



Review

AI-Based Prediction of Gene Expression in Single-Cell and Multiscale Genomics and Transcriptomics

Ema Andreea Pălăștea ^{1,2}, Irina-Mihaela Matache ^{2,†} , Eugen Radu ^{2,3} , Octavian Henegariu ^{4,*} and Octavian Bucur ^{1,2,5,*}

¹ Genomics Research and Development Institute, Bucharest 030167, Romania; ema-andreea.palastea2022@stud.umfcd.ro

² Faculty of Medicine, Carol Davila University of Medicine and Pharmacy, Bucharest 030167, Romania; irina.m.matache@gmail.com (I.-M.M.); eugen.radu@umfcd.ro (E.R.)

³ Emergency University Hospital, Bucharest 050098, Romania

⁴ Department of Neurosurgery, Yale School of Medicine, New Haven, CT 06520, USA

⁵ Viron Molecular Medicine Institute, Boston, MA 02108, USA

* Correspondence: octavian.henegariu@yale.edu (O.H.); octavian.bucur@genomica.gov.ro (O.B.)

† Current address: Department of Neurological Surgery, University of Chicago, Chicago, IL 60637, USA.

Abstract

Omics research is changing the way medicine develops new strategies for diagnosis, prevention, and treatment. With the surge of advanced machine learning models tailored for omics analysis, recent research has shown improved results and pushed the progress towards personalized medicine. The dissection of multiple layers of genetic information has provided new insights into precision medicine, at the same time raising issues related to data abundance. Studies focusing on single-cell scale have upgraded the knowledge about gene expression, revealing the heterogeneity that governs the functioning of multicellular organisms. The amount of information gathered through such sequencing techniques often exceeds the human capacity for analysis. Understanding the underlying network of gene expression regulation requires advanced computational tools that can deal with the complex analytical data provided. The recent emergence of artificial intelligence-based frameworks, together with advances in quantum algorithms, has the potential to enhance multi-omics analyses, increasing the efficiency and reliability of the gene expression profile prediction. The development of more accurate computational models will significantly reduce the error rates in interpreting large datasets. By making analytical workflows faster and more precise, these innovations make it easier to integrate and interrogate multi-omics data at scale. Deep learning (DL) networks perform well in terms of recognizing complex patterns and modeling non-linear relationships that enable the inference of gene expression profiles. Applications range from direct prediction of DNA sequence-informed predictive modeling to transcriptomic and epigenetic analysis. Quantum computing, particularly through quantum machine learning methods, is being explored as a complementary approach for predictive modeling, with potential applications to complex gene interactions in increasingly large and high-dimensional biological datasets. Together, these tools are reshaping the study of complex biological data, while ongoing innovation in this field is driving progress towards personalized medicine. Overall, the combination of high-resolution omics and advanced computational tools marks an important shift toward more precise and data-driven clinical decision-making.

Keywords: machine learning; AI-based prediction; single-cell omics; quantum computing; deep learning; spatial transcriptomics; gene expression



Academic Editors: Francesca Oliviero and Anna Lisa Giuliani

Received: 29 November 2025

Revised: 30 December 2025

Accepted: 8 January 2026

Published: 13 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The process of gene expression is determined by the dynamic relationship between the cell and various intrinsic and extrinsic factors. The emergence of new technologies might help with deciphering the underlying mechanisms that direct a cell towards changing its state [1]. Understanding this process requires both high-resolution experimental techniques and computational tools capable of capturing the complexity of cellular diversity. Single-cell sequencing plays an important role in the research of gene expression [2], as it enables the analysis of genomic, epigenomic, and transcriptomic heterogeneity in cellular populations and the tracking of changes over time. Single-cell genome sequencing reveals genetic variability, compared to bulk sequencing, which only offers an average gene expression profile in populations of cells [3]. Bulk sequencing remains a cost-effective alternative, but it masks particular and rare patterns. The function of each cell is dictated by a regulatory network which comprises transcription factors (TFs) and regulatory sequences, which makes the approach of single-cell multi-omics critical for unveiling the biological contribution of every cell. The analysis of the transcriptome, the proteome, and the genome is employed in identifying the potential trajectory of a cell. Single-cell studies can be performed either one at a time or in a massive parallel fashion. Nonetheless, they have improved the way in which we understand the biological networks created in heterogeneous tissues [4]. Building on this expanding resolution of cellular states, it becomes important to clarify how these technologies connect the gap between genetic variation and functional consequences.

While genome-wide association studies (GWASs) managed to identify thousands of variants associated with the risk of genetically inherited traits and diseases, it did not reveal the mechanisms underlying these links [5]. Consequently, integrating transcriptomic and epigenomic modalities has emerged as a powerful strategy to explain how genetic variants exert regulatory effects.

Transcriptomics profiling can be combined with multimodal epigenetic analysis techniques such as single-cell chromatin overall omics-scale landscape sequencing—scCOOL-seq [4]; nucleosome occupancy and methylome sequencing (cNOME-seq) [6]; single-cell chromatin accessibility and transcriptome sequencing (scCAT-seq) [7]; and combined assay for transposase-accessible chromatin using sequencing and RNA (scATAC-RNA-seq) [8]. These methods, crossing multiple epigenetic layers, facilitate the mapping of accessible chromatin regions in single cells, another powerful approach for dissecting tissue heterogeneity [4].

Although scRNA-seq [9] (single-cell transcriptomics) and other single-cell techniques have become indispensable tools in multiple research areas, they have limitations (see Benefits and Limitations). One would be the risk of being presented only with a partial image of the complexity of the gene regulation process [4]. Nevertheless, its limitations arise from the lack of spatial context. To combat these issues, spatial transcriptomics (ST) technologies were introduced. This concept relies on retaining spatial information that highly contributes to gene expression analysis and prediction [9].

Dynamic cellular processes do not depend on the expression of a single gene but are determined by the temporal expression and function of multiple genes. This hints at the importance of the spatial interaction between different sites of the genome, the interaction with other cells, and with the surrounding environment. Even though single-cell omics and spatial genomics give us an unprecedented view on gene programs and cell states, these methods require the alteration or lysis of the cells (to obtain single cells) and prevent us from directly observing the dynamics of molecular profiles in living cells. Certain pseudo-time algorithms, such as Monocle2 [10] or Waddington-OT [11], can deduce dynamic processes from static snapshots of molecular profiles. Pseudo-time inferring represents the progression of single cells along a biological process, based on gene expression profiles,

in the conditions of lacking explicit time points [12]. These methods rely on assumptions about RNA kinetics that might be violated when the sampling timescale does not align with the pace of biological changes. Technologies like Live-seq [13,14] provide real-time sampling, but with a low throughput.

The complex datasets have catalyzed the adaptation of artificial intelligence (AI) for suiting data interpretation [15]. A subset of AI, deep learning (DL), leverages artificial neural networks to extract patterns from data, their architecture being inspired by the biological model of animal brains. Each layer serves as a non-linear transformation function of its input, processing the result from the previous layer as its input and generating progressively complex representations as the number of layers increases. The training of DL models is performed by them fitting themselves on the training data. DL frameworks perform tasks with a deep and hierarchical architecture, making them suitable for biological applications. These features encouraged scientists to build DL models performing various predictive tasks, ranging from predictions regarding crop yields [16] to frameworks that estimate mortality rates [17] or assess risk stratification among patients with a certain condition, such as COVID-19 [18]. The similarity between biological sequences and natural languages enables DL architectures to process omics datasets [19], thus increasing their applications designed to predict regulatory interactions, model transcriptional regulation, or identify non-coding variant effects.

The training of neural networks to recognize and predict gene expression patterns represent an important upgrade towards precision and preventive medicine. From a clinical point of view, such models could perform early diagnosis, patient clustering, and treatment design [20]. Minor genetic variability can be incriminated for decreased response to treatment and AI can help identify such variants and find an alternative that suits the patient's phenotype.

Nonetheless, machine learning (ML) and deep learning (DL) models encounter significant challenges in predicting mammalian transcriptional regulation due to the complex sequence logic and the need to analyze interactions over large genomic sequences [21].

The complexity of this type of analysis originates from the multiple variables that must be taken into account when inferring different predictions. mRNA vs. protein abundance relationship studies reveal interesting correlations [18,22,23], while mapping quantitative trait loci [24] allowed the switch from bulk to single-cell focus. The AI revolution brought interest in developing models that work with existing datasets and improve the quality of the gene expression profiles; DL platforms were created for understanding transcriptional regulation, the impact of noncoding variants, and many other directions of investigation [25]. As an example, DeepBind [26] is able to predict transcription factor-binding sites (TFBSs).

This review aims to summarize the recent advances and applications of AI algorithms in the study of gene expression, especially discussing the ability to predict expression patterns, highlighting the importance of single-cell omics in the development of this research area. This paper also tackles the topic of promising clinical applications that can result from gathering data through these innovative technologies; a brief description of the possible improvement that can be brought by quantum computing in this field is included as well.

2. Single-Cell Sequencing Principles and Protocols

Single-cell omics upgraded the resolution at which scientists look at tissue heterogeneity [27]. Compared to bulk sequencing, despite being more statistically sensitive and harder to analyze, it also reveals a more realistic and detailed version of tissue complexity [28]. There is a continuous quest for computational methods to link cell heterogeneity at gene expression level to phenotype [27]. Single-cell sequencing brought new

perspectives not only in the diagnostic and identification of possible gene variants but also suggested new applications in drug discovery and development [29]. To understand how these advantages are realized in practice, it is necessary to examine the experimental workflows and technological variants that underlie single-cell assays.

Since the introduction of single-cell technology, various approaches have been applied in order to capture cells and to amplify RNA, each having advantages and disadvantages [30]. Provided that a single-cell technique by itself might only offer an incomplete frame of the complexity of the gene regulation process, the introduction of the multi-omics approaches certainly enhances the quality of the determinations [4].

While single-cell sequencing methods can be applied at multiple levels and spots, one of the most used technologies is the single-cell RNA sequencing (scRNA-seq), which captures the transcriptome for assessing gene expression levels.

The main steps in scRNA-seq involve single-cell isolation and capture, cell lysis, reverse transcription, cDNA amplification, and library preparation. Isolation and capture steps aim to yield high-quality individual cells from tissues, forming a basis for precise genetic and molecular analyses. ScRNA-seq preserves the individuality of each cell, empowering the study of heterogeneous populations such as tumors, unlike bulk sequencing, which offers an average signal from a large number of cells.

There are different methods of isolating and capturing cells, depending on the organisms, cell properties, or tissues [31]. Common strategies include isolating whole cells, nuclei, organelles, or using specific marker proteins to separate particular cell types. The most widely used methods are fluorescence-activated cell sorting (FACS), magnetic-activated cell-sorting (MACS), microfluidic systems, and laser microdissection. Each single cell is finally captured in an isolated reaction chamber where all transcripts from this single cell are uniquely barcoded after being converted into cDNA molecules.

Gradually, it has been shown that the dissociation process might induce the expression of stress genes, revealing an artificial alteration of the transcriptional process. Experiments demonstrated that protease dissociation at 37 °C triggers the expression of stress genes [30], whereas performing dissociation at 4 °C minimizes the artifact. As an alternative to the scRNA-seq for single-cell sequencing, single-nucleus RNA sequencing (snRNA-seq) can be used. This method avoids the need for intact cells, as it only targets the mRNAs from the cells' nuclei. It can be applied to frozen or otherwise difficult tissue samples, besides minimalizing the artificial transcriptional stress response [32]. snRNA-seq reduces the artifacts induced by dissociation, has broad applicability, and the isolation process is easier, but it lacks information on mRNA processing and RNA stability.

Following RNA extraction and reverse transcription, cDNA amplification is performed using either polymerase chain reaction (PCR) [33–39] or in vitro transcription (IVT) [40–43]. Both methods can lead to amplification biases; therefore, the introduction of unique molecular identifiers (UMIs) that are used to barcode each mRNA molecule during reverse transcription marked a major improvement [44]. The listed sequencing protocols use adapted UMIs [37,38,45].

The next step involves sequencing barcoded cDNA libraries using high throughput platforms. In DNA nanoballs, DNA fragments are processed to generate a 3' adenine overhang. DNA fragments are ligated with dTTP-tailed adapters, circularized and amplified to generate dense libraries. The final product of this protocol is a single-strand circular DNA library [1].

The isolation and capture strategies divide scRNA-seq techniques into two main approaches, the plate-/microfluidic-based methods and droplet-based methods. The plate- and microfluidic-based methods are similar and often limited to ~50 to ~500 cells per analysis. Plate-based systems use fluorescence-activated cell sorting, while microfluidic-

based ones are automated platforms, such as Fluidigm C1, which use parallel microfluidic channels. Fluidigm C1 uses valve-based isolation, which controls the fluidic circuit; this valve system is placed under the connection site of the nanoliter chambers [46]. These methods generally present high sensitivity, quantifying up to ~10,000 genes per cell. Meanwhile, droplet-based methods barcode single cells and use UMIs to tag each transcript in individual oil droplets [2]. This increases the throughput to up to 10,000 cells per run. However, it only detects 1000–3000 genes per cell, due to technical limitations. To put it briefly, the first method performs a more sensitive gene quantification but can cope with a smaller number of cells in an analysis, while the latter extracts less genetic information from each cell, but provides a transcriptional landscape from a larger population of cells. Droplet microfluidic protocols have also been developed, such as spinDrop [47], which mixes fluorescence-activated droplet sorting (FADS) for obtaining the probe with picoinjection. Regardless of the chosen platform, rigorous quality control is indispensable to ensure that downstream analyses reflect true biological signal rather than technical artifacts.

Choosing a scRNA-seq method requires considering a few factors, such as the expected number of cells, cell size, and acceptable technical biases. Quality control (QC) is crucial and is divided into cell QC and gene QC. For cell QC, barcodes that might not belong to intact individual cells should be eliminated. These altered cells are filtered out by analyzing UMI counts, the number of expressed genes, the total detected counts, and the proportion of mitochondrial RNA fractions. For high throughput scRNA-seq, doublets or multiplets must also be excluded—these are cells that present very high counts and a large number of expressed genes and occur in the case of two or more cells that are accidentally directioned in the same droplet or well [26,48]. Since none of the sequencing methods alone can tell the differences between doublets and true single cells, computational tools, such as scMODD (a model-driven doublet detector in scRNA-seq data) [49], scrublet [50], and scds [51], have been developed to infer and remove these artifacts [30].

3. Machine Learning and Deep Learning Models Used for Gene Expression Levels Prediction

3.1. Machine Learning and Deep Learning Architectures in Omics Research

AI systems are designed to perform tasks in the same way a human being would. These models imitate human cognition and are constructed in such a manner that they develop learning, reasoning, and perception like a human. This concept empowered the scientific world to develop AI-based systems that are more and more complex and that can analyze large amounts of data, recognize patterns, and make predictions based on the learnt and extracted information. Furthermore, ML and DL frameworks can learn from previous mistakes and adapt afterward even without being programmed to do so. Machine learning is a subset of AI that focuses on building algorithms that enable computers to learn from data and make predictions. DL is a subset of ML which uses artificial neural networks to learn from data. The similarity of artificial neural networks to the structure and function of the human brain underscores their utility in handling vast and complex datasets that exceed the processing capacity of humans. Another advantage is that DL models improve performance, as they are exposed to more data [20].

Given these foundational principles, the transition from general-purpose AI to domain-specific genomics frameworks becomes increasingly intuitive. The promising results of applying DL algorithms in various areas of study have led to the development of dedicated platforms using this technology in genomics, a field where researchers encounter highly complex, multi-dimensional data. There are numerous DL models used for genomic data analysis and others are emerging. Artificial neural networks resemble the biological model of the human cortex, both in terms of architectural design and function. DL models are

trained using a process named backpropagation based on mathematical calculations that adjust the parameters or weights relating to prediction errors [52]. DL is also divided into supervised, semi-supervised, and unsupervised learning. There are different applications of these types of learning. For example, deep feed-forward neural networks (DFNNs) aim at solving tasks such as classification or regression performed by mapping input data to a representation and use supervised learning. Semi-supervised learning is applied when there is only a small number of data points labeled which will be used to label the unlabeled data. Unsupervised learning is commonly involved in single-cell sequencing data analysis, reducing dimensionality and clustering cells, basically identifying patterns in data without supervision [52]. Genomic studies such as cell-type classification using single-cell RNA-seq data have widely used semi-supervised learning methods with deep neural networks (DNNs) [53].

