High-impact discovery through gene expression and regulation research

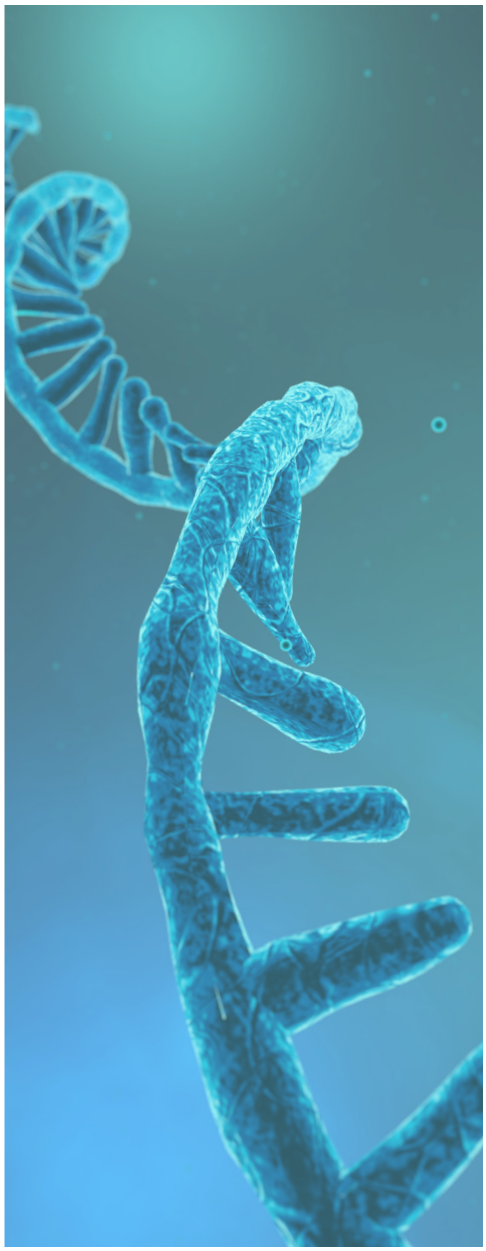
Advancing RNA sequencing, chromatin structure, and DNA methylation studies with the power of next-generation sequencing

illumina®

Table of contents

Introduction

The wondrous diversity of life is evident at every level, from complex ecosystems down to specialized cells and molecular networks. DNA may get most of the public's attention, but it is gene expression and regulation that orchestrate the dynamics of cell function and physiology. The 90% of the human genome once regarded as "junk DNA" is now appreciated for its role in controlling which genes are expressed, including where, when, and how much. The majority of variants identified in genome-wide association studies (GWAS) occur in noncoding regions of DNA, underscoring the significance of gene expression and regulation in the mechanisms of disease.[1] This eBook highlights discoveries powered by modern transcriptomics and epigenetics methods that are impacting our understanding of biology.



Next-generation sequencing accelerates transcriptomics research

The dynamic landscape of transcriptomics

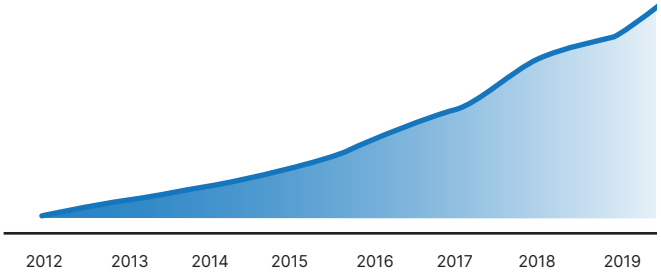
Our curiosity and quest for answers has always been the driving force for discovery. As our tools have evolved from the basic light microscope to high-throughput DNA sequencers, so has our understanding of the world around us. Next-generation sequencing (NGS) and NGS-based RNA sequencing (RNA-Seq) are technological advances driving scientists to push beyond the limits of traditional methods. As researchers seek to understand how the transcriptome shapes biology, RNA-Seq is one of the most significant and powerful tools in modern science.

