Med. 217 (2).
Emrich, S.J., et al., 2007. Gene discovery and annotation using LCM-454 transcriptome sequencing. Genome Res. 17 (1), 69–73.
Eng, C.-H.-L., et al., 2019. Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. Nature 568 (7751), 235–239.
Engström, P.G., et al., 2013. Systematic evaluation of spliced alignment programs for RNA-seq data. Nat. Methods 10 (12), 1185–1191.
Felli, C., et al., 2022. Circulating microRNAs as novel non-invasive biomarkers of paediatric celiac disease and adherence to gluten-free diet. EBioMedicine, 76.
Ferguson, A., Arranz, E., O'mahony, S., 1993. Clinical and pathological spectrum of coeliac disease–active, silent, latent, potential. Gut 34 (2), 150.
Ferguson, A., Murray, D., 1971. Quantitation of intraepithelial lymphocytes in human jejunum. Gut 12 (12), 988–994.
Francavilla, A., et al., 2023. Gluten-free diet affects fecal small non-coding RNA profiles and microbiome composition in celiac disease supporting a host-gut microbiota crosstalk. Gut Microbes 15 (1).
Frazee, A.C., et al., 2015. Ballgown bridges the gap between transcriptome assembly and expression analysis. Nat. Biotechnol. 33 (3), 243–246.
Freire, R., et al., 2019. Human gut derived-organoids provide model to study gluten response and effects of microbiota-derived molecules in celiac disease. Sci. Rep. 9 (1), 7029.
Freitag, T.L., et al., 2020. Gliadin nanoparticles induce immune tolerance to gliadin in mouse models of celiac disease. Gastroenterology 158 (6), 1667–1681.e12.
Garalde, D.R., et al., 2018. Highly parallel direct RNA sequencing on an array of nanopores. Nat. Methods 15 (3), 201–206.
Gierahn, T.M., et al., 2017. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat. Methods 14 (4), 395–398.

Green, P.H., 2005. The many faces of celiac disease: clinical presentation of celiac disease in the adult population. Gastroenterology 128 (4), S74–S78.
Haberman, Y., et al., 2021. Mucosal genomics implicate lymphocyte activation and lipid metabolism in refractory environmental enteric dysfunction. Gastroenterology 160 (6), 2055–2071.e0.
Hao, Z., et al., 2019. Screening for differential expression of genes for resistance to Sitodiplosis mosellana in bread wheat via BSR-seq analysis. Theor. Appl. Genet. 132, 3201–3221.
Hashimshony, T., et al., 2016. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol. 17, 1–7.
He, S. et al., 2022. High-plex Multiomic Analysis in FFPE at Subcellular Level by Spatial Molecular Imaging. bioRxiv, p. 2021.11.03.467020.
Hoilat, G.J., et al., 2022. Larazotide acetate for treatment of celiac disease: a systematic review and meta-analysis of randomized controlled trials. Clin. Res. Hepatol. Gastroenterol. 46 (1), 101782.
Hong, M., et al., 2020. RNA sequencing: new technologies and applications in cancer research. J. Hematol. Oncol. 13 (1), 166.
Huber, W., et al., 2015. Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods 12 (2), 115–121.
Jabri, B., Sollid, L.M., 2017. T cells in celiac disease. J. Immunol. 198 (8), 3005–3014.
Jain, M., et al., 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 17, 1–11.
Jain, M., et al., 2018. Nanopore sequencing and assembly of a human genome with ultra-long reads. Nat. Biotechnol. 36 (4), 338–345.
Jin, Y., et al., 2020. Exploring the mechanisms of larazotide in the regulation of the tight junction barrier. Gastroenterology 158 (6), S-274–S-275.
Kaarbø, M., et al., 2023. Duodenal inflammation in common variable immunodeficiency has altered transcriptional response to viruses. J. Allergy Clin. Immunol. 151 (3), 767–777.
Kim, D., et al., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, 1–13.
Kim, D., et al., 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37 (8), 907–915.
Klein, A.M., et al., 2015. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161 (5), 1187–1201.