As these learning paradigms diversified, specialized architectures emerged to address specific biological problems. There are different types of DNNs used in gene regulation studies, the most popular being multi-layer perceptrons (MLPs). MLPs receive input data and direct them through a variable number of hidden layers. Other major deep-learning architectures in gene regulation research include convolutional neural networks (CNNs), recurrent neural networks (RNNs), graph neural networks (GNNs), and transformers, each presenting certain properties that prove their utility in the domain [53]. RNNs such as gated recurrent units (GRUs) or long short-term memory (LSTM) can handle long-range interactions of sequential data. RNNs present hidden states within the network which “remember sequential information at earlier locations” [15,54], being the equivalent of the memory of the network [54]. Transformers can be used for self-supervised learning on biological sequences [19]. Equipped with these architectures, researchers began applying them directly to regulatory genomics, leading to the first generation of sequence-to-expression models.

The implementation of DL models to infer gene expression patterns promises to help with deciphering existent problems in genetics research. It might also lead the way to uncovering new approaches in terms of phenotype prediction. Many of these models use DNA sequences as input rather than observed genetic variants, learning from the sequences and then being able to recognize patterns and predict outcomes for arbitrary sequences. The principle behind the learning algorithm is recognizing motifs (short repetitive patterns of nucleotides) and the elaboration of predictions is based on combinations of such motifs. This is known as the “sequence-to-expression” model. Multiple DL models were developed using this principle, the most relevant examples being Enformer [55], ExPecto [56], Xpresso [57], and Basenji [58], all of them performing gene expression prediction directly from the sequence [55].

3.2. Deep Learning Frameworks for Sequence-to-Expression Prediction

The earliest DL frameworks used for gene expression analysis operated on relatively short input sequences, under 500 bp length. They were capable of capturing local features with spatially invariant patterns among these sequences, and the architecture that best suits this kind of approach is CNN. Tools like DeepBind [59] and DeepSEA [60] predict DNA- and RNA-binding specificities of proteins (see Figure 1). They were designed to present a convolutional layer, a rectification layer, and a pooling layer. These three layers are described as a CNN block, followed by a fully connected layer. Nowadays, longer genomic regions, up to 200 kb, long enough to contain most determinants of gene expression, can be subjected to DL modeling [61].

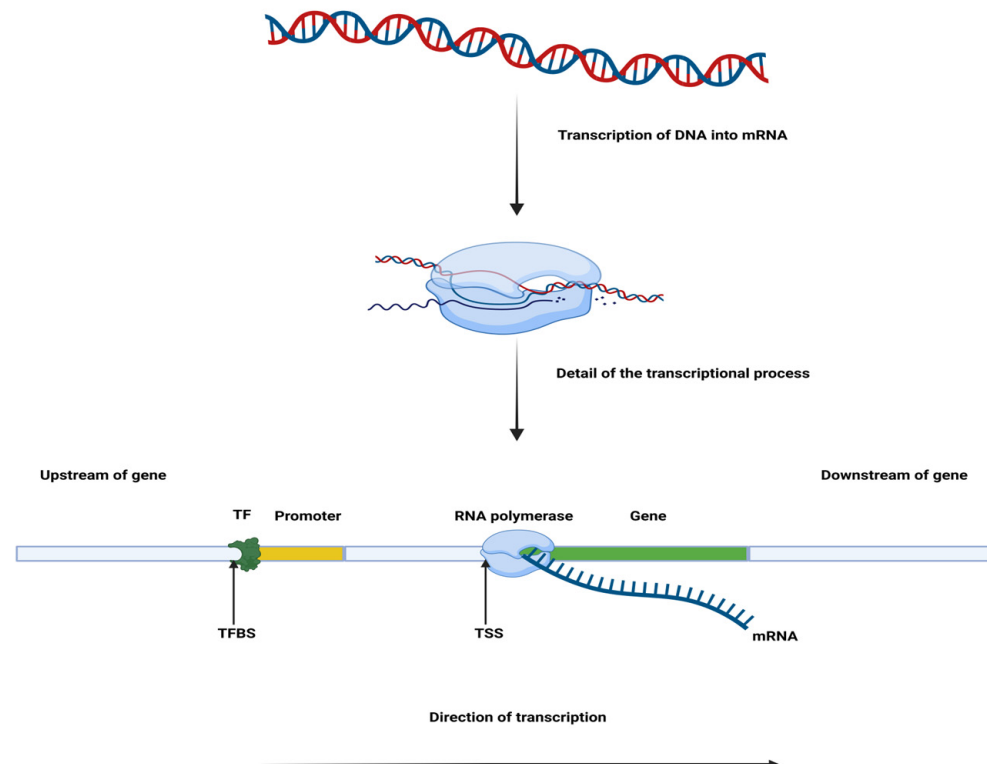


Figure 1. A simplified representation of transcription; the main elements that are used in gene expression prediction by the frameworks mentioned in the paper. TF, transcription factor; TFBS, transcription factor-binding site; and TSS, transcription start site. Created in BioRender. Palastea, E. <https://BioRender.com/dvtafpy> (accessed on 29 December 2025).

Many later deep-learning models adopted CNN architectures (Basset [62], DanQ [63] (CNN + BiLSTM), CpGenie [64], DeFine [65], Basenji, Expecto, Xpresso, Enformer (CNN + transformer), BPNet [19], and others). DanQ, for example, has effectively employed a hybrid architecture of CNN and bidirectional LSTM (BiLSTM), outperforming DeepSEA, despite being trained on the same dataset [63]. All of the previously mentioned methods regard the genome as linear sequences per chromosome (see Table 1).

As the field continued to evolve, researchers increasingly turned their attention towards modeling 3D genome architecture and extending predictions across species with frameworks such as Akita [66] and Orca [67,68]. Orca is a tool designed for predicting 3D genome architecture from sequence-dependent features of structures such as chromatin compartments and topologically associating domains (TADs). It operates on sequence data spanning from kilobases to entire chromosomes [67].

Dalla-Torre et al. introduced nucleotide transformer (NT) [69], a family of transformer-based foundation models which are pre-trained on DNA sequences from multiple genomes. It brought a different approach from the usual training of DL models, which was based on one reference genome. This initiative was conducted in an effort to overcome the poor performance of DL systems in predicting individualized gene expression patterns. The study demonstrated that NT outperforms other methods in predicting splice sites, histone modifications, and gene regulatory elements, and can even identify functionally important variants. For personalized gene expression prediction [55], paired DNA sequence and gene expression data were used as input to a specialized downstream transformer model. The reason behind training this single model by feeding it with the embeddings from a large set of genes is the theory that it might capture information that explains variation in gene expression across individuals and across genes. The performance was shown to be

improved as more attention layers were added to the transformer blocks that are integrated within the model's architecture [69].

Enformer [21], a pre-trained epigenomic model, is still considered the reference in terms of DL frameworks performing gene expression prediction. Its workflow is different from that of the NT, inferring separate predictions for each haplotype. In single-cell gene expression prediction analysis, the model captures cell-specific gene expression and enables variant effect prediction at the scale of single-cell and single-base pair. Thus, the heterogeneity of variant effect helps with the initiative of cell-type annotation and with deciphering genetic variation [21]. Enformer can capture variation across genes, but struggles with cross-individual variation, showing an average correlation of roughly zero for all tissues. Compared to traditional statistical models such as PrediXcan [55], which uses the regularized linear regression, Enformer often failed to offer a valid attribute to the variants' direction of effect but was more exact in predicting the magnitude of the effect [52,70]. Enformer is the first DL model to embrace a hybrid architecture for gene expression and epigenetic feature prediction, having been structured on both CNN and transformer artificial neural networks [19].

Table 1. The table summarizes short descriptions of the datasets used and the output of each deep learning model mentioned.

Deep Learning Model	Target Analysis	Prediction Output	Advantages	Disadvantages
DeepSEA [56,60]	Chromatin features such as TF binding and histone marks from large-scale chromatin profiling data	DNase hypersensitivity, TF-binding sites, and histone modifications	Simple CNN architecture, directly learning from sequence (automated learning) Scalability Predictive power Good performance on large datasets	Computational costs, expensive resources High-quality demands for training data Interpretability—black box limitations available for all DL models
DeepBind [26]	Protein binding to DNA/RNA (TFs, RBPs; works on both microarray and sequencing data)	Probability for TFs/RBPs of binding on sequences prediction of sequence specificity of protein binding	Simple CNN architecture with automatic feature learning Efficiency Performs well in predicting DNA/RNA protein-binding sites Efficiency in analyzing complex patterns	Costs and resources Black box configuration The simple CNN architecture has troubles capturing long-range genomic interactions
Basenji [58]	Epigenetic and transcriptional analysis	Quantitative prediction of gene expression profiles and epigenetic marks (CAGE, TFBS, and histone modifications)	CNN architecture, but capable of analyzing longer sequences, provides information on distant enhancers Noise reduction Multitasking	Struggles with very long-range interactions, lower resolution compared to transformer-based models Analyzes the regulatory processes, but does not interpret the effect of the mutations in coding regions

Table 1. Cont.

Deep Learning Model	Target Analysis	Prediction Output	Advantages	Disadvantages
Basset [62]	Chromatin accessibility, such as DNase hypersensitivity	Binary classification for chromatin accessibility in 164 cell types using sequencing data (DNase-seq)	A CNN capable of multitasking Prediction of variant effect Highly accurate predictions compared to traditional ML algorithms Automated feature extraction Pretraining accelerates the learning of new data	Designed only for processing short sequences It is more limited due to the binary prediction, characteristic of the original CNN architecture, compared to Basenji Data dependency
Enformer [55]	Epigenomic profiles and gene expression	Predicts transcriptional activity and chromatin features across large genomic regions	Transformer-based model capable of analyzing large genomic sequences Captures enhancer-promoter relationships Variant effect prediction Infers cross-gene correlation	Often fails to correctly attribute the direction of the effect of a mutation (failure to assess cross-individual correlation) High computational resources Trained on a fixed set of data, hard to generalize Limited resolution
DanQ [63]	TFBS prediction, mapping chromatin accessibility and histone marks	Predicts chromatin features and TFBS at base level	Hybrid architecture, both CNN and BiLSTM More precise than DeepSEA in predicting certain regulatory markers Predictions directly from DNA sequence Captures local and long-range dependencies	Computational expenses Needs training on large sets of labeled data to obtain accurate predictions Black box limitations
GraphReg [68]	Gene expression via 3D chromatin structure	Makes use of GATs to incorporate chromatin interaction graphs	GAT that uses 3D chromatin data to infer superior predictions on enhancer–promoter interactions Captures massive long-range interactions Higher accuracy in the prediction of TFBS	Data dependency Limited resolution Sensitive to noisy data

Table 1. Cont.

Deep Learning Model	Target Analysis	Prediction Output	Advantages	Disadvantages
ExPecto [56]	Infers cell-type specific predictions for gene expression levels directly from DNA sequence based on histone marks, TF binding, and chromatin accessibility mapping	Epigenomic effect prediction using DNA sequences; predicting cell-type specific gene expression and effects of genomic variants	Directly from sequence predictions Tissue-type specific predictions Variant effect prediction Generalization across any human population variant	Limited to short DNA sequences Lower accuracy than benchmark models Fails to explain cross-individual variability Fails to attribute the right direction of variant's effect Complex architecture
Xpresso [57]	Analyzing the sequence elements located within +/− 1500 bp around a TSS (TFBS, TSS, and chromatin accessibility) can reveal the expected mRNA abundance for the target gene	Predicts transcriptional activity (mRNA abundance) using TSS annotations and CAGE	DL model that predicts mRNA abundance directly from DNA sequence Can infer significant predictions based only on promoter sequences and mRNA stability Simple structure Efficient Can quantify non-promoter contributions to gene expression	Narrow genomic window Incorrect attributions of the direction of the variant's effect on gene expression (increase or decrease) Limited interpretability (both due to fixed training dataset and the restricted analysis of the model)
DeFine [65]	Prediction of cell-type-specific DNA binding of TFs based on TF ChIP-seq data	Classification of TF-DNA binding or unbinding in the context of a genomic variant and prediction of the functional effect of the altered sequence	Multi-modal integration—DNA sequence, chromatin accessibility, and histone marks High accuracy of TFBS specificity prediction Captures long-range interactions	Data dependency Limited resolution High computational resources demanded Interpretability

Abbreviations: CAGE, cap analysis of gene expression; GATs, graph attention networks; RBP, RNA-binding proteins; TF, transcription factor; TFBS, transcription factor binding sites; TSS, transcription start site.

Traditional linear methods still play a complementary role. Elastic net linear regression [52,71] predicts gene expression levels from SNPs (single nucleotide polymorphisms) and remains useful in transcriptome-wide association studies.

Another approach involved in the study of genomic data is the mapping of expression quantitative trait loci (eQTL). Due to the high dimensionality of genomic data, the challenges of mapping eQTL are supposed to be overcome by the emergence of ML systems like random forests [72], which is an ensemble learning method. Random forests can perform classification and mean regression of the individual trees after building multiple decision trees. Another method for performing linear regression to simplify eQTL analysis is lasso [73], which favors solutions with fewer parameter values. Predictions of DeepSEA were shown to prioritize functional non-coding regulatory mutations in Human Gene Mutation Database (HGMD) and eQTL in the GRASP search tool [19]. These improvements to eQTL mapping present the potential to infer gene expression prediction from genetic variant genotypes [72].

3.3. Models Focusing on Epigenetic Analysis

Beyond variant interpretation, several models predict epigenetic modifications. Cp-Genie [64] predicts DNA methylation status using input DNA sequences (see Figure 2). Training for this model was performed on restricted representation bisulfite sequencing (RRBS) and whole-genome bisulfite sequencing (WGBS) profiles from ENCODE database. DeFine [65] specializes in TF-binding prediction, by training on cell-line-specific genomes instead of the human reference genome sequence. Another framework, called REUNION [74], extracts information from single-cell multi-omics data for TFBS prediction on a genome-wide scale. Multi-omics single-cell data gathering and analysis is also performed by a relatively new pipeline called MultiSC [75], developed by Lin et al., and based on a multi-modal constraint autoencoder (see Table 2).

Table 2. Deep learning frameworks that predict DNA methylation at CpG sites—a short description of the main target and the prediction output for each of them.

Model	Target Analysis	Prediction Output	Advantages	Disadvantages
CpGenie [64]	CpG methylation status	Infers predictions regarding CpG sites based on regulatory regions (enhancers, promoters, etc.)	CNN architecture capable of single-nucleotide resolutions (high sensitivity) Learns directly from DNA sequence and can produce predictions for the effect of previously unknown variants Variant prioritization	Not suitable for analyzing long-distance genomic interactions Lower performance in CpG-poor regions Data dependency It is an older model and has a less accurate performance compared to modern, transformer-based frameworks
MethylNet [76]	Methylation status at CpG sites and non-CpG sites	Enhanced prediction accuracy of methylation sites using additional chromatin data	Structured on VAEs, it can perform both supervised and unsupervised learning Multitasking model Can capture nonlinear relationships between CpG islands High-fidelity data	Requires advanced computational resources Limited interpretability Sensitive to batch effects
DeepMethyl [77]	DNA methylation status (CpG and non-CpG sites)	A hybrid model (CNN-RNN) that evaluates the methylation levels across both CpG and non-CpG regions	Integrates 3D genomic data Has integrated denoising autoencoders It performs better than benchmark ML models	In the absence of data about methylation status of the neighboring regions, its accuracy decreases Limited genomic coverage Data dependency Interpretability issues
DeepSignal [78]	CpG methylation	Distinguishes between methylated and non-methylated CpG based on nanopore sequencing data	Hybrid CNN and BiLSTM architecture which ensures a dual-feature extraction with a higher accuracy than traditional ML algorithms Very efficient for low-depth sequencing data Performs better than bisulfite sequencing	High computational demand It is outperformed by newer models Requires retraining Black box configuration limitations Data dependency

Abbreviations: BiLSTM, bidirectional long short-term memory; CNN, convolutional neural network; ML, machine learning; RNN, recurrent neural network; VAE, variational autoencoder.