A highly sensitive and accurate method for gene expression analysis, RNA-Seq reveals the full transcriptome, not just a few selected transcripts. RNA-Seq provides visibility into previously undetectable changes in gene expression and enables the characterization of multiple forms of noncoding RNA.[2,3] Researchers can use RNA-Seq to detect the fine architecture of the transcriptome, such as transcript isoforms, gene fusions, single nucleotide variants, and other features—without the limitation of prior knowledge.[3-5]

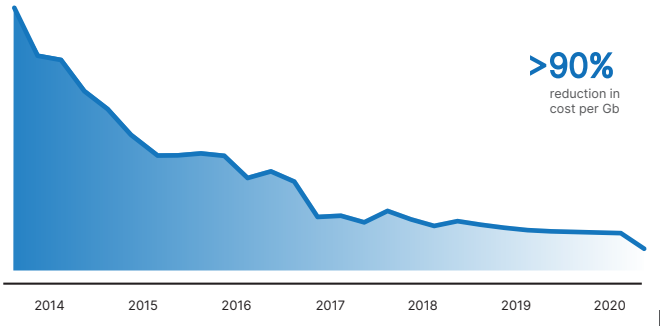
RNA-Seq has quickly emerged as the paramount approach to high-throughput transcriptome profiling

RNA-Seq drives high-impact research

With the true discovery power of unbiased RNA detection, RNA-Seq has quickly emerged as the paramount approach to high-throughput transcriptome profiling.[5-8] Grant funding and publication trends over the last decade demonstrate the rapid adoption and increasing impact of RNA-Seq–based research. Shrinking costs per gigabase (Gb) for NGS have made RNA-Seq accessible to more researchers.



The number of RNA-Seq publications grew by six-fold from 2012 to 2020.[9]



Sequencing costs per Gb decreased by over 90% from 2014 to 2020.[9]

Gene expression analysis with RNA-Seq is considered a vital tool to uncover the mechanisms of cancer and aid genetic disease research.[7,10] RNA-Seq also provides a view of noncoding transcripts and illuminates their role in complex disease.[11,12]

High-resolution methods like single-cell RNA-Seq and spatial transcriptomics continue to transform how the research community studies gene expression.[13] These RNA-Seq advances have enabled researchers to examine the details of cancer, development, and infectious disease at the single-cell level with tissue context.[14-22] Studies have combined single-cell and spatial RNA-Seq to map the architecture of skin cancer,[23] characterize human intestinal development,[24] and track COVID-19 pathology.[14,15] Many more high-impact publications will be powered by RNA-Seq.

The advantages of RNA-Seq vs alternative technologies

Hypothesis-free study design and higher discovery power

RNA-Seq is a powerful sequencing-based method that captures a full and informative spectrum of gene expression data.[25] RNA-Seq provides many significant advantages over legacy technologies such as reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR) and gene expression microarrays. Because RNA-Seq does not require predesigned probes, the data sets are unbiased, allowing for hypothesis-free experimental design.[3,4,6] This type of NGS analysis is an effective tool for transcript and variant discovery studies, which are not possible using traditional methods that require known targets. Similarly, RNA-Seq is considered an accurate alternative to immunohistochemistry (IHC) for measuring biomarker genes without needing specific antibodies.[8]

Wider range of detection and higher sensitivity

RNA-Seq quantifies individual sequence reads aligned to a reference sequence to produce discreet read counts.[3] By increasing or decreasing the number of sequencing reads (coverage depth), researchers can fine-tune the sensitivity of an experiment to accommodate different study objectives. The quantitative nature of this process and the ability to control coverage levels supports an extremely broad dynamic range, with absolute rather than relative expression values.[2-4] RNA-Seq offers finer coverage of the transcriptome and lower technical variability than microarrays with the ability to detect a higher percentage of differentially expressed genes, especially genes with low abundance.[26,27] RNA-Seq also shows high agreement with gold-standard RT-qPCR, but on a significantly larger scale.[6]

Advantages of RNA-Seq

- Enables true discovery power across the transcriptome
- Provides sensitive, accurate measurement of gene expression at the transcript level
- Generates both qualitative and quantitative data
- Captures splice junctions, fusions, coding, and multiple forms of noncoding RNA such as siRNA, miRNA, snoRNA, tRNA, and lncRNA
- Covers an extremely broad dynamic range
- Delivers excellent performance with degraded RNA such as FFPE tissue samples
- Maintains and tracks strand-specific information in the data
- Scales for large studies with high sample throughput

RNA-Seq vs alternative technologies

RNA-Seq offers higher discovery power and wider dynamic range for transcriptome studies.