Kleshchevnikov, V., et al., 2022. Cell 2location maps fine-grained cell types in spatial transcriptomics. Nat. Biotechnol. 40 (5), 661–671.
Korn, T., et al., 2009. IL-17 and Th17 Cells. Annu. Rev. Immunol. 27, 485–517.
Kovaka, S., et al., 2019. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol. 20, 1–13.
Kumar, V., et al., 2015. Systematic annotation of celiac disease loci refines pathological pathways and suggests a genetic explanation for increased interferon-gamma levels. Hum. Mol. Genet. 24 (2), 397–409.
Lamarre, S., et al., 2018. Optimization of an RNA-Seq differential gene expression analysis depending on biological replicate number and library size. Front. Plant Sci. 9, 108.
Lein, E., Borm, L.E., Linnarsson, S., 2017. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. Science 358 (6359), 64–69.
Leinonen, R., et al., 2010. The sequence read archive. Nucleic Acids Res. 39 (suppl_1), D19–D21.
Leonard, M.M., et al., 2019. RNA sequencing of intestinal mucosa reveals novel pathways functionally linked to celiac disease pathogenesis. PLoS ONE 14 (4).
Li, J., et al., 2022. KIR+CD8+ T cells suppress pathogenic T cells and ar active in autoimmune diseases and COVID-19. Science 376 (6590).
Liao, Y., Smyth, G.K., Shi, W., 2013. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30 (7), 923–930.
Lin, M., et al., 2021. BSR-Seq analysis provides insights into the cold stress response of Actinidia arguta F1 populations. BMC Genomics 22 (1), 72.
Lindeman, I., et al., 2021. Longevity, clonal relationship, and transcriptional program of celiac disease–specific plasma cells. J. Exp. Med. 218 (2).
Lister, R., et al., 2008. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133 (3), 523–536.
Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 1–21.
Luca, S.V., et al., 2020. Bioactivity of dietary polyphenols: the role of metabolites. Crit. Rev. Food Sci. Nutr. 60 (4), 626–659.
Luecken, M.D., Theis, F.J., 2019. Current best practices in single-cell RNA-seq analysis: a tutorial. Mol. Syst. Biol. 15 (6), e8746.
Macosko, E.Z., et al., 2015. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161 (5), 1202–1214.
Margulies, M., et al., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437 (7057), 376–380.
Marín-Sanz, M., Barro, F., 2022. RNAi silencing of wheat gliadins alters the network of transcription factors that regulate the synthesis of seed storage proteins toward maintaining grain protein levels. Front. Plant Sci. 13, 935851.
Marsh, M.N., 1988. Studies of intestinal lymphoid tissue. XI–The immunopathology of cell-mediated reactions in gluten sensitivity and other enteropathies. Scanning Microsc. 2 (3), 41.
McCarthy, D.J., et al., 2017. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33 (8), 1179–1186.
McKenna, A., et al., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20 (9), 1297–1303.
Merritt, C.R., et al., 2020. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Nat. Biotechnol. 38 (5), 586–599.
Mifsud, B., et al., 2015. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. Nat. Genet. 47 (6), 598–606.
Moehs, C.P., et al., 2019. Development of decreased-gluten wheat enabled by determination of the genetic basis of lys3a barley. Plant Physiol. 179 (4), 1692–1703.
Molinie, B., et al., 2014. Multiple autoimmune susceptibility genes are γ-interferon responsive in human monocytes. Gastroenterology 146 (5), S-902.
Montoro, D.T., et al., 2018. A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. Nature 560 (7718), 319–324.
Moor, A.E., Itzkovitz, S., 2017. Spatial transcriptomics: paving the way for tissue-level systems biology. Curr. Opin. Biotechnol. 46, 126–133.
Morris, K.V., Mattick, J.S., 2014. The rise of regulatory RNA. Nat. Rev. Genet. 15 (6), 423–437.
Nagafuchi, Y., Yanaoka, H., Fujio, K., 2022. Lessons from transcriptome analysis of autoimmune diseases. Front. Immunol. 13, 857269.