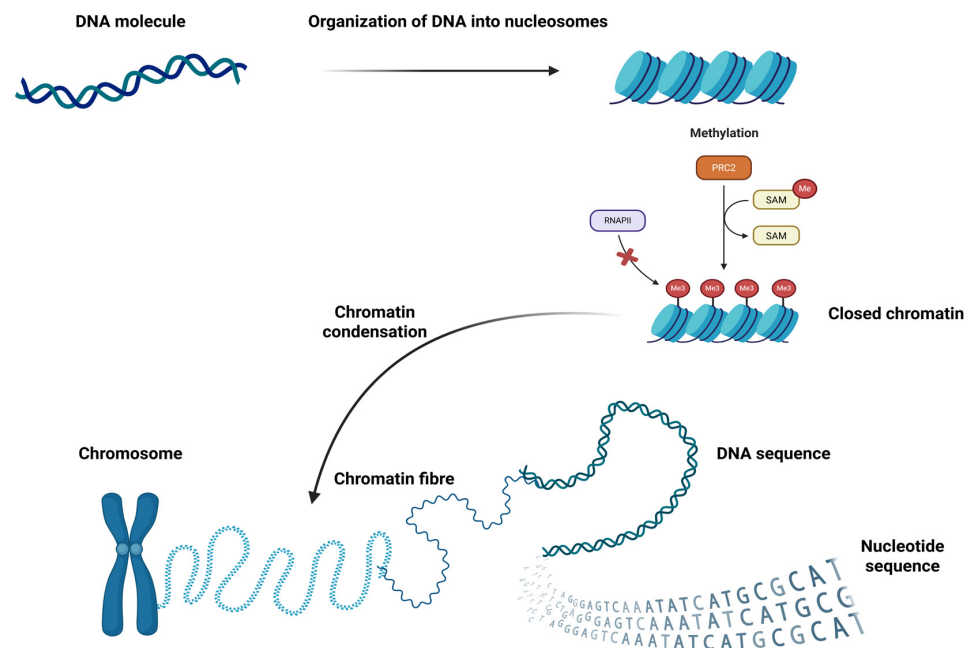


Figure 2. This figure depicts one of the most important aspects of chromatin organization, the closed chromatin state achieved by DNA methylation and histone methylation. This process is very important in gene expression studies. The prediction of DNA methylation status is the main focus for several deep learning frameworks described in this paper. Me, methyl group; PRC2, polycomb repressive complex 2; RNAPII, RNA polymerase II; SAM, S-adenosyl methionine. Created in BioRender. Palastea, E. <https://BioRender.com/iifgvui> (accessed on 29 December 2025).

Recently conducted studies aim to integrate genomic and transcriptomic predictions, such as promoter activity and gene expression levels. Models such as Expecto and Xpresso are also gene expression level predictors. Expecto is, in fact, a redesign of DeepSEA in the form of a three-stage model, and Xpresso is projected to evaluate the mRNA expression levels based solely on the genomic sequence surrounding the promoter. Another approach is that of DeepMEL [79], constructed especially for the profiling of chromatin accessibility. BPNet [19], built as a ten-layer CNN, provides base-resolution-binding affinity predictions for genomic sequences, being trained on four TFs.

3.4. Models for Transcriptomic Analysis of Gene Expression

The transcriptomic-level applications are of utter importance, pinpointing the central role of the transcriptome in the gene regulation process. The diversity of the eukaryotic transcriptome emerges from the multiple levels of regulation of the transcription, thus affecting the rate of gene expression in various ways. Essential aspects that should be assessed in this kind of analysis are the existence of multiple promoters for a gene which produce RNA transcripts with different 5'-UTRs, the RNA splicing regulation process, the possibility of existence of multiple polyadenylation sites (PASs) which produce different 3'-UTRs, and others. The majority of the DL models developed for transcriptomic-level predictions use a CNN architecture. CNNProm [80], DeeReCT-PromID [81], and DeeReCT-TSS [80] are designed mainly for promoter recognition, while other applications perform splicing predictions, especially of cassette exons (DARTS [82], SpliceAI [83], and Pangolin [84]). PAS quantification can be handled by DeeReCT-PolyA [19], APARENT [19], and DeeReCT-APA. There is also the possibility of using programs that predict subcellular localization, like RNATracker, or those that predict microRNA targets—MiRTDL [85]. Additionally, DeepBind can be employed for sequence-based RNA-protein-binding prediction. Other potential research directions in the deep learning model predictions in omics are represented by the proteomic and phenotypic-level applications [19].

3.5. Models Specifically Used for Single-Cell Studies

Most of the aforementioned models have been trained on bulk-sequencing data, which somewhat limits their ability to capture the heterogeneity found in tissues or organs subjected to genetic analysis. Single-cell transcriptomic analysis has paved the way to understanding the differentiation during development, regeneration, and disease, but lineage tracing (which assigns single-cell transcriptomes to cellular lineage trees) remains technically demanding. Single-cell technologies also generate high-dimensional, complex amounts of data that challenge even the traditional computational approaches, making them unfeasible. This explains the tendency towards upgrading from traditional statistical models and ML to DL alternatives [52].

ML algorithms have been introduced for identifying cell lineages from scRNA-seq data. A computational tool already used in this attempt is GEMLI [86] (gene expression memory-based lineage inference). GEMLI facilitates the study of heritable gene expression and is able to perform discrimination of symmetric and asymmetric cell fate decisions and reconstruction of multicellular structures from pooled scRNA-seq datasets. GEMLI can be applied to any scRNA-seq dataset, only requiring exonic reads, and is based on the characteristic expression distributions of memory genes across cell types. It is best suitable for analyzing only small to medium-sized cell lineages related over several generations, but will perform less robustly in cell populations that contain distantly related cells [86]. Another algorithm, available as an R package (compatible with R version 4.1.3 or higher), for predicting TF-gene interaction using scRNA-seq is scGATE [87].

One application was to infer single-cell gene expression profiles from TF analysis using a tree-guided multi-tasking approach, a framework used by multiple models such as TRIANGULATE [88] or SCENIC [88]. They have a similar concept, deriving a TF activity score per cell and studying the associations between single-cell gene expression and TFs. These statistical models are trained to consider the gene expression measurements of genes across single cells as being the tasks in the MTL set-up.

Further work in this field was directed towards integrating DL resources. Efforts have been undertaken in an attempt to fill in the missing temporal information of single-cell gene expression. Current techniques do not enable researchers to monitor the profile of scRNA-seq data continuously over time, therefore performing measurements at discrete time points. A realistic *in silico* prediction at any time is needed in order to have a more exact representation of the dynamics of gene expression. One framework which attempts to predict accurate gene expression at any time point is scNODE [89], which combines a variational autoencoder (VAE) with neural ordinary differential equation (ODE) [89]. A desired approach is using statistical methods for approximating ODEs that serve for modeling gene regulatory networks (GRNs) [90].

Several computational approaches based on DL algorithms have emerged lately for the various stages of scRNA-seq analysis. Most of them focus on normalization, data correction, and downstream analysis steps. Removing technical artifacts or unintended effects requires detecting and removing changes in measurements between samples and features through normalization. DL models designed for this task include DCA [91], SAUCIE [92], Auto Impute [93], DeepMc [94], Deep Impute [95], and scVI [96]. Further research on this topic intends to overcome the challenges of technical noise from scRNA-seq data.

Data correction algorithms such as ResNets [97], MNNs [97], BERMUDA [98], DESC [99], and others promise to account for variables represented by batch, dropout, and cell cycle effects. Dropout occurs when a gene is observed at a moderate or high expression level in one cell but is not observed in another cell [52,100]. Imputation methods for scRNA-seq are divided into two categories. The first type includes those who change all gene expression levels (e.g., MAGIC [101]) and single-cell analysis via expression recovery

(e.g., SAVER [102]), while the second category is based on methods that impute dropout events, such as scImpute. Most algorithms developed for dropout events identification are based on AEs (autoencoders) and the examples are AutoImpute, Deep Count AE Network (DCA), and SAUCIE. Other levels where DL algorithms can analyze scRNA-seq data are dimensionality reduction (scvis [103], BasisVAE [104], and SAUCIE), clustering, and cell annotation (DESC, scDCCA [105], and scDeepCluster [106]). Cell–cell communication can be inferred using programs such as CellChat [107] or CellPhoneDB [108]. Multi-modal integration is possible with models like scMVAE [109], DCCA, and DeepMaps [110]. DeepVelo [111] is a platform that uses RNA velocity, a technique evaluating a cell's fate based on the ratio between newly synthesized, unspliced mRNA and mature mRNA, for dimensionality reduction in the dataset [111].

Gene expression is subject to certain changes in the context of external influences. Single-cell gene expression studies can be used to explore cellular responses to infection, gene editing, and drug exposure. Understanding these effects at single-cell resolution may provide deeper insight into future clinical, medical, and pharmacological approaches. The difficulties of obtaining perturbed tissue samples and inferring a prediction from this analysis require the use of bioinformatics tools, of which DL models have a promising outlook. One of the most recent initiatives in this area uses scPRAM [112], a model for predicting perturbation responses in single-cell gene expression based on attention mechanisms. scPRAM can infer an accurate prediction of gene expression responses to perturbation for unseen cell types after aligning cell states before and after perturbation. The VAE used for this model reduces the high dimensionality and data sparsity of single-cell gene expression data, mapping this matrix to a lower-dimensional latent space [112]. Other similar approaches and deep learning models designed for the same applications are scVIDR [113] and scGen [114], also using VAE, but in different manners. Nonetheless, they appear to have a lower performance than scPRAM. Another model employed in the study of perturbation's role in gene expression prediction in single-cell techniques is CellOT [115], having a different architecture from the already mentioned frameworks. It integrates optimal transport and input convex neural architectures and seems to be very efficient in predicting single-cell drug responses. CellOT directly learns and uncovers maps between control and perturbed cell states and then predicts perturbation responses [115].

Attention-based neural networks are introduced in the development of deep learning predictive models applied for phenotype prediction. One model with this architecture is ScRAT [116] which uses scRNA-seq data for training. ScRAT consists of the following three modules: sample mixup, attention layer, and phenotype classifier. It can learn from a limited number of training samples and is independent of cell-type annotations, being a promising tool for finding phenotypic-driver cell types that can lead to the discovery of novel molecular mechanisms and targeted therapies.

3.6. The Potential of Quantum Computing Approaches for Gene Expression Prediction

Quantum computing and its intersection with ML are emerging computational paradigms that are being investigated for their applicability to high-dimensional biological data analysis. Although most applications remain in early, proof-of-concept stages and are not yet clinically deployable, these methods have the potential to become complementary tools for classical bioinformatics techniques [117].

One representative example is the development of a quantum circuit model for inferring GRNs from single-cell transcriptomic data derived from human lymphoblastoid cells [118]. Quantum models use a basic unit of information to encode data, named qubit. In this quantum single-cell GRN (qscGRN) model, each gene corresponds to a qubit. Leveraging phenomena like superposition and entanglement, scRNA sequencing data are used

to determine regulatory gene–gene dependencies within a quantum framework [118]. In addition to circuit-based models, quantum annealing approaches have been explored for optimization tasks, including single-cell RNA-seq data clustering. Alternative low-energy solutions could be obtained and may facilitate exploration of different gene grouping configurations [119].

Beyond clustering, other studies have explored the comparative capabilities of quantum machine learning (QML) versus classical frameworks for gene expression classification, assessing pattern extraction, feature relevance, and computational constraints [120]. Furthermore, quantum machine learning models, through quantum kernel methods, have been evaluated for breast cancer molecular subtype classification, showing performance comparable to classical approaches while requiring fewer data points [121].

Complementing these classification-focused efforts, other quantum frameworks focused on the identification of attractors in GRNs modeled using Boolean networks. Because the number of possible network states grows exponentially with system size, identifying these attractors (stable states associated with biological phenotypes) is computationally challenging for classical methods. To address this limitation, quantum search strategies have been proposed for attractor identification in synchronous Boolean networks, indicating resilience to noise on current noisy intermediate-scale quantum (NISQ) devices [122]. Altogether, although these approaches remain in the proof-of-concept stage, they provide early evidence that quantum techniques could offer complementary perspectives for the analysis of complex gene expression data.

4. Insight into the Role of Spatial Transcriptomics in Phenotype Prediction

There are multiple neural networks architectures that are currently used for genomic studies. Each of them is developed according to certain characteristics and designed for specific purposes. For example, MLPs, RNNs, and CNNs are frequently used in these research areas. RNNs perform sequential data analysis for text, speech, or DNA sequences, while CNNs can draw spatial relationships from image data. CNNs use convolution filters to extract granular features from a traversed image. Such applications suggest the potential of these neural networks to aid in histopathology and genomic diagnostics. Certain CNNs have already been proven efficient in predicting a few types of cancer based on digital histopathology [123].

Building on these methodological strengths, incorporating spatially resolved molecular data has become essential for placing gene expression back into its native tissue context. Spatial multi-omics technologies might be key to the characterization of a sample, which is usually a fixed fresh-frozen or formalin-fixed and paraffin-embedded tissue section that can also be stained. Most of the time, hematoxylin and eosin (HE) staining is used. When obtaining sections from the target tissue, serial sections can be specifically dissociated into single cells or nuclei which will be subjected to single-cell sequencing. Data obtained from this analysis will serve for the optimal deconvolution of spatial information and additional data integration. Apart from the previously listed methods of performing single-cell sequencing and analyzing the results, including the DL models involved in the analysis of the sparse data, there are also integrative methods which additionally perform the association with tissue morphology. Some rely on microscopy-based methods, such as single-molecule FISH (smFISH) (multiplexed error-robust FISH (MERFISH [124])), sequential FISH (seqFISH [125]), and OligoFISSEQ [126]), while others use laser capture microdissection-based isolation of single cells from the sections and techniques of single-cell sequencing (epigenomic, transcriptomic, proteomic, etc). MERFISH is based on a system where each gene has an associated binary code, where 1 signifies fluorescence and 0 means

no fluorescence. seqFISH relies on a system that assigns a color sequence code for each gene, having twenty-four color probes per gene and sixty different pseudo-color options [127]. The integration of data generally requires different computational models which can deal with the large amount of sparse genomics data, and the current focus is on DL models that use various types of neural networks (VAE, adversarial autoencoders—AAEs, ViT, and LSTM) [128].

ST offers the opportunity of studying gene expression patterns while preserving the spatial configuration of the intact tissue [9]. Applications developed for ST mapping of cell types are Bayesian models, such as Cell2location [129], which uses scRNA-seq reference of cell types integrated with 10x Visium, having a 55 μm resolution and roughly 3–30 cells per capture spot to spatially resolve fine-grained cell types. ST techniques such as Visium and Slide-seq [130] present the disadvantage of their resolution being far lower than single-cell level, each pixel in Slide-seq covering multiple cells (~10 microns), while Visium spots average around 10–20 cells for each 50 microns. Consequently, the measured gene expression at a spot reflects a mixture of cells [131].

The main focus of research regarding gene expression prediction is the design of methods of studying and obtaining omics data without damaging the tissue, preserving its initial architecture and pattern of gene expression. Another important target would be inferring such data from viable, living cells, providing information from a dynamic point of view. These methods attempt to extract features that are much closer to the real-time events happening within a tissue or organ.

A useful modality to monitor live cells is offered by Raman microscopy, “a unique tool that relies on the collection of vibrational energy levels of molecules” [13]. Raman microscopy has the advantages of being nondestructive and label-free but lacks genetic and molecular interpretability. Recent research by Kobayashi-Kirschvink et al. introduced a computational framework named Raman2RNA (R2R) [13] that can infer single-cell RNA expression profiles from label-free nondestructive Raman hyperspectral images. The approach integrates single-molecule fluorescence in situ hybridization (smFISH)-anchor-based integration data of selected markers from the same cells, and scRNA-seq profiles to train the model. Prediction of scRNA profiles can also be performed in an anchor-free manner with AAEs [132]. The final step is the translation of Raman images into single-cell expression profiles. R2R ultimately offers a label-free live-cell inference of single-cell expression profiles over time, thanks to tracking live single cells by time-lapse Raman imaging. Briefly, R2R could provide expression profiles at single-cell resolution by using multiplex smFISH as an anchor between the images and the scRNA-seq profiles [13].

ST plays an increasingly important role in today’s research in gene expression prediction, as it mixes gene expression data with the localization of cells in a tissue. It unveils aspects regarding cell communication and interaction in their natural environment. On a single-cell scale, it further develops an improved image of the tissue heterogeneity and the cooperation and function of each cell inside a particular organ or tissue. Besides Raman, plenty of ST-focused frameworks have been implemented. The use of DL applications in this field is of utter interest, considering the complexity and quantity of the extracted data, which sometimes exceeds the human analysis capacity.