Compare to qPCR

Compare to microarrays

Perspectives on RNA-Seq

See how a lab studying host-pathogen responses gained deeper insights through adopting RNA-Seq.

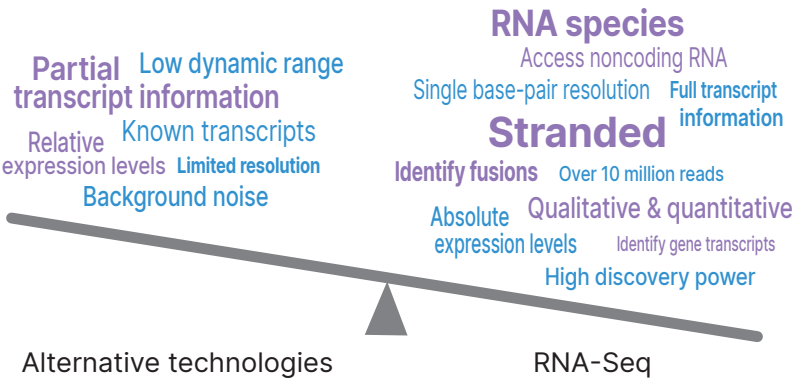
Watch video

Detection of alternative splice sites, novel isoforms, and noncoding RNA

A major advantage of RNA-Seq is that it provides a rich view of transcriptome activity well beyond basic abundance measurements. With RNA-Seq, researchers can detect alternative splice sites, characterize novel gene fusions, and identify allele-specific expression—all in a single experiment.[2,3,5]

Certain library preparation methods for RNA-Seq also allow researchers to detect and sequence multiple forms of noncoding RNA, including small interfering RNA (siRNA), microRNA (miRNA), small nucleolar RNA (snoRNA), transfer RNA (tRNA), and long noncoding RNA (lncRNA).[3,5,11] The ability to sequence small RNA fragments enables high-quality data generation with degraded RNA samples, such as formalin-fixed, paraffin-embedded (FFPE) tissues.[28] The sequencing data can also be reanalyzed as novel features of the transcriptome are discovered over time, without rerunning the experiment.

In summary, RNA-Seq offers many advantages over alternative experimental approaches. It provides a unique combination of transcriptome-wide coverage, broad dynamic range, and high sensitivity that can empower researchers to investigate and understand the molecular mechanisms of normal development and disease.



Weigh your options—RNA-Seq provides many advantages over alternative technologies like RT-qPCR, gene expression arrays, or immunohistochemistry.

Gene expression studies with RNA-Seq

Differential expression for complex disease research

To understand normal cell development and disease, researchers frequently investigate differential expression over time, in specific tissues, or in response to varying conditions. Changes in gene expression levels can reflect the effects of genetic variation, environment, and disease processes or treatments.[25] RNA-Seq delivers a comprehensive picture of gene expression across the full transcriptome at a specific time, and shows exceptional performance in profiling genes with low expression levels, increasing the discovery power of these studies.[29]

Data from RNA-Seq experiments can offer insight into the gene networks and pathways involved in complex disease and cell biology mechanisms.[30,31] For example, transcriptomic analysis is helping researchers compare brain regions with different pathology to identify meaningful gene expression changes in Alzheimer's disease (AD).[30,32] Differential expression profiling is also revealing the pathogenesis of heart failure and identifying gene signatures to detect heart disease.[25,33-36]

More applications for RNA-Seq

Profile drug response RNA biomarkers for drug discovery and development research.