Nashiry, M.A., et al., 2021. Bioinformatics and system biology approaches to identify the diseasome and comorbidities complexities of SARS-CoV-2 infection with the digestive tract disorders. Brief. Bioinform. 22 (6).
Oikonomopoulos, S., et al., 2016. Benchmarking of the Oxford Nanopore MinION sequencing for quantitative and qualitative assessment of cDNA populations. Sci. Rep. 6 (1), 31602.
Patrick, M.F., et al., 2021. The phenotype and TCR repertoire of intestinal CD8+ T cells is altered in coeliac disease. Gut 70 (SUPPL 1), A24.
Patro, R., et al., 2017. Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods 14 (4), 417–419.
Picelli, S., et al., 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods 10 (11), 1096–1098.
Pimentel, H., et al., 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods 14 (7), 687–690.
Quinn, E.M., et al., 2015. Transcriptome analysis of CD4+ T cells in coeliac disease reveals imprint of BACH2 and IFNI regulation. PLoS ONE 10 (10).
Ramírez-Sánchez, A.D., et al., 2022. Single-cell RNA sequencing of peripheral blood mononuclear cells from pediatric coeliac disease patients suggests potential pre-seroconversion markers. Front. Immunol. 13.
Ricaño-Ponce, I., et al., 2016. Refined mapping of autoimmune disease associated genetic variants with gene expression suggests an important role for non-coding RNAs. J. Autoimmun. 68, 62–74.
Rizzo, F., et al., 2020. Small-bowel carcinomas associated with celiac disease: transcriptomic profiling shows predominance of microsatellite instability-immune and mesenchymal subtypes. Virchows Arch. 476 (5), 711–723.
Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26 (1), 139–140.
Rodriques, S.G., et al., 2019. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science 363 (6434), 1463–1467.
Saeidian, A.H., et al., 2020. Research techniques made simple: whole-transcriptome sequencing by RNA-Seq for diagnosis of monogenic disorders. J. Invest. Dermatol. 140 (6), 1117–1126.e1.
Sakula, J., Shiner, M., 1957. Coeliac disease with atrophy of the small-intestine mucosa. Lancet 273, 876–877.
Salentijn, E.M.J., et al., 2013. Quantitative and qualitative differences in celiac disease epitopes among durum wheat varieties identified through deep RNA-amplicon sequencing. BMC Genomics 14 (1).
Sharon, D., et al., 2013. A single-molecule long-read survey of the human transcriptome. Nat. Biotechnol. 31 (11), 1009–1014.
Shewry, P., 2019. What is gluten—why is it special? Front. Nutr. 101.
Shoda, T., et al., 2023. Molecular analysis of duodenal eosinophilia. J. Allergy Clin. Immunol. 151 (4), 1027–1039.
Simonis, M., Kooren, J., De Laat, W., 2007. An evaluation of 3C-based methods to capture DNA interactions. Nat. Methods 4 (11), 895–901.
Slifer, Z.M., et al., 2021. Larazotide acetate: a pharmacological peptide approach to tight junction regulation. Am. J. Physiol. Gastrointest. Liver Physiol. 320 (6), G983–G989.
Smith, A.M., et al., 2019. Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing. PLoS One 14 (5), e0216709.
Snir, O., et al., 2019. Transcriptional profiling of human intestinal plasma cells reveals effector functions beyond antibody production. United Eur. Gastroenterol. J. 7 (10), 1399–1407.
Sollid, L.M., Jabri, B., 2005. Is celiac disease an autoimmune disorder? Curr. Opin. Immunol. 17 (6), 595–600.
Sollid, L.M., Jabri, B., 2013. Triggers and drivers of autoimmunity: lessons from coeliac disease. Nat. Rev. Immunol. 13 (4), 294–302.
Sosnovski, K.E., et al., 2023. GATA6-AS1 regulates intestinal epithelial mitochondrial functions, and its reduced expression is linked to intestinal inflammation and less favourable disease course in ulcerative colitis. J. Crohns Colitis 17 (6), 960–971.