To overcome these analytical limitations, several neural frameworks now explicitly integrate patch-level morphology with spatial gene expression matrices. TISSUE [132] is a DL framework implemented as a spatial gene prediction model, relying on ST and scRNA-seq data. It estimates uncertainty for spatial gene expression prediction for unmeasured transcripts and is based on supervised learning. Histology images used as a support for inferring gene expression patterns present a growing interest, considering the reduced costs of slide preparation compared to the complexity of other protocols for

similar purposes. HGGEp [133] is a hypergraph neural network model developed for predicting gene expression levels from histology images. HGGEp partitions the slide in multiple patches centered around spots. Cell morphological information perception of the model is enhanced by integrating a gradient enhancement module (GEM). The modular structure employing attention mechanisms and LSTM enables this model to refine spatial information and produce detailed gene expression landscapes [133]. To further exploit the link between tissue morphology and gene expression, Pham et al. created stLearn (version v1.2.2) [134]. This software was designed for identifying cell types, observing cell–cell interactions and the influence neighboring cells inflict upon transcriptional processes at their location in the tissue. This framework integrates a pseudo-time-space algorithm that traces the changes in gene expression over time across tissue regions, which can reveal links between tumor progression and local responses [134].

Prediction of gene expression from whole-slide images (WSIs) data is gaining more and more interest from the research field, introducing methods such as ST-Net [135], a DL model that captures high-resolution gene expression heterogeneity from histology slides and ST data. Pizurica et al. worked on a DL framework capable of leveraging gene expression predictions from histology slides. Their application, named SEQUOIA [136], focuses on prediction of cancer-associated genes and uses a DL architecture with linearized attention, performing a digitalized histopathological interpretation. Other DL programs designed for similar purposes are HisToGene [137], DeepPT [138], Hist2ST [139], and THItToGene [140]. WSI morphological features have been associated with genomic mutations [141], gene expression profiles, and methylation patterns [142].

Deep generative models have also emerged. They are used to generate WSI tiles after being infused with matched gene expression profiles. Carrillo-Perez et al. designed a model called RNA-GAN that utilizes a VAE to learn a latent representation of the multi-tissue gene expression profiles which is then transferred to a generative adversarial network (GAN). GAN generates synthetic tissue tiles based on the learnt information [143]. Expanding beyond generative approaches, Zheng et al. identified a particular role of artificial intelligence in predicting immune and inflammatory gene expression signatures from hepatocellular carcinoma histology, proving a possible diagnostic purpose [144]. Building on this momentum, Zhao et al. developed SpiRiT, a vision transformer (ViT) framework to infer spatial single-cell transcriptomic signatures from HE-stained histology slides and named it SpiRiT [145] (Spatial Omics Prediction and Reproducibility Integrated Transformer). This model was tested for predictions for human breast cancer and whole mouse pup histology slides [145].

Some initiatives to increase the resolution of ST to single-cell level have emerged and they rely on AI computational models. SPOTlight [146] is a deconvolution algorithm built upon a non-negative matrix factorization regression algorithm and least squares regression to determine a spot's composition, defining cell-type-specific topic profiles and the characteristic gene distribution of a cell type. It enables an automated interpretation through the possibility to integrate unpaired ST and scRNA-seq data. The limitations consist of the ineffective learning and use of the intrinsic topological information of cell-type constitutions within spots. A potential tool to overcome these challenges seems to be the development of semi-supervised models based on graph convolutional networks (GCNs), in which intrinsic topological information can be represented as graphs. Deconvoluting ST data through graph-based convolutional networks (DSTGs) models have shown good results, are based on scRNA-seq datasets, and are able to learn the precise composition of ST data using a semi-supervised GCN [131].

Another model, STEM [147] (SpaTially aware EMbedding), applies deep transfer learning to jointly encode both ST and single-cell data into a shared space, allowing pseudo-

spatial mapping between cells. This approach compensates for the lack of spatial resolution in single-cell methods and for the limited coverage of ST techniques.

scResolve [148], developed by Chen et al., is a method for recovering single-cell expression profiles from ST measurements at multi-cellular resolution, restoring expression profiles of single cells at their specific location, a task which cannot be performed by deconvolution. scResolve combines spot-level expression profiles with the paired histology slide image, resulting in subcellular gene maps. It then segments these maps into individual cells and produces expression profiles. Segmentation is performed firstly by identifying the nuclei from tissue-staining images; then, a DL transformer model infers, for each spot from the gene expression map, if it is a part of a cell or of the extracellular matrix and its relative position related to the center of its nucleus. Finally, the spots identified as parts of the cells are grouped according to the relative positions to nucleus centers.

Stellaris web server [149] offers fast and accurate spatial mapping of user-uploaded scRNA-seq data. Many additional AI applications, such as SOMDE [150], scGCO [151], SEDR [152], coSTA [153], STAGATE [154], and GCNG [155], support ST analysis but do not reach true single-cell resolution. Single-cell resolution was achieved with ST data by using NCEM [156].

ST has shed light on multiple aspects of omics, solving problems related to the dimensionality of data and providing a holistic approach to gene expression. Another factor that contributes to reducing the bias of transcriptomic gene expression characterization in ST data is retaining the spatial configuration, as it does not need tissue dissociation. This helps with providing the tissue context that is very relevant for understanding the signaling between cells at a known location [9,134]. The advancement towards single-cell resolution certainly increased the performance of the techniques by deconvolution and referencing single-cell datasets. Morphological identification relying directly on pixel-level analysis might further reduce the error occurrence observed in the single-cell segmentation. Limitations still occur, with many AI models employed for such tasks failing to identify consensus tissue domains (TDs). In an attempt to improve clustering of cell types based on similarities in gene expression patterns or molecular functions, Kaur et al. proposed MILWRM [157], an algorithm designed for TDs characterization using multiplex immunofluorescence and ST data.

5. Translational Implications

AI applications have been introduced in the clinical settings, gradually becoming a powerful tool for healthcare providers. No-code AI platforms, such as the one described by Hoseini et al. [158], are very useful for people that do not have expertise in coding and can perform tasks such as white blood cells classification. Other AI frameworks are tested for dental medicine applications [159], boosting diagnostic accuracy, treatment, and monitoring through teledentistry [160]. The great enhancement in interpreting imaging data brought by AI also encouraged initiatives to implement such models in cardiovascular imaging [161]. Further research for fine-tuning these models targets optimization of microscopy techniques and nanoscale imaging, as it is described in the work of Zhao et al., where they used ML for expansion microscopy samples from clinical specimens [162].

As these clinical applications multiply, single-cell and spatial omics technologies are beginning to reshape the landscape of translational research by providing unprecedented molecular detail. Single-cell omics technologies uncover tissue and organ heterogeneity at an unprecedented scale and have shown promising results in conjunction with various AI programs for predicting gene expression. Clinical application should certainly benefit from the thorough analysis that can be performed by these innovative methods. The huge ability of single-cell techniques to dissect tissue heterogeneity has impacted the way of

analyzing malignancy and certainly helped with understanding the processes that govern the tumoral microenvironment and cell communication at unprecedented resolution. The details revealed by using these techniques have brought new insights in studying targets, drugs, and molecular pathways that play important roles in producing pathologies [163]. Oncology is one of the fields that will massively use the evolution of single-cell sequencing and AI algorithms for diagnostic and treatment purposes, taking personalized medicine a step further. There have already been initiatives that integrated such technologies, with Fan et al. using ML algorithms for analyzing single-cell RNA-seq data for osteosarcoma, particularly targeting CD8⁺ T cells [164].

Zhao et al. applied their work with SpiRiT [145] to improve targeted therapy strategies, integrating ST data with the histopathological landscape of tumors, to provide a more precise map of gene expression. SpiRiT predicts spatial gene expression profiles at a single-cell resolution for human breast cancer and also works for multiple species and tissues. Patient outcome in oncology can see important improvement in the future after implementing these tools in cancer therapy and diagnosis [145].

AI modeling promises not only to enhance the drug development and diagnostic area, but it also aims to predict the outcomes of different diseases. Personalized medicine holds the premises for uncovering the underlying mechanisms of diseases in individuals in a patient-tailored manner. AI-driven models may soon ensure dynamic risk prediction, allowing the anticipation of disease progression or treatment response based on leveraging genomics and transcriptomic signatures. The role of minor mutations may be established thanks to these frameworks, such as the influence of SNPs on drug susceptibility in different patients. One of the main goals of AI-based omics analysis remains the identification of clinically relevant variants. In clinical practice, promising models are being adapted for patient monitoring, selection of candidates for clinical trials, and risk stratification [20]. AI algorithms and models have demonstrated powerful clustering capacity, proving to be increasingly useful for patient risk stratification based on selected biomarkers. For example, applications for sepsis phenotyping speed up the diagnosis, predictions for clinical outcome of the patient, and bring new insights into optimization of treatment thanks to the analysis of drug response, dosage in different sub-phenotypes, and potential drug targets [165–167]. Such frameworks can integrate clinical parameters and omics data for a thorough selection of relevant information used in the predictions.

In parallel, spatial and single-cell deconvolution methods are being tested directly on patient tissues to identify cellular niches and subpopulations linked to clinical behavior. DSTG [131] was tested on pancreatic ductal adenocarcinoma sections, providing insight into its capacity of spatially mapping the prevalence of different cell subpopulations. The experiment additionally revealed the ability of detecting unique cytoarchitectures and cancer-linked phenomena.

The majority of DL systems tested for single-cell gene expression in the context of ST analysis have been applied to cerebral cortex regions and rendered a topological structure similar to the real distribution. Building on these early demonstrations of spatial fidelity, the most important applications, as mentioned above, are the oncology-related ones. Hao et al. used their tool called STEM [147] for characterization of tumor microenvironments on a single-cell scale based on a dataset of human squamous cell carcinoma. The results were comparable to manual gene selection but managed to reduce the analysis time and the error rates encountered in the human-curated study. STEM also showed promising results in retrieving gene expression patterns from liver sections and endothelial cells, analyzing spatial variation. These advances underscore how AI is becoming indispensable for interpreting complex transcriptional programs in disease.

It is worth noting that machine and DL platforms proved to be useful in the context of cancer immunotherapy. This is especially critical because cancer immunotherapy has emerged as one of the most transformative treatment paradigms in modern oncology, yet its efficacy is profoundly shaped by the cellular heterogeneity and immune landscape that only single-cell technologies can resolve. In this research area, the shift from bulk to single-cell technology has contributed to the scale at which tissue heterogeneity was observed and brought more accuracy to the predictions. The quantity of encountered data also turned the researchers' attention to AI alternatives that can work out a solution for the rich transcriptomic information. AI approach is both able to remove technical noise and to provide novel insights into omics profiles. These algorithms can decipher the underlying mechanisms for tumor resistance to immunotherapy and also uncover the weak spots that can be attacked with treatment alternatives [168]. By illuminating these mechanisms of therapeutic resistance, AI-driven single-cell analysis provides a foundation for more rational and personalized immunotherapy strategies.

In parallel with these immunotherapy-focused advances, researchers have increasingly used DL to explore additional regulatory layers that influence tumor behavior, particularly alternative splicing programs involved in development and metastasis. Zhang et al. developed a DL framework that adds data from empirical RNA sequencing to produce predictions of alternative splicing sites. This platform is called DARTS [82] and was used for studying processes encountered in the development of embryos and cancer metastasis. In this case, because the process studied concerns the epithelial–mesenchymal transformation, the system performed best when trained on combined information from two databases, ENCODE and Roadmap [82].

Other DL algorithms were designed for gene expression-based cancer prediction, such as WT-GAN [169], which operates on microarray gene expression cancer data as input (see Table 3 for a summary of the models that have translational applications). Together, these translational frameworks highlight the accelerating convergence of single-cell biology, spatial omics, and DL in modern precision medicine (Figure 3).

Table 3. Translational applications for ML and DL frameworks trained on single-cell sequencing data. This table offers a short description of the architecture of the deep learning model and the potential applications for diagnostic purposes.

Framework	Architecture	Application	Reference
LASSO, XGboost, GBM, Boruta, CoxBoost, survival-SVM	ML algorithms	Predicting T cell exhaustion signature from genes in osteosarcoma	Fan et al. [164]
SpiRiT	ViT	Predicting spatial gene expression profiles for human breast cancer	Zhao et al. [145]
DSTG	Graph-based convolutional network	Deconvolution of pancreatic cancer tissue sections	Song and Su [131]
STEM	MLP encoder Transfer learning	Characterization of tumor microenvironment on single-cell scale based on datasets of squamous cell carcinoma; retrieving gene expression patterns from liver sections	Hao et al. [147]

Table 3. Cont.

Framework	Architecture	Application	Reference
iMAP	GANs and deep autoencoders	Removes batch effects and identifies batch-specific cells	Gui et al. [168]
DARTS	DNN and a Bayesian hypothesis testing statistical model	Useful for studying the development of embryos and cancer metastasis	Zhang et al. [82]
WT-GAN	GAN	Predicting cancer types based on MGCD datasets	Ravindran and Gunavathi [169]
Seurat [82]	ML algorithm	Characterization of epithelial cell lineages in lung adenocarcinoma	Zhang et al. [82]
PMG training framework	5 PMGs ViT-L model, Segment Anything Model, and ResNeXt framework	Differentiation between benign and malignant cells and leukemia typing	Yan et al. [170]

Abbreviations: DNN, deep neural network; GAN, generative adversarial network; ML, machine learning; MLP, multi-layer perceptron; PMG, multistage progressive multigranularity; and ViT, vision transformer.

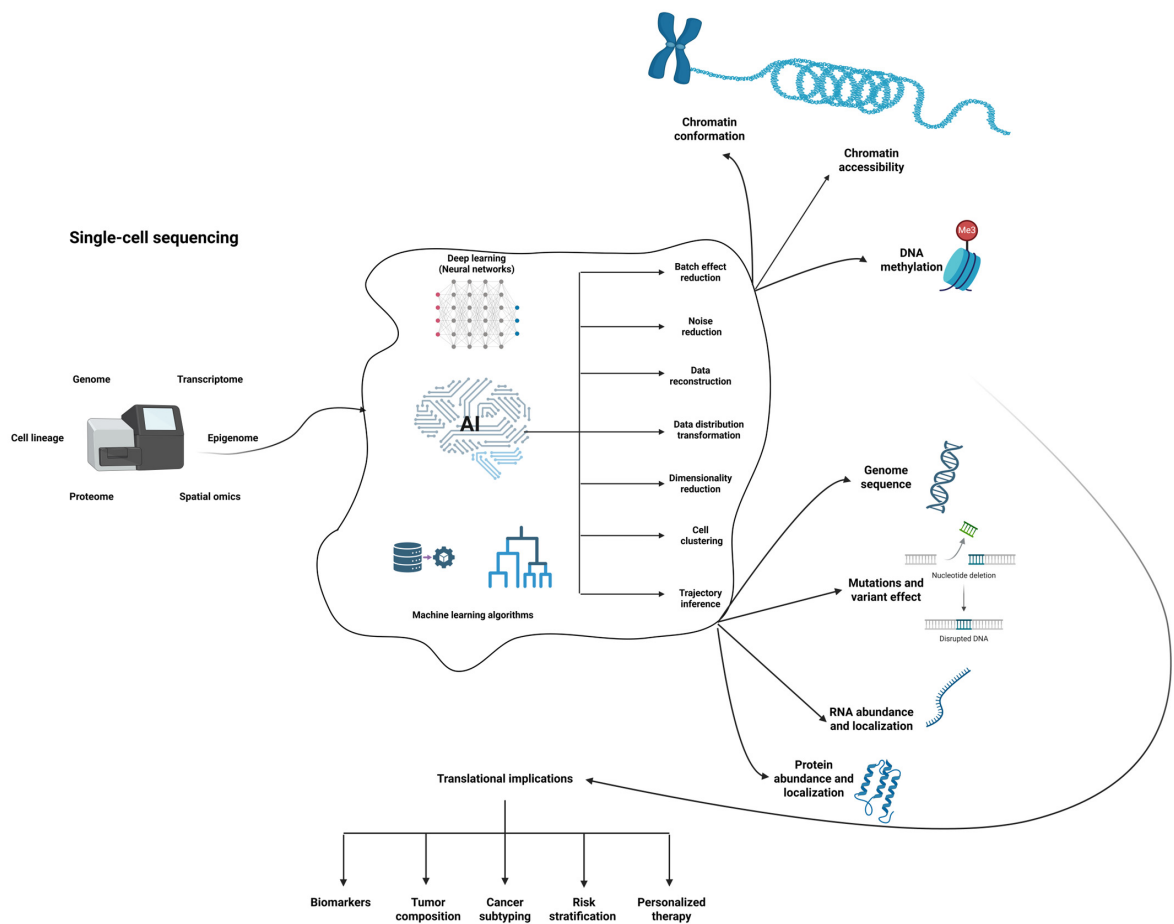


Figure 3. Summarizing diagram that presents a simplified workflow from single-cell sequencing to gene expression prediction using AI methods and the translational applications resulting from the inferred cell phenotypes. Abbreviations: AI, artificial intelligence. Created in BioRender. Palastea, E. <https://BioRender.com/asl891t> (accessed on 7 January 2026).

6. Benefits and Limitations

The rapidly growing interest in single-cell technologies and in developing novel computational approaches to handle the immense quantity of sequencing and spatial data are driving an accelerated evolution in these domains. The need for continuous improvement is demonstrated by the huge potential of these methods to fill the gaps in omics research and omics-related clinical approach. However, every step forward comes with inherent limitations, and the technological surge requires careful validation to define each application's role and prevent errors (see Table 4).

For example, single-cell RNA sequencing (scRNA-seq) encountered issues regarding the temperature conditions used in the protocol. The dissociation process at 37 °C induced artificial changes in the expression profiles following the activation of stress genes transcription distorting gene expression profiles. This problem was mitigated by lowering the temperature at 4 °C during dissociation, but there may be other factors that can trigger the expression of stress genes [1]. scRNA-seq, even with the help of DL platforms, cannot offer enough details related to the spatial context from which the cells were extracted, due to the tissue dissociation process. Nonetheless, even with improved protocols, the loss of spatial context during dissociation remains a major gap in single-cell data. This limitation naturally shifts attention toward spatial methods and supports their integration with single-cell approaches for a more complete biological picture.

The single-cell technologies raise issues connected to the statistical errors that may interfere with the final results' reading. Transcriptomic technologies have broadened the knowledge regarding the complexity of certain tissues and the social network that establishes between cells. It complements single-cell sequencing, because ST itself lacks the desired resolution. Although single-cell omics fail to preserve spatial integrity, thereby losing details, the information retrieved from spatially resolved transcriptomics combined with the single-cell techniques can produce an enhanced picture of the organization within tissues and also offer a dynamic perspective [163].

Another limitation of scRNA-seq is the clustering resolution. DL frameworks such as stLearn improve clustering by integrating spatial information and tissue morphology for gene expression prediction, using SMEclust, to find subclusters of cells that were not identified by classical scRNA-seq methods [134].

Single-cell multi-omics have paved the way to deciphering cell population heterogeneity within tissues but encountered several limitations over time, like high costs and technical difficulty. The downsides of the combined approach of single-cell multi-omics and histology slide imaging reside in the heterogeneity of the sample (not all of the omics techniques are compatible with all sample types) and the encounter of single-cell statistical uncertainty. Using histology slides for inferring gene expression prediction is a cost-reducing alternative, but nonetheless a challenging task, having its limitations in ignoring the complex relationship between cell morphology and gene expression patterns and not utilizing the full features extracted from the images. There has certainly been huge improvement, also with the aid of advanced computational methods and new technology that have emerged [171]. Nonetheless, one aspect that still restricts the number of applications for single-cell omics is the statistical part that generates computational errors. Technical limitations raise concerns when analyzing results from scRNA sequencing, as there is a high dropout rate. A dropout event is, in fact, a false "zero" result, which translates into the inability to detect an expressed gene because of a low RNA capture rate. A true zero result represents a gene that is not expressed in a particular cell type. Shorter genes present a greater dropout rate and have lower counts [52]. The need for discerning between real and false zero values has led to the development of statistical models and DL methods, but this area still deals with a lot of cases where a suitable model has not been introduced yet. Data

sparsity in scRNA-seq is a challenge that can impede downstream analyses and confronts scientists with multiple statistical problems. Some of them are related to the uncertainty of whether the observed zero values from the analysis are a reflection of the absence of gene expression in the cell or they occur due to a lack of data regarding that gene [171].

Variations in scRNA-seq data stem from multiple factors, including sequencing depth, variation in cell type, and technical artifacts. To overcome such limitations, the statistical models should account for error distribution and parameter uncertainty. Models suitable for this kind of analysis often rely on Poisson distribution, negative binomial distribution, generalized linear models (GLMs), and Pearson residuals for dimensionality reduction. Whilst this approach solves statistical issues that hinder the full recovery of data from scRNA-seq, it cannot be directly applied to new single-cell platforms [172].

Table 4. Benefits and limitations of the discussed methods.

Methods Used	Benefits	Limitations
scRNA-seq	• High resolution • Reveals tissue heterogeneity • Captures rare cell populations • Lineage and trajectory analysis • Enables understanding of microenvironment interactions	• Artificial transcriptional stress responses • Lack of spatial context • Statistical errors • Low-resolution clustering • High dropout rate and false zero values • Data sparsity • Variations in cell sequencing depth
ST	• Preserves spatial integrity • Maps tissue architecture • Infers spatial gene expression patterns • Provides a link between transcriptomics and morphology of a tissue • Suitable for pathology studies (tumors, fibrosis, immune cells infiltration, etc.)	• Lacks desired resolution • Low transcript detection • Complexity • High cost • Complex interpretation tools—pipelines—for data analysis
DL platforms [173]	• Can find subclusters and patterns that are not identified by classical scRNA-seq methods • Efficient use of sparse, high-dimensional data • Can extract non-linear relationships • Can be fitted for large datasets • Can integrate multiomics data • Improved denoising	• “Black box” configuration—hidden states of the network that often impede a trustworthy biological prediction. The complex layered architecture masks interactions between millions of parameters within the networks, rendering their decision-making process hardly comprehensible to most human observers • Require a suitable dataset • Require computational resources • Must be fed with large labeled datasets • Selection of a suitable model for the task • Ethical concerns regarding sensitive data usage for training the models (data privacy) • Lack of suitably structured data of the clinical information systems for feeding AI models • Biased predictions produced by training on clinical data retrieved from limited cohorts of study

Abbreviations: AI, artificial intelligence; DL, deep learning; scRNA-seq, single-cell RNA sequencing; and ST, spatial transcriptomics.

Interpretation of (ST) data also raised a lot of challenges and some issues still persist, with methods like cellular segmentation and annotation being prone to error. While these techniques are gaining popularity, especially because they have been complemented by AI-powered curation lately, there are multiple aspects that are not adequately tackled by the current technologies. Many of them are unable to identify consensus tissue domains (TDs), which refer to grouping tissues or cell types based on similar gene expression patterns, molecular functions, or biological roles [157].

Particularities of both ML algorithms and DL models as AI subsets indicate the advantages and limitations of the methods. On one side, ML depends on careful data preprocessing that usually requires multiple steps to select suitable data. On the other side, DL has the advantage of automated learning directly from raw or minimally processed data. This feature allows DL to deal better with complex and non-linear data and makes the process less time-consuming. Regarding the training data, ML can usually be trained on small and medium datasets, while DL requires large datasets. DL presents the risk of overfitting, thus needing the large sets of data to train on. Thanks to the statistical and mathematical structure of ML models, these present more interpretability and are easier to follow by a human observer, while the DL frameworks have the “black box” configuration that usually prevents the observers from understanding how predictions are made by the model. This issue is one of the main reasons for reticent implementation of DL in clinical applications. A further difference can be seen in the costs, as the DL models require high computational resources and longer training times. The comparative performance of ML and DL depends on the task; DL can outperform classical ML in the case of large datasets or where complex genomic interactions are analyzed, but it does not consistently perform better. ML tends to have a more stable performance across datasets thanks to the underlying algorithms, and this feature makes it suitable for classification tasks and disease phenotype prediction [174].

7. Conclusions

The continued development of innovative approaches for omics data analysis has substantially advanced our understanding of cellular heterogeneity, gene regulation, and disease-associated molecular mechanisms. Omics medicine lies at the intersection of genomics and bioinformatics, showing that computational methods are essential for understanding complex biological processes. AI-based algorithms have already made a significant contribution to gene expression prediction, enabling the identification of regulatory patterns that were previously difficult to capture using traditional analytical frameworks.

Although gene regulation remains incompletely understood, ongoing refinements in ML and DL architectures, together with improved optimization strategies, continue to reduce prediction error rates. Single-cell sequencing technologies have improved data consistency and reproducibility, strengthening the reliability of computational modeling. The inclusion of spatial omics technologies, designed to extract gene expression profiles at single-cell resolution, combined with the increasingly proficient AI platforms, represent major steps toward precision medicine. Together, these integrative approaches allow us to reconstruct cellular states in multiple dimensions, revealing how transcriptional programs interact with the tissue environments that shape them.

These advances build a strong foundation for patient-tailored predictions of disease progression and therapeutic response. Continued progress in experimental technologies, complemented by emerging computational paradigms, including quantum computing, is expected to further accelerate biomedical research and deepen our understanding of GRNs underlying human health and disease.

Author Contributions: E.A.P. performed the literature search and drafted the initial version of the manuscript. I.-M.M. and E.R. contributed with their expertise and ideas and improved the first draft of the manuscript. O.B. and O.H. supervised the project and critically revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Publication of this manuscript is supported by the Genomics Research and Development Institute, the University of Medicine and Pharmacy Carol Davila, through institutional Open access program, and from The Ministry of Investment and European Projects, through the Managing Authority for the Health Program, PS/272/PS_P5/OP1/RSO1.1/PS_P5_RSO1.1_A9-ROGEN Project (MySMIS 324809).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: Authors would like to thank host institutions for support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

AI, artificial intelligence; BiLSTM, bidirectional long short-term memory; CAGE, cap analysis of gene expression; cDNA, complementary DNA; CNNs, convolutional neural networks; DL, deep learning; DNA, deoxyribonucleic acid; DNNs, deep neural networks; eQTL, expression quantitative trait locus; FISH, fluorescence in situ hybridization; GANs, generative adversarial networks; GAT, graph attention network; GEM, gradient enhancement module; GNNs, graph neural networks; GRNs, gene regulatory networks; GRUs, gated recurrent units; GWAS, genome-wide association studies; HE, hematoxylin and eosin; LSTM, long short-term memory; ML, machine learning; MLPs, multi-layer perceptrons; mRNA, messenger ribonucleic acid; QC, quality control; RNA, ribonucleic acid; RNNs, recurrent neural networks; sc, single-cell; scRNA-seq, single-cell RNA sequencing; snRNA-seq, single-nucleus RNA sequencing; ST, spatial transcriptomics; TF, transcription factor; TFBS, transcription factor-binding sites; VAE, variational autoencoder; WSI, whole-slide images.

References

1. Jovic, D.; Liang, X.; Zeng, H.; Lin, L.; Xu, F.; Luo, Y. Single-cell RNA Sequencing Technologies and Applications: A Brief Overview. *Clin. Transl. Med.* **2022**, *12*, e694. [\[CrossRef\]](#)
2. Kashima, Y.; Sakamoto, Y.; Kaneko, K.; Seki, M.; Suzuki, Y.; Suzuki, A. Single-Cell Sequencing Techniques from Individual to Multiomics Analyses. *Exp. Mol. Med.* **2020**, *52*, 1419–1427. [\[CrossRef\]](#)
3. Li, X.; Wang, C.Y. From Bulk, Single-Cell to Spatial RNA Sequencing. *Int. J. Oral Sci.* **2021**, *13*, 36. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Zhu, C.; Preissl, S.; Ren, B. Single-Cell Multimodal Omics: The Power of Many. *Nat. Methods* **2020**, *17*, 11–14. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Cuomo, A.S.E.; Nathan, A.; Raychaudhuri, S.; MacArthur, D.G.; Powell, J.E. Single-Cell Genomics Meets Human Genetics. *Nat. Rev. Genet.* **2023**, *24*, 535–549. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Wasney, M.; Pott, S. Simultaneous Measurement of DNA Methylation and Nucleosome Occupancy in Single Cells Using ScNOME-Seq. In *Chromatin Accessibility*; Humana: New York, NY, USA, 2023; Volume 2611.
7. Liu, L.; Liu, C.; Quintero, A.; Wu, L.; Yuan, Y.; Wang, M.; Cheng, M.; Leng, L.; Xu, L.; Dong, G.; et al. Deconvolution of Single-Cell Multi-Omics Layers Reveals Regulatory Heterogeneity. *Nat. Commun.* **2019**, *10*, 470. [\[CrossRef\]](#)
8. Bai, Y.; Deng, X.; Chen, D.; Han, S.; Lin, Z.; Li, Z.; Tong, W.; Li, J.; Wang, T.; Liu, X.; et al. Integrative Analysis Based on ATAC-Seq and RNA-Seq Reveals a Novel Oncogene PRPF3 in Hepatocellular Carcinoma. *Clin. Epigenet.* **2024**, *16*, 154. [\[CrossRef\]](#)
9. Williams, C.G.; Lee, H.J.; Asatsuma, T.; Vento-Tormo, R.; Haque, A. An Introduction to Spatial Transcriptomics for Biomedical Research. *Genome Med.* **2022**, *14*, 68. [\[CrossRef\]](#)
10. Trapnell, C.; Cacchiarelli, D.; Grimsby, J.; Pokharel, P.; Li, S.; Morse, M.; Lennon, N.J.; Livak, K.J.; Mikkelsen, T.S.; Rinn, J.L. The Dynamics and Regulators of Cell Fate Decisions Are Revealed by Pseudotemporal Ordering of Single Cells. *Nat. Biotechnol.* **2014**, *32*, 381–386. [\[CrossRef\]](#)
11. Schiebinger, G.; Shu, J.; Tabaka, M.; Cleary, B.; Subramanian, V.; Solomon, A.; Gould, J.; Liu, S.; Lin, S.; Berube, P.; et al. Optimal-Transport Analysis of Single-Cell Gene Expression Identifies Developmental Trajectories in Reprogramming. *Cell* **2019**, *176*, 928–943.e22. Erratum in *Cell* **2019**, *176*, 1517. [\[CrossRef\]](#)