Sowińska, A., et al., 2020. Transcriptional and ultrastructural analyses suggest novel insights into epithelial barrier impairment in celiac disease. Cells 9 (2).
Ståhl, P.L., et al., 2016. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science 353 (6294), 78–82.
Stark, R., Grzelak, M., Hadfield, J., 2019. RNA sequencing: the teenage years. Nat. Rev. Genet. 20 (11), 631–656.
Stegle, O., Teichmann, S.A., Marioni, J.C., 2015. Computational and analytical challenges in single-cell transcriptomics. Nat. Rev. Genet. 16 (3), 133–145.
Svensson, V., Vento-Tormo, R., Teichmann, S.A., 2018. Exponential scaling of single-cell RNA-seq in the past decade. Nat. Protoc. 13 (4), 599–604.
Tan, I.L., et al., 2021a. Circulating miRNAs as potential biomarkers for celiac disease development. Front. Immunol. 12.

Tan, I.L., et al., 2021b. A combined mRNA- and miRNA-sequencing approach reveals miRNAs as potential regulators of the small intestinal transcriptome in celiac disease. Int. J. Mol. Sci. 22 (21).
Tang, F., et al., 2009. mRNA-Seq whole-transcriptome analysis of a single cell. Nat. Methods 6 (5), 377–382.
Tuomela, S., et al., 2016. Comparative analysis of human and mouse transcriptomes of Th17 cell priming. Oncotarget 7 (12), 13416–13428.
van der Graaf, A., et al., 2020. Systematic prioritization of candidate genes in disease loci identifies TRAFD1 as a master regulator of IFNγ signaling in celiac disease. Front. Genet. 11.
Vickovic, S., et al., 2019. High-definition spatial transcriptomics for in situ tissue profiling. Nat. Methods 16 (10), 987–990.
Wang, E.T., et al., 2008. Alternative isoform regulation in human tissue transcriptomes. Nature 456 (7221), 470–476.
Wang, D.W., et al., 2017a. Genome-wide analysis of complex wheat gliadins, the dominant carriers of celiac disease epitopes. Sci. Rep. 7, 44609.
Wang, Y., et al., 2017b. Mapping stripe rust resistance gene YrZH22 in Chinese wheat cultivar Zhoumai 22 by bulked segregant RNA-Seq (BSR-Seq) and comparative genomics analyses. Theor. Appl. Genet. 130, 2191–2201.
Weirather, J.L., et al., 2017. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. F1000Res 6, 100.
Withoff, S., et al., 2016. Understanding celiac disease by genomics. Trends Genet. 32 (5), 295–308.
Workman, R.E., et al., 2019. Nanopore native RNA sequencing of a human poly(A) transcriptome. Nat. Methods 16 (12), 1297–1305.
Wu, H., Wang, C., Wu, Z., 2015. PROPER, comprehensive power evaluation for differential expression using RNA-seq. Bioinformatics 31 (2), 233–241.
Xin, Y., et al., 2016. Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. Proc. Natl. Acad. Sci. 113 (12), 3293–3298.
Yao, Y., et al., 2021. Differential expression profile of glutenspecific T cells identified by single-cell RNAseq. PLoS ONE 16 (10 October 2021).
Yohannes, D.A., et al., 2020. Effects of in vivo gluten challenge on PBMC gene expression profiles in diet treated celiac disease. Front. Immunol. 11.
Yu, T., et al., 2022. The nutritional intervention of resveratrol can effectively alleviate the intestinal inflammation associated with celiac disease induced by wheat gluten. Front. Immunol. 13.
Zhang, Y., et al., 2021. Single-cell RNA sequencing in cancer research. J. Exp. Clin. Cancer Res. 40 (1), 81.
Zhang, W., Ciclitira, P., Messing, J., 2014. PacBio sequencing of gene families — a case study with wheat gluten genes. Gene 533 (2), 541–546.
Zheng, G.X., et al., 2017. Massively parallel digital transcriptional profiling of single cells. Nat. Commun. 8 (1), 14049.