12. Hou, W.; Ji, Z.; Chen, Z.; Wherry, E.J.; Hicks, S.C.; Ji, H. A Statistical Framework for Differential Pseudotime Analysis with Multiple Single-Cell RNA-Seq Samples. *Nat. Commun.* **2023**, *14*, 7286. [\[CrossRef\]](#)
13. Kobayashi-Kirschvink, K.J.; Comiter, C.S.; Gaddam, S.; Joren, T.; Grody, E.I.; Ounadjela, J.R.; Zhang, K.; Ge, B.; Kang, J.W.; Xavier, R.J.; et al. Prediction of Single-Cell RNA Expression Profiles in Live Cells by Raman Microscopy with Raman2RNA. *Nat. Biotechnol.* **2024**, *42*, 1726–1734. [\[CrossRef\]](#)
14. Chen, W.; Guillaume-Gentil, O.; Rainer, P.Y.; Gäbelein, C.G.; Saelens, W.; Gardeux, V.; Klaeger, A.; Dainese, R.; Zachara, M.; Zambelli, T.; et al. Live-Seq Enables Temporal Transcriptomic Recording of Single Cells. *Nature* **2022**, *608*, 733–740. [\[CrossRef\]](#)
15. Keyl, P.; Bischoff, P.; Dernbach, G.; Bockmayr, M.; Fritz, R.; Horst, D.; Blüthgen, N.; Montavon, G.; Müller, K.R.; Klauschen, F. Single-Cell Gene Regulatory Network Prediction by Explainable AI. *Nucleic Acids Res.* **2023**, *51*, E20. [\[CrossRef\]](#)
16. Elshewey, A.M. Enhancing Crop Yield Prediction Based on Dove Optimization Algorithm and Gradient Boosting Model. *Signal Image Video Process.* **2025**, *19*, 951. [\[CrossRef\]](#)
17. El-Rashidy, N.; Tarek, Z.; Elshewey, A.M.; Shams, M.Y. Multitask Multilayer-Prediction Model for Predicting Mechanical Ventilation and the Associated Mortality Rate. *Neural Comput. Appl.* **2025**, *37*, 1321–1343. [\[CrossRef\]](#)
18. Tarek, Z.; Shams, M.Y.; Towfek, S.K.; Alkahtani, H.K.; Ibrahim, A.; Abdelhamid, A.A.; Eid, M.M.; Khodadadi, N.; Abualigah, L.; Khafaga, D.S.; et al. An Optimized Model Based on Deep Learning and Gated Recurrent Unit for COVID-19 Death Prediction. *Biomimetics* **2023**, *8*, 552. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Li, Z.; Gao, E.; Zhou, J.; Han, W.; Xu, X.; Gao, X. Applications of Deep Learning in Understanding Gene Regulation. *Cell Rep. Methods* **2023**, *3*, 100384. [\[CrossRef\]](#)
20. Nosrati, H.; Nosrati, M. Artificial Intelligence in Regenerative Medicine: Applications and Implications. *Biomimetics* **2023**, *8*, 442. [\[CrossRef\]](#)
21. Schwessinger, R.; Deasy, J.; Woodruff, R.T.; Young, S.; Branson, K.M. Single-Cell Gene Expression Prediction from DNA Sequence at Large Contexts. *BioRxiv* **2023**. [\[CrossRef\]](#)
22. Buccitelli, C.; Selbach, M. MRNAs, Proteins and the Emerging Principles of Gene Expression Control. *Nat. Rev. Genet.* **2020**, *21*, 630–644. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Wang, D.; Eraslan, B.; Wieland, T.; Hallström, B.; Hopf, T.; Zolg, D.P.; Zecha, J.; Asplund, A.; Li, L.; Meng, C.; et al. A Deep Proteome and Transcriptome Abundance Atlas of 29 Healthy Human Tissues. *Mol. Syst. Biol.* **2019**, *15*, e8503. [\[CrossRef\]](#)
24. Chick, J.M.; Munger, S.C.; Simecek, P.; Huttlin, E.L.; Choi, K.; Gatti, D.M.; Raghupathy, N.; Svenson, K.L.; Churchill, G.A.; Gygi, S.P. Defining the Consequences of Genetic Variation on a Proteome-Wide Scale. *Nature* **2016**, *534*, 500–505, Correction in *Nature* **2022**, *606*, E16. [\[CrossRef\]](#)
25. Avsec, Ž.; Agarwal, V.; Visentin, D.; Ledsam, J.R.; Grabska-Barwinska, A.; Taylor, K.R.; Assael, Y.; Jumper, J.; Kohli, P.; Kelley, D.R. Effective Gene Expression Prediction from Sequence by Integrating Long-Range Interactions. *Nat. Methods* **2021**, *18*, 1196–1203. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Alipanahi, B.; Delong, A.; Weirauch, M.T.; Frey, B.J. Predicting the Sequence Specificities of DNA- and RNA-Binding Proteins by Deep Learning. *Nat. Biotechnol.* **2015**, *33*, 831–838. [\[CrossRef\]](#)
27. Auerbach, B.J.; Hu, J.; Reilly, M.P.; Li, M. Applications of Single-Cell Genomics and Computational Strategies to Study Common Disease and Population-Level Variation. *Genome Res.* **2021**, *31*, 1728–1741. [\[CrossRef\]](#)
28. Chen, G.; Ning, B.; Shi, T. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Front. Genet.* **2019**, *10*, 317. [\[CrossRef\]](#)
29. Van de Sande, B.; Lee, J.S.; Mutasa-Gottgens, E.; Naughton, B.; Bacon, W.; Manning, J.; Wang, Y.; Pollard, J.; Mendez, M.; Hill, J.; et al. Applications of Single-Cell RNA Sequencing in Drug Discovery and Development. *Nat. Rev. Drug Discov.* **2023**, *22*, 496–520. [\[CrossRef\]](#)
30. He, J.; Lin, L.; Chen, J. Practical Bioinformatics Pipelines for Single-Cell RNA-Seq Data Analysis. *Biophys Rep.* **2022**, *8*, 158–169. [\[PubMed\]](#)
31. van den Brink, S.; Sage, F.; Vertesy, A.; Spanjaard, B.; Peterson-Maduro, J.; Baron, C.; Robin, C.; van Oudenaarden, A. Single-Cell Sequencing Reveals Dissociation-Induced Gene Expression in Tissue Subpopulations. *Nat. Methods* **2017**, *14*, 935–936. [\[CrossRef\]](#)
32. Ding, J.; Adiconis, X.; Simmons, S.K.; Kowalczyk, M.S.; Hession, C.C.; Marjanovic, N.D.; Hughes, T.K.; Wadsworth, M.H.; Burks, T.; Nguyen, L.T.; et al. Systematic Comparison of Single-Cell and Single-Nucleus RNA-Sequencing Methods. *Nat. Biotechnol.* **2020**, *38*, 737–746, Correction in *Nat. Biotechnol.* **2020**, *38*, 756. [\[CrossRef\]](#) [\[PubMed\]](#)
33. Picelli, S.; Faridani, O.R.; Björklund, Å.K.; Winberg, G.; Sagasser, S.; Sandberg, R. Full-Length RNA-Seq from Single Cells Using Smart-Seq2. *Nat. Protoc.* **2014**, *9*, 171–181. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Bageritz, J.; Raddi, G. Single-Cell RNA Sequencing with Drop-Seq. In *Single Cell Methods. Methods in Molecular Biology*; Humana: New York, NY, USA, 2019; Volume 1979, pp. 73–85.
35. DeLaughter, D. The Use of the Fluidigm C1 for RNA Expression Analyses of Single Cells. *Curr. Protoc. Mol. Biol.* **2018**, *122*, e55. [\[CrossRef\]](#)

36. Danielski, K. Guidance on Processing the 10× Genomics Single Cell Gene Expression Assay. In *Methods in Molecular Biology*; Humana: New York, NY, USA, 2022; Volume 2584.
37. Sheng, K.; Cao, W.; Niu, Y.; Deng, Q.; Zong, C. Effective Detection of Variation in Single-Cell Transcriptomes Using MATQ-Seq. *Nat. Methods* **2017**, *14*, 267–270. [[CrossRef](#)] [[PubMed](#)]
38. Gierahn, T.M.; Wadsworth, M.H.; Hughes, T.K.; Bryson, B.D.; Butler, A.; Satija, R.; Fortune, S.; Christopher Love, J.; Shalek, A.K. Seq-Well: Portable, Low-Cost Rna Sequencing of Single Cells at High Throughput. *Nat. Methods* **2017**, *14*, 395–398. Erratum in *Nat. Methods* **2017**, *14*, 752. [[CrossRef](#)]
39. Liu, C.; Wu, T.; Fan, F.; Liu, Y.; Wu, L.; Junkin, M.; Wang, Z.; Yu, Y.; Wang, W.; Wei, W.; et al. A Portable and Cost-Effective Microfluidic System for Massively Parallel Single-Cell Transcriptome Profiling. *BioRxiv* **2019**. [[CrossRef](#)]
40. Kolodziejczyk, A.A.; Kim, J.K.; Svensson, V.; Marioni, J.C.; Teichmann, S.A. The Technology and Biology of Single-Cell RNA Sequencing. *Mol. Cell* **2015**, *58*, 610–620. [[CrossRef](#)]
41. Keren-Shaul, H.; Kenigsberg, E.; Jaitin, D.A.; David, E.; Paul, F.; Tanay, A.; Amit, I. MARS-Seq2.0: An Experimental and Analytical Pipeline for Indexed Sorting Combined with Single-Cell RNA Sequencing. *Nat. Protoc.* **2019**, *14*, 1841–1862. [[CrossRef](#)]
42. Hashimshony, T.; Wagner, F.; Sher, N.; Yanai, I. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep.* **2012**, *2*, 666–673. [[CrossRef](#)]
43. Juzenas, S.; Goda, K.; Kisieliovas, V.; Zvirblyte, J.; Quintinal-Villalonga, A.; Siurkus, J.; Nainys, J.; Mazutis, L. InDrops-2: A Flexible, Versatile and Cost-Efficient Droplet Microfluidic Approach for High-Throughput ScRNA-Seq of Fresh and Preserved Clinical Samples. *Nucleic Acids Res.* **2025**, *53*, gkae1312. [[CrossRef](#)]
44. Heumos, L.; Schaar, A.C.; Lance, C.; Litnetskaya, A.; Drost, F.; Zappia, L.; Lücken, M.D.; Strobl, D.C.; Henao, J.; Curion, F.; et al. Best Practices for Single-Cell Analysis across Modalities. *Nat. Rev. Genet.* **2023**, *24*, 550–572. [[CrossRef](#)]
45. Zheng, G.X.Y.; Terry, J.M.; Belgrader, P.; Ryvkin, P.; Bent, Z.W.; Wilson, R.; Ziraldo, S.B.; Wheeler, T.D.; McDermott, G.P.; Zhu, J.; et al. Massively Parallel Digital Transcriptional Profiling of Single Cells. *Nat. Commun.* **2017**, *8*, 14049. [[CrossRef](#)] [[PubMed](#)]
46. Xu, X.; Zhang, Q.; Li, M.; Lin, S.; Liang, S.; Cai, L.; Zhu, H.; Su, R.; Yang, C. Microfluidic Single-Cell Multiomics Analysis. *View* **2023**, *4*, 20220034. [[CrossRef](#)]
47. De Jonghe, J.; Kaminski, T.S.; Morse, D.B.; Tabaka, M.; Ellermann, A.L.; Kohler, T.N.; Amadei, G.; Handford, C.E.; Findlay, G.M.; Zernicka-Goetz, M.; et al. SpinDrop: A Droplet Microfluidic Platform to Maximise Single-Cell Sequencing Information Content. *Nat. Commun.* **2023**, *14*, 4788. [[CrossRef](#)] [[PubMed](#)]
48. Hong, R.; Koga, Y.; Bandyadka, S.; Leshchik, A.; Wang, Y.; Akavoor, V.; Cao, X.; Sarfraz, I.; Wang, Z.; Alabdullatif, S.; et al. Comprehensive Generation, Visualization, and Reporting of Quality Control Metrics for Single-Cell RNA Sequencing Data. *Nat. Commun.* **2022**, *13*, 1688. [[CrossRef](#)] [[PubMed](#)]
49. Zhao, X.; Du, A.; Qiu, P. ScMODD: A Model-Driven Algorithm for Doublet Identification in Single-Cell RNA-Sequencing Data. *Front. Syst. Biol.* **2022**, *2*, 1082309. [[CrossRef](#)]
50. Wolock, S.L.; Lopez, R.; Klein, A.M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.* **2019**, *8*, 281–291.e9. [[CrossRef](#)]
51. Bais, A.S.; Kostka, D. Gene Expression Scds: Computational Annotation of Doublets in Single-Cell RNA Sequencing Data. *Bioinformatics* **2020**, *36*, 1150–1158. [[CrossRef](#)]
52. Erfanian, N.; Heydari, A.A.; Feriz, A.M.; Iañez, P.; Derakhshani, A.; Ghasemigol, M.; Farahpour, M.; Razavi, S.M.; Nasseri, S.; Safarpour, H.; et al. Deep Learning Applications in Single-Cell Genomics and Transcriptomics Data Analysis. *Biomed. Pharmacother.* **2023**, *165*, 115077. [[CrossRef](#)]
53. Shen, X.; Jiang, C.; Wen, Y.; Li, C.; Lu, Q. A Brief Review on Deep Learning Applications in Genomic Studies. *Front. Syst. Biol.* **2022**, *2*, 877717. [[CrossRef](#)]
54. Yue, T.; Wang, Y.; Zhang, L.; Gu, C.; Xue, H.; Wang, W.; Lyu, Q.; Dun, Y. Deep Learning for Genomics: From Early Neural Nets to Modern Large Language Models. *Int. J. Mol. Sci.* **2023**, *24*, 15858. [[CrossRef](#)]
55. Ramprasad, P.; Pai, N.; Pan, W. Enhancing Personalized Gene Expression Prediction from DNA Sequences Using Genomic Foundation Models. *Hum. Genet. Genom. Adv.* **2024**, *5*, 100347. [[CrossRef](#)]
56. Zhou, J.; Theesfeld, C.L.; Yao, K.; Chen, K.M.; Wong, A.K.; Troyanskaya, O.G. Deep Learning Sequence-Based Ab Initio Prediction of Variant Effects on Expression and Disease Risk. *Nat. Genet.* **2018**, *50*, 1171–1179. [[CrossRef](#)]
57. Agarwal, V.; Shendure, J. Predicting mRNA Abundance Directly from Genomic Sequence Using Deep Convolutional Neural Networks. *Cell Rep.* **2020**, *31*, 107663. [[CrossRef](#)] [[PubMed](#)]
58. Kelley, D.R. Cross-Species Regulatory Sequence Activity Prediction. *PLoS Comput. Biol.* **2020**, *16*, e1008050. [[CrossRef](#)]
59. Zeng, H.; Edwards, M.D.; Liu, G.; Gifford, D.K. Convolutional Neural Network Architectures for Predicting DNA-Protein Binding. *Bioinformatics* **2016**, *32*, i121–i127. [[CrossRef](#)] [[PubMed](#)]
60. Zhou, J.; Troyanskaya, O.G. Predicting Effects of Noncoding Variants with Deep Learning-Based Sequence Model. *Nat. Methods* **2015**, *12*, 931–934. [[CrossRef](#)] [[PubMed](#)]

61. Hu, X.; Fernie, A.R.; Yan, J. Deep Learning in Regulatory Genomics: From Identification to Design. *Curr. Opin. Biotechnol.* **2023**, *79*, 102887. [\[CrossRef\]](#)
62. Kelley, D.R.; Snoek, J.; Rinn, J.L. Basset: Learning the Regulatory Code of the Accessible Genome with Deep Convolutional Neural Networks. *Genome Res.* **2016**, *26*, 990–999. [\[CrossRef\]](#)
63. Quang, D.; Xie, X. DanQ: A Hybrid Convolutional and Recurrent Deep Neural Network for Quantifying the Function of DNA Sequences. *Nucleic Acids Res.* **2016**, *44*, e107. [\[CrossRef\]](#)
64. Zeng, H.; Gifford, D.K. Predicting the Impact of Non-Coding Variants on DNA Methylation. *Nucleic Acids Res.* **2017**, *45*, e99. [\[CrossRef\]](#)
65. Wang, M.; Tai, C.; Weinan, E.; Wei, L. DeFine: Deep Convolutional Neural Networks Accurately Quantify Intensities of Transcription Factor-DNA Binding and Facilitate Evaluation of Functional Non-Coding Variants. *Nucleic Acids Res.* **2018**, *46*, E69. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Fudenberg, G.; Kelley, D.R.; Pollard, K.S. Predicting 3D Genome Folding from DNA Sequence with Akita. *Nat. Methods* **2020**, *17*, 1111–1117. [\[CrossRef\]](#)
67. Zhou, J. Sequence-Based Modeling of Three-Dimensional Genome Architecture from Kilobase to Chromosome Scale. *Nat. Genet.* **2022**, *54*, 725–734. [\[CrossRef\]](#)
68. Karbalayghareh, A.; Sahin, M.; Leslie, C.S. Chromatin Interaction-Aware Gene Regulatory Modeling with Graph Attention Networks. *Genome Res.* **2022**, *32*, 930–944. [\[CrossRef\]](#) [\[PubMed\]](#)
69. Dalla-Torre, H.; Gonzalez, L.; Mendoza-Revilla, J.; Lopez Carranza, N.; Grzywaczewski, A.H.; Oteri, F.; Dallago, C.; Trop, E.; de Almeida, B.P.; Sirelkhatim, H.; et al. Nucleotide Transformer: Building and Evaluating Robust Foundation Models for Human Genomics. *Nat. Methods* **2025**, *22*, 287–297. [\[CrossRef\]](#)
70. Sasse, A.; Ng, B.; Spiro, A.E.; Tasaki, S.; Bennett, D.A.; Gaiteri, C.; De Jager, P.L.; Chikina, M.; Mostafavi, S. Benchmarking of Deep Neural Networks for Predicting Personal Gene Expression from DNA Sequence Highlights Shortcomings. *Nat. Genet.* **2023**, *55*, 2060–2064. [\[CrossRef\]](#)
71. Mai, J.; Lu, M.; Gao, Q.; Zeng, J.; Xiao, J. Transcriptome-Wide Association Studies: Recent Advances in Methods, Applications and Available Databases. *Commun. Biol.* **2023**, *6*, 899. [\[CrossRef\]](#)
72. Xie, R.; Quitadamo, A.; Cheng, J.; Shi, X. A Predictive Model of Gene Expression Using a Deep Learning Framework. In Proceedings of the 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Shenzhen, China, 15–18 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 676–681.
73. Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Statist. Soc. B* **1996**, *58*, 267–288. [\[CrossRef\]](#)
74. Yang, Y.; Pe'er, D. REUNION: Transcription Factor Binding Prediction and Regulatory Association Inference from Single-Cell Multi-Omics Data. *Bioinformatics* **2024**, *40*, i567–i575. [\[CrossRef\]](#)
75. Lin, X.; Jiang, S.; Gao, L.; Wei, Z.; Wang, J. MultiSC: A Deep Learning Pipeline for Analyzing Multiomics Single-Cell Data. *Brief. Bioinform.* **2024**, *25*, bbae492. [\[CrossRef\]](#)
76. Levy, J.J.; Titus, A.J.; Petersen, C.L.; Chen, Y.; Salas, L.A.; Christensen, B.C. MethylNet: An Automated and Modular Deep Learning Approach for DNA Methylation Analysis. *BMC Bioinform.* **2020**, *21*, 108. [\[CrossRef\]](#)
77. Wang, Y.; Liu, T.; Xu, D.; Shi, H.; Zhang, C.; Mo, Y.Y.; Wang, Z. Predicting DNA Methylation State of CpG Dinucleotide Using Genome Topological Features and Deep Networks. *Sci. Rep.* **2016**, *6*, 19598. [\[CrossRef\]](#)
78. Ni, P.; Huang, N.; Zhang, Z.; Wang, D.P.; Liang, F.; Miao, Y.; Xiao, C.L.; Luo, F.; Wang, J. DeepSignal: Detecting DNA Methylation State from Nanopore Sequencing Reads Using Deep-Learning. *Bioinformatics* **2019**, *35*, 4586–4595. [\[CrossRef\]](#)
79. Minnoye, L.; Taskiran, I.I.; Mauduit, D.; Fazio, M.; van Aershot, L.; Hulselmans, G.; Christiaens, V.; Makhzami, S.; Seltenhammer, M.; Karras, P.; et al. Cross-Species Analysis of Enhancer Logic Using Deep Learning. *Genome Res.* **2020**, *31*, 1815–1834. [\[CrossRef\]](#)
80. Zhou, J.; Zhang, B.; Li, H.; Zhou, L.; Li, Z.; Long, Y.; Han, W.; Wang, M.; Cui, H.; Li, J.; et al. Annotating TSSs in Multiple Cell Types Based on DNA Sequence and RNA-Seq Data via DeeReCT-TSS. *Genom. Proteom. Bioinform.* **2022**, *20*, 959–973. [\[CrossRef\]](#) [\[PubMed\]](#)
81. Umarov, R.; Kuwahara, H.; Li, Y.; Gao, X.; Solovyev, V. Promoter Analysis and Prediction in the Human Genome Using Sequence-Based Deep Learning Models. *Bioinformatics* **2019**, *35*, 2730–2737. [\[CrossRef\]](#)
82. Zhang, Z.; Pan, Z.; Ying, Y.; Xie, Z.; Adhikari, S.; Phillips, J.; Carstens, R.P.; Black, D.L.; Wu, Y.; Xing, Y. Deep-Learning Augmented RNA-Seq Analysis of Transcript Splicing. *Nat. Methods* **2019**, *16*, 307–310. [\[CrossRef\]](#)
83. Jaganathan, K.; Kyriazopoulou Panagiotopoulou, S.; McRae, J.F.; Darbandi, S.F.; Knowles, D.; Li, Y.I.; Kosmicki, J.A.; Arbelaez, J.; Cui, W.; Schwartz, G.B.; et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **2019**, *176*, 535–548.e24. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Zeng, T.; Li, Y.I. Predicting RNA Splicing from DNA Sequence Using Pangolin. *Genome Biol.* **2022**, *23*, 103. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Cheng, S.; Guo, M.; Wang, C.; Liu, X.; Liu, Y.; Wu, X. MiRTDL: A Deep Learning Approach for MiRNA Target Prediction. *IEEE ACM Trans. Comput. Biol. Bioinform.* **2015**, *13*, 1161–1169. [\[CrossRef\]](#)

86. Eisele, A.S.; Tarbier, M.; Dormann, A.A.; Pelechano, V.; Suter, D.M. Gene-Expression Memory-Based Prediction of Cell Lineages from ScRNA-Seq Datasets. *Nat. Commun.* **2024**, *15*, 2744, Correction in *Nat. Commun.* **2024**, *15*, 4752. [\[CrossRef\]](#)
87. Malekpour, S.A.; Haghverdi, L.; Sadeghi, M. Single-Cell Multi-Omics Analysis Identifies Context-Specific Gene Regulatory Gates and Mechanisms. *Brief. Bioinform.* **2024**, *25*, bbae180. [\[CrossRef\]](#)
88. Ardakani, F.B.; Kattler, K.; Heinen, T.; Schmidt, F.; Feuerborn, D.; Gasparoni, G.; Lepikhov, K.; Nell, P.; Hengstler, J.; Walter, J.; et al. Prediction of Single-Cell Gene Expression for Transcription Factor Analysis. *Gigascience* **2020**, *9*, giaa113. [\[CrossRef\]](#)
89. Zhang, J.; Larschan, E.; Bigness, J.; Singh, R. ScNODE: Generative Model for Temporal Single Cell Transcriptomic Data Prediction. *Bioinformatics* **2024**, *40*, ii146–ii154. [\[CrossRef\]](#)
90. Hossain, I.; Fanfani, V.; Fischer, J.; Quackenbush, J.; Burkholz, R. Biologically Informed NeuralODEs for Genome-Wide Regulatory Dynamics. *Genome Biol.* **2024**, *25*, 127. [\[CrossRef\]](#)
91. Eraslan, G.; Simon, L.M.; Mircea, M.; Mueller, N.S.; Theis, F.J. Single-Cell RNA-Seq Denoising Using a Deep Count Autoencoder. *Nat. Commun.* **2019**, *10*, 390. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Amodio, M.; van Dijk, D.; Srinivasan, K.; Chen, W.S.; Mohsen, H.; Moon, K.R.; Campbell, A.; Zhao, Y.; Wang, X.; Venkataswamy, M.; et al. Exploring Single-Cell Data with Deep Multitasking Neural Networks. *Nat. Methods* **2019**, *16*, 1139–1145. [\[CrossRef\]](#) [\[PubMed\]](#)
93. Talwar, D.; Mongia, A.; Sengupta, D.; Majumdar, A. AutoImpute: Autoencoder Based Imputation of Single-Cell RNA-Seq Data. *Sci. Rep.* **2018**, *8*, 16329. [\[CrossRef\]](#)
94. Mongia, A.; Sengupta, D.; Majumdar, A. DeepMc: Deep Matrix Completion for Imputation of Single Cell RNA-Seq Data. *J. Comput. Biol.* **2018**, *27*, 1011–1019. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Arisdakessian, C.; Poirion, O.; Yunits, B.; Zhu, X.; Garmire, L.X. DeepImpute: An Accurate, Fast, and Scalable Deep Neural Network Method to Impute Single-Cell RNA-Seq Data. *Genome Biol.* **2019**, *20*, 211. [\[CrossRef\]](#)
96. Lopez, R.; Regier, J.; Cole, M.B.; Jordan, M.I.; Yosef, N. Deep Generative Modeling for Single-Cell Transcriptomics. *Nat. Methods* **2018**, *15*, 1053–1058. [\[CrossRef\]](#)
97. Zou, B.; Zhang, T.; Zhou, R.; Jiang, X.; Yang, H.; Jin, X.; Bai, Y. DeepMNN: Deep Learning-Based Single-Cell RNA Sequencing Data Batch Correction Using Mutual Nearest Neighbors. *Front. Genet.* **2021**, *12*, 708981. [\[CrossRef\]](#)
98. Wang, T.; Johnson, T.S.; Shao, W.; Lu, Z.; Helm, B.R.; Zhang, J.; Huang, K. BERMUDA: A Novel Deep Transfer Learning Method for Single-Cell RNA Sequencing Batch Correction Reveals Hidden High-Resolution Cellular Subtypes. *Genome Biol.* **2019**, *20*, 165. [\[CrossRef\]](#)
99. Li, X.; Wang, K.; Lyu, Y.; Pan, H.; Zhang, J.; Stambolian, D.; Susztak, K.; Reilly, M.P.; Hu, G.; Li, M. Deep Learning Enables Accurate Clustering with Batch Effect Removal in Single-Cell RNA-Seq Analysis. *Nat. Commun.* **2020**, *11*, 2338. [\[CrossRef\]](#)
100. Qiu, P. Embracing the Dropouts in Single-Cell RNA-Seq Analysis. *Nat. Commun.* **2020**, *11*, 1169. [\[CrossRef\]](#)
101. van Dijk, D.; Sharma, R.; Nainys, J.; Yim, K.; Kathail, P.; Carr, A.J.; Burdziak, C.; Moon, K.R.; Chaffer, C.L.; Pattabiraman, D.; et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **2018**, *174*, 716–729.e27. [\[CrossRef\]](#)
102. Huang, M.; Wang, J.; Torre, E.; Dueck, H.; Shaffer, S.; Bonasio, R.; Murray, J.I.; Raj, A.; Li, M.; Zhang, N.R. SAVER: Gene Expression Recovery for Single-Cell RNA Sequencing. *Nat. Methods* **2018**, *15*, 539–542. [\[CrossRef\]](#) [\[PubMed\]](#)
103. Ding, J.; Condon, A.; Shah, S.P. Interpretable Dimensionality Reduction of Single Cell Transcriptome Data with Deep Generative Models. *Nat. Commun.* **2018**, *9*, 2002. [\[CrossRef\]](#) [\[PubMed\]](#)
104. Märtens, K.; Yau, C. BasisVAE: Translation-Invariant Feature-Level Clustering with Variational Autoencoders. In Proceedings of the International Conference on Artificial Intelligence and Statistics, Online, 26–28 August 2020.
105. Wang, J.; Xia, J.; Wang, H.; Su, Y.; Zheng, C.H. ScDCCA: Deep Contrastive Clustering for Single-Cell RNA-Seq Data Based on Auto-Encoder Network. *Brief. Bioinform.* **2023**, *24*, bbac625. [\[CrossRef\]](#) [\[PubMed\]](#)
106. Tian, T.; Wan, J.; Song, Q.; Wei, Z. Clustering Single-Cell RNA-Seq Data with a Model-Based Deep Learning Approach. *Nat. Mach. Intell.* **2019**, *1*, 191–198. [\[CrossRef\]](#)
107. Jin, S.; Guerrero-Juarez, C.F.; Zhang, L.; Chang, I.; Ramos, R.; Kuan, C.H.; Myung, P.; Plikus, M.V.; Nie, Q. Inference and Analysis of Cell-Cell Communication Using CellChat. *Nat. Commun.* **2021**, *12*, 1088. [\[CrossRef\]](#)
108. Efremova, M.; Vento-Tormo, M.; Teichmann, S.A.; Vento-Tormo, R. CellPhoneDB: Inferring Cell-Cell Communication from Combined Expression of Multi-Subunit Ligand-Receptor Complexes. *Nat. Protoc.* **2020**, *15*, 1484–1506. [\[CrossRef\]](#) [\[PubMed\]](#)
109. Zuo, C.; Chen, L. Deep-Joint-Learning Analysis Model of Single Cell Transcriptome and Open Chromatin Accessibility Data. *Brief. Bioinform.* **2021**, *22*, bbaa287. [\[CrossRef\]](#) [\[PubMed\]](#)
110. Ma, A.; Wang, X.; Li, J.; Wang, C.; Xiao, T.; Liu, Y.; Cheng, H.; Wang, J.; Li, Y.; Chang, Y.; et al. Single-Cell Biological Network Inference Using a Heterogeneous Graph Transformer. *Nat. Commun.* **2023**, *14*, 964. [\[CrossRef\]](#)
111. Cui, H.; Maan, H.; Vladoiu, M.C.; Zhang, J.; Taylor, M.D.; Wang, B. DeepVelo: Deep Learning Extends RNA Velocity to Multi-Lineage Systems with Cell-Specific Kinetics. *Genome Biol.* **2024**, *25*, 27. [\[CrossRef\]](#) [\[PubMed\]](#)
112. Jiang, Q.; Chen, S.; Chen, X.; Jiang, R. Gene Expression ScPRAM Accurately Predicts Single-Cell Gene Expression Perturbation Response Based on Attention Mechanism. *Bioinformatics* **2024**, *40*, btae265. [\[CrossRef\]](#)

113. Kana, O.; Nault, R.; Filipovic, D.; Marri, D.; Zacharewski, T.; Bhattacharya, S. Generative Modeling of Single-Cell Gene Expression for Dose-Dependent Chemical Perturbations. *Patterns* **2023**, *4*, 100817. [\[CrossRef\]](#)
114. Lotfollahi, M.; Wolf, F.A.; Theis, F.J. ScGen Predicts Single-Cell Perturbation Responses. *Nat. Methods* **2019**, *16*, 715–721. [\[CrossRef\]](#)
115. Bunne, C.; Stark, S.G.; Gut, G.; del Castillo, J.S.; Levesque, M.; Lehmann, K.V.; Pelkmans, L.; Krause, A.; Rättsch, G. Learning Single-Cell Perturbation Responses Using Neural Optimal Transport. *Nat. Methods* **2023**, *20*, 1759–1768, Correction in *Nat. Methods* **2023**, *20*, 1830. [\[CrossRef\]](#)
116. Mao, Y.; Lin, Y.Y.; Wong, N.K.Y.; Volik, S.; Sar, F.; Collins, C.; Ester, M. Phenotype Prediction from Single-Cell RNA-Seq Data Using Attention-Based Neural Networks. *Bioinformatics* **2024**, *40*, btae067. [\[CrossRef\]](#)
117. Nałecz-Charkiewicz, K.; Charkiewicz, K.; Nowak, R.M. Quantum Computing in Bioinformatics: A Systematic Review Mapping. *Brief. Bioinform.* **2024**, *25*, bbae391. [\[CrossRef\]](#)
118. Roman-Vicharra, C.; Cai, J.J. Quantum Gene Regulatory Networks. *Npj Quantum Inf.* **2023**, *9*, 67. [\[CrossRef\]](#)
119. Kubacki, M.; Niranjani, M. Quantum Annealing-Based Clustering of Single Cell RNA-Seq Data. *Brief. Bioinform.* **2023**, *24*, bbad377. [\[CrossRef\]](#) [\[PubMed\]](#)
120. Ghosh, A.; Fuad, M.M.; Bhattacharjee, S. Empirical Quantum Advantage Analysis of Quantum Kernel in Gene Expression Data. *arXiv* **2024**, arXiv:2411.07276. [\[CrossRef\]](#)
121. Repetto, V.; Ceroni, E.G.; Buonaiuto, G.; D'Aurizio, R. Quantum Enhanced Stratification of Breast Cancer: Exploring Quantum Expressivity for Real Omics Data. *Quantum Mach. Intell.* **2025**, *7*, 81. [\[CrossRef\]](#)
122. Rossini, M.; Weidner, F.M.; Ankerhold, J.; Kestler, H.A. A Novel Quantum Algorithm for Efficient Attractor Search in Gene Regulatory Networks. *Patterns* **2025**, *6*, 101295. [\[CrossRef\]](#)
123. Tran, K.A.; Kondrashova, O.; Bradley, A.; Williams, E.D.; Pearson, J.V.; Waddell, N. Deep Learning in Cancer Diagnosis, Prognosis and Treatment Selection. *Genome Med.* **2021**, *13*, 152. [\[CrossRef\]](#) [\[PubMed\]](#)
124. Xia, C.; Babcock, H.P.; Moffitt, J.R.; Zhuang, X. Multiplexed Detection of RNA Using MERFISH and Branched DNA Amplification. *Sci. Rep.* **2019**, *9*, 7721. [\[CrossRef\]](#) [\[PubMed\]](#)
125. Eng, C.H.L.; Lawson, M.; Zhu, Q.; Dries, R.; Kouloua, N.; Takei, Y.; Yun, J.; Cronin, C.; Karp, C.; Yuan, G.C.; et al. Transcriptome-Scale Super-Resolved Imaging in Tissues by RNA SeqFISH+. *Nature* **2019**, *568*, 235–239. [\[CrossRef\]](#)
126. Nguyen, H.Q.; Chatteraj, S.; Castillo, D.; Nguyen, S.C.; Nir, G.; Lioutas, A.; Hershsberg, E.A.; Martins, N.M.C.; Reginato, P.L.; Hannan, M.; et al. 3D Mapping and Accelerated Super-Resolution Imaging of the Human Genome Using in Situ Sequencing. *Nat. Methods* **2020**, *17*, 822–832. [\[CrossRef\]](#)
127. Longo, S.K.; Guo, M.G.; Ji, A.L.; Khavari, P.A. Integrating Single-Cell and Spatial Transcriptomics to Elucidate Intercellular Tissue Dynamics. *Nat. Rev. Genet.* **2021**, *22*, 627–644. [\[CrossRef\]](#) [\[PubMed\]](#)
128. Vandereyken, K.; Sifrim, A.; Thienpont, B.; Voet, T. Methods and Applications for Single-Cell and Spatial Multi-Omics. *Nat. Rev. Genet.* **2023**, *24*, 494–515. [\[CrossRef\]](#)
129. Kleshchevnikov, V.; Shmatko, A.; Dann, E.; Aivazidis, A.; King, H.W.; Li, T.; Elmentaite, R.; Lomakin, A.; Kedlian, V.; Gayoso, A.; et al. Cell2location Maps Fine-Grained Cell Types in Spatial Transcriptomics. *Nat. Biotechnol.* **2022**, *40*, 661–671. [\[CrossRef\]](#)
130. Rodriques, S.G.; Stickels, R.R.; Goeva, A.; Martin, C.A.; Murray, E.; Vanderburg, C.R.; Welch, J.; Chen, L.M.; Chen, F.; Macosko, E.Z. Slide-Seq: A Scalable Technology for Measuring Genome-Wide Expression at High Spatial Resolution. *Science* **2019**, *363*, 1463–1467. [\[CrossRef\]](#)
131. Song, Q.; Su, J. DSTG: Deconvoluting Spatial Transcriptomics Data through Graph-Based Artificial Intelligence. *Brief. Bioinform.* **2021**, *22*, bbaa414. [\[CrossRef\]](#)
132. Sun, E.D.; Ma, R.; Navarro Negredo, P.; Brunet, A.; Zou, J. TISSUE: Uncertainty-Calibrated Prediction of Single-Cell Spatial Transcriptomics Improves Downstream Analyses. *Nat. Methods* **2024**, *21*, 444–454. [\[CrossRef\]](#) [\[PubMed\]](#)
133. Li, B.; Zhang, Y.; Wang, Q.; Zhang, C.; Li, M.; Wang, G.; Song, Q. Gene Expression Prediction from Histology Images via Hypergraph Neural Networks. *Brief. Bioinform.* **2024**, *25*, bbae500. [\[CrossRef\]](#)
134. Pham, D.; Tan, X.; Balderson, B.; Xu, J.; Grice, L.F.; Yoon, S.; Willis, E.F.; Tran, M.; Lam, P.Y.; Raghubar, A.; et al. Robust Mapping of Spatiotemporal Trajectories and Cell–Cell Interactions in Healthy and Diseased Tissues. *Nat. Commun.* **2023**, *14*, 7739. [\[CrossRef\]](#) [\[PubMed\]](#)
135. He, B.; Bergensträhle, L.; Stenbeck, L.; Abid, A.; Andersson, A.; Borg, Å.; Maaskola, J.; Lundeberg, J.; Zou, J. Integrating Spatial Gene Expression and Breast Tumour Morphology via Deep Learning. *Nat. Biomed. Eng.* **2020**, *4*, 827–834. [\[CrossRef\]](#)
136. Pizurica, M.; Zheng, Y.; Carrillo-Perez, F.; Noor, H.; Yao, W.; Wohlfart, C.; Vladimirova, A.; Marchal, K.; Gevaert, O. Digital Profiling of Gene Expression from Histology Images with Linearized Attention. *Nat. Commun.* **2024**, *15*, 9886. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Pang, M.; Su, K.; Li, M. Leveraging Information in Spatial Transcriptomics to Predict Super-Resolution Gene Expression from Histology Images in Tumors. *BioRxiv* **2021**. [\[CrossRef\]](#)

138. Hoang, D.T.; Dinstag, G.; Shulman, E.D.; Hermida, L.C.; Ben-Zvi, D.S.; Elis, E.; Caley, K.; Sammut, S.J.; Sinha, S.; Sinha, N.; et al. A Deep-Learning Framework to Predict Cancer Treatment Response from Histopathology Images through Imputed Transcriptomics. *Nat. Cancer* **2024**, *5*, 1305–1317. [\[CrossRef\]](#)
139. Zeng, Y.; Wei, Z.; Yu, W.; Yin, R.; Yuan, Y.; Li, B.; Tang, Z.; Lu, Y.; Yang, Y. Spatial Transcriptomics Prediction from Histology Jointly through Transformer and Graph Neural Networks. *Brief. Bioinform.* **2022**, *23*, bbac297. [\[CrossRef\]](#)
140. Jia, Y.; Liu, J.; Chen, L.; Zhao, T.; Wang, Y. THltoGene: A Deep Learning Method for Predicting Spatial Transcriptomics from Histological Images. *Brief. Bioinform.* **2024**, *25*, bbad464. [\[CrossRef\]](#)
141. Qu, H.; Zhou, M.; Yan, Z.; Wang, H.; Rustgi, V.K.; Zhang, S.; Gevaert, O.; Metaxas, D.N. Genetic Mutation and Biological Pathway Prediction Based on Whole Slide Images in Breast Carcinoma Using Deep Learning. *NPJ Precis. Oncol.* **2021**, *5*, 87. [\[CrossRef\]](#)
142. Zheng, H.; Momeni, A.; Cedoz, P.L.; Vogel, H.; Gevaert, O. Whole Slide Images Reflect DNA Methylation Patterns of Human Tumors. *NPJ Genom. Med.* **2020**, *5*, 11. [\[CrossRef\]](#)
143. Carrillo-Perez, F.; Pizurica, M.; Ozawa, M.G.; Vogel, H.; West, R.B.; Kong, C.S.; Herrera, L.J.; Shen, J.; Gevaert, O. Synthetic Whole-Slide Image Tile Generation with Gene Expression Profile-Infused Deep Generative Models. *Cell Rep. Methods* **2023**, *3*, 100534. [\[CrossRef\]](#) [\[PubMed\]](#)
144. Zeng, Q.; Klein, C.; Caruso, S.; Maille, P.; Laleh, N.G.; Sommacale, D.; Laurent, A.; Amaddeo, G.; Gentien, D.; Rapinat, A.; et al. Artificial Intelligence Predicts Immune and Inflammatory Gene Signatures Directly from Hepatocellular Carcinoma Histology. *J. Hepatol.* **2022**, *77*, 116–127. [\[CrossRef\]](#)
145. Zhao, Y.; Alizadeh, E.; Taha, H.B.; Liu, Y.; Xu, M.; Mahoney, J.M.; Li, S. Inferring Single-Cell Spatial Gene Expression with Tissue Morphology via Explainable Deep Learning. *BioRxiv* **2024**. [\[CrossRef\]](#)
146. Elosua-Bayes, M.; Nieto, P.; Mereu, E.; Gut, I.; Heyn, H. SPOTlight: Seeded NMF Regression to Deconvolute Spatial Transcriptomics Spots with Single-Cell Transcriptomes. *Nucleic Acids Res.* **2021**, *49*, E50. [\[CrossRef\]](#) [\[PubMed\]](#)
147. Hao, M.; Luo, E.; Chen, Y.; Wu, Y.; Li, C.; Chen, S.; Gao, H.; Bian, H.; Gu, J.; Wei, L.; et al. STEM Enables Mapping of Single-Cell and Spatial Transcriptomics Data with Transfer Learning. *Commun. Biol.* **2024**, *7*, 56. [\[CrossRef\]](#)
148. Chen, H.; Lee, Y.J.; Ovando-Ricardez, J.A.; Rosas, L.; Rojas, M.; Mora, A.L.; Bar-Joseph, Z.; Lugo-Martinez, J. Recovering Single-Cell Expression Profiles from Spatial Transcriptomics with scResolve. *Cell Rep. Methods* **2024**, *4*, 100864. [\[CrossRef\]](#)
149. Li, X.; Xiao, C.; Qi, J.; Xue, W.; Xu, X.; Mu, Z.; Zhang, J.; Li, C.Y.; Ding, W. STellaris: A Web Server for Accurate Spatial Mapping of Single Cells Based on Spatial Transcriptomics Data. *Nucleic Acids Res.* **2023**, *51*, W560–W568. [\[CrossRef\]](#) [\[PubMed\]](#)
150. Hao, M.; Hua, K.; Zhang, X. SOMDE: A Scalable Method for Identifying Spatially Variable Genes with Self-Organizing Map. *Bioinformatics* **2021**, *37*, 4392–4398. [\[CrossRef\]](#) [\[PubMed\]](#)
151. Zhang, K.; Feng, W.; Wang, P. Identification of Spatially Variable Genes with Graph Cuts. *Nat. Commun.* **2022**, *13*, 5488. [\[CrossRef\]](#)
152. Xu, H.; Fu, H.; Long, Y.; Ang, K.S.; Sethi, R.; Chong, K.; Li, M.; Uddamvathanak, R.; Lee, H.K.; Ling, J.; et al. Unsupervised Spatially Embedded Deep Representation of Spatial Transcriptomics. *Genome Med.* **2024**, *16*, 12. [\[CrossRef\]](#)
153. Xu, Y.; McCord, R.P. CoSTA: Unsupervised Convolutional Neural Network Learning for Spatial Transcriptomics Analysis. *BMC Bioinform.* **2021**, *22*, 397. [\[CrossRef\]](#)
154. Dong, K.; Zhang, S. Deciphering Spatial Domains from Spatially Resolved Transcriptomics with an Adaptive Graph Attention Auto-Encoder. *Nat. Commun.* **2022**, *13*, 1739. [\[CrossRef\]](#)
155. Yuan, Y.; Bar-Joseph, Z. GCNG: Graph Convolutional Networks for Inferring Gene Interaction from Spatial Transcriptomics Data. *Genome Biol.* **2020**, *21*, 300. [\[CrossRef\]](#)
156. Li, Y.; Stanojevic, S.; Garmire, L.X. Emerging Artificial Intelligence Applications in Spatial Transcriptomics Analysis. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 2895–2908. [\[CrossRef\]](#) [\[PubMed\]](#)
157. Kaur, H.; Heiser, C.N.; McKinley, E.T.; Ventura-Antunes, L.; Harris, C.R.; Roland, J.T.; Farrow, M.A.; Selden, H.J.; Pingry, E.L.; Moore, J.F.; et al. Consensus Tissue Domain Detection in Spatial Omics Data Using Multiplex Image Labeling with Regional Morphology (MILWRM). *Commun. Biol.* **2024**, *7*, 1295. [\[CrossRef\]](#) [\[PubMed\]](#)
158. Hoseini, S.S.; Dewar, R. Empowering Healthcare Professionals with No-Code Artificial Intelligence Platforms for Model Development, a Practical Demonstration for Pathology. *Discoveries* **2024**, *12*, e182. [\[CrossRef\]](#) [\[PubMed\]](#)
159. Tagra, H.; Batra, P. Dentistry 4.0: A Whole New Paradigm. *Discoveries* **2021**, *4*, e19. [\[CrossRef\]](#)
160. Kaushik, R.; Rapaka, R. A Patient-Centered Perspectives and Future Directions in AI-Powered Teledentistry. *Discoveries* **2024**, *12*, e199. [\[CrossRef\]](#)
161. Edpuganti, S.; Shamim, A.; Gangolli, V.H.; Weerasekara, R.A.D.K.N.W.; Yellamilli, A. Artificial Intelligence in Cardiovascular Imaging: Current Landscape, Clinical Impact, and Future Directions. *Discoveries* **2025**, *13*, e211. [\[CrossRef\]](#)
162. Zhao, Y.; Bucur, O.; Irshad, H.; Chen, F.; Weins, A.; Stancu, A.L.; Oh, E.Y.; Distasio, M.; Torous, V.; Glass, B.; et al. Nanoscale Imaging of Clinical Specimens Using Pathology-Optimized Expansion Microscopy. *Nat. Biotechnol.* **2017**, *35*, 757–764. [\[CrossRef\]](#)
163. Chen, J.; Wang, Y.; Ko, J. Single-Cell and Spatially Resolved Omics: Advances and Limitations. *J. Pharm. Anal.* **2023**, *13*, 833–835. [\[CrossRef\]](#)

164. Fan, Q.; Wang, Y.; Cheng, J.; Pan, B.; Zang, X.; Liu, R.; Deng, Y. Single-Cell RNA-Seq Reveals T Cell Exhaustion and Immune Response Landscape in Osteosarcoma. *Front. Immunol.* **2024**, *15*, 1362970. [[CrossRef](#)]
165. Antcliffe, D.B.; Harte, E.; Hussain, H.; Jiménez, B.; Browning, C.; Gordon, A.C. Metabolic Septic Shock Sub-Phenotypes, Stability over Time and Association with Clinical Outcome. *Intensive Care Med.* **2025**, *51*, 529–541. [[CrossRef](#)]
166. Yang, J.; Ou, F.; Li, B.; Zeng, L.; Chen, Q.; Gan, H.; Yu, J.; Guo, Q.; Feng, J.; Zhang, J. Machine Learning Based Screening of Biomarkers Associated with Cell Death and Immunosuppression of Multiple Life Stages Sepsis Populations. *Sci. Rep.* **2025**, *15*, 30302. [[CrossRef](#)]
167. Guo, F.; Zhu, X.; Wu, Z.; Zhu, L.; Wu, J.; Zhang, F. Clinical Applications of Machine Learning in the Survival Prediction and Classification of Sepsis: Coagulation and Heparin Usage Matter. *J. Transl. Med.* **2022**, *20*, 265. [[CrossRef](#)]
168. Gui, Y.; He, X.; Yu, J.; Jing, J. Artificial Intelligence-Assisted Transcriptomic Analysis to Advance Cancer Immunotherapy. *J. Clin. Med.* **2023**, *12*, 1279. [[CrossRef](#)]
169. Ravindran, U.; Gunavathi, C. Deep Learning Assisted Cancer Disease Prediction from Gene Expression Data Using WT-GAN. *BMC Med. Inform. Decis. Mak.* **2024**, *24*, 311. [[CrossRef](#)]
170. Yan, G.; Mingyang, G.; Wei, S.; Hongping, L.; Liyuan, Q.; Ailan, L.; Xiaomei, K.; Huilan, Z.; Juanjuan, Z.; Yan, Q. Diagnosis and Typing of Leukemia Using a Single Peripheral Blood Cell through Deep Learning. *Cancer Sci.* **2025**, *116*, 533–543. [[CrossRef](#)] [[PubMed](#)]
171. Lähnemann, D.; Köster, J.; Szczurek, E.; McCarthy, D.J.; Hicks, S.C.; Robinson, M.D.; Vallejos, C.A.; Campbell, K.R.; Beerenwinkel, N.; Mahfouz, A.; et al. Eleven Grand Challenges in Single-Cell Data Science. *Genome Biol.* **2020**, *21*, 31. [[CrossRef](#)]
172. Choudhary, S.; Satija, R. Comparison and Evaluation of Statistical Error Models for ScRNA-Seq. *Genome Biol.* **2022**, *23*, 27. [[CrossRef](#)] [[PubMed](#)]
173. Hwang, H.; Jeon, H.; Yeo, N.; Baek, D. Big Data and Deep Learning for RNA Biology. *Exp. Mol. Med.* **2024**, *56*, 1293–1321. [[CrossRef](#)] [[PubMed](#)]
174. Bhandari, N.; Khare, S.; Walambe, R.; Kotecha, K. Comparison of Machine Learning and Deep Learning Techniques in Promoter Prediction across Diverse Species. *PeerJ Comput. Sci.* **2021**, *7*, e365. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.