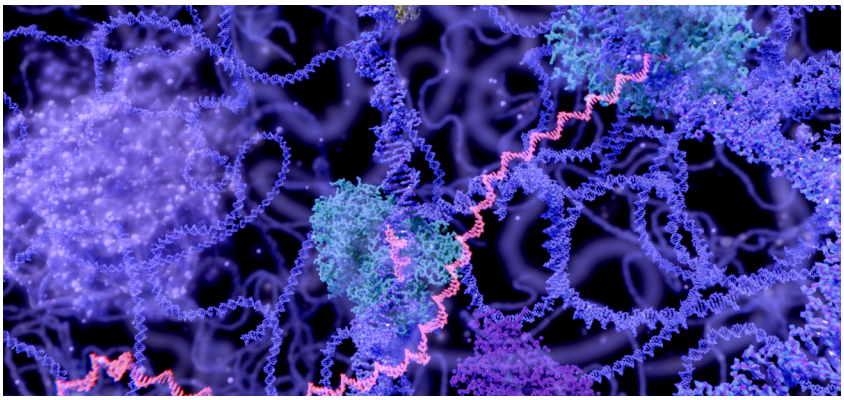
Learn more

Analyze viral and bacterial transcriptome signatures with microbial RNA-Seq.

Learn more

Biomarker and drug target discovery with RNA-Seq

Gene expression studies are laying the groundwork for advances in precision medicine by identifying potential therapeutic biomarkers and drug targets.[29] RNA-Seq has shown that noncoding transcripts, including lncRNAs and miRNAs, play important roles in cancer and kidney disease.[37] Whole-transcriptome sequencing allowed clinical researchers to identify lncRNAs, miRNAs, and mRNAs associated with metastasis of breast cancer to the brain.[38] Systematic tissue-specific gene expression studies also helped establish over 800 lncRNA gene-trait associations with inflammatory bowel disease, type 1 and type 2 diabetes, and other complex diseases.[12]



RNA-Seq for cancer research

RNA-Seq is a critical tool for direct measurement of the functional consequence of mutations. Despite the average cancer containing about 46 mutations, only 5 to 8 are necessary for initiation.[39] Genomic profiling alone is insufficient to differentiate these driver mutations from passenger mutations, or those which do not influence cancer initiation or progression. Measurement of gene expression patterns and mutation consequences using RNA-Seq enables large-scale, unbiased differentiation of factors crucial for cancer progression, resulting in more thorough and accurate cancer modeling.

Gene fusion detection is particularly significant for cancer research, as 20% of all human tumors carry translocations and gene fusions.[40] The majority of gene fusions have a significant impact on tumorigenesis and a strong association with morphological phenotype, making them useful as potential diagnostic and prognostic markers.[40] These phenomena were demonstrated in a clinical study that used whole-transcriptome sequencing to detect novel fusion genes and classify pathogenic drivers in patients with acute leukemia.[41]

Clinical and translational researchers are using RNA-Seq to identify clinically relevant mechanisms of disease, discover predictive biomarkers, and identify responsive subpopulations for cancer and complex disease.[42-45] For example, transcriptomic analysis is helping to classify tumor microenvironments that correlate with patient response to immunotherapy.[46,47] One such study revealed differentially expressed genes, including *STAT1*, *TLR3*, and *IL10*, as putative drug targets. Combination therapy targeting these three pathways increased the immune checkpoint blockade response rate from 10% up to 80%, demonstrating the power of gene expression profiling to inform patient treatment approaches.[46]

Detecting changes in the cancer transcriptome

RNA-Seq provides functional information about cancer gene expression and the gene fusions that drive tumor progression.

Learn more

RNA-Seq for genetic disease research

RNA-Seq offers a complementary approach to GWAS for genetic disease research that increases diagnostic yield.[48-50] Measuring expression abundance in specific tissues can reveal the functional impact of pathogenic mutations and identify which genes mediate the genotype's effect on phenotype.[1,48-50] RNA-Seq can also validate computational predictions of splicing and increase confidence in the reclassification of variants of unknown significance (VUS).

Analysis of expression quantitative trait loci with RNA-Seq

Although GWAS have identified tens of thousands of genetic variants associated with many complex traits, the causal variants and genes impacted by these GWAS-identified loci remain largely unknown.[1,51] Many studies have shown that GWAS risk variants co-localize with genes that regulate expression.[1,52] These genes, known as expression quantitative trait loci (eQTLs), suggest that regulation is an important molecular mechanism used by GWAS risk variants, the majority of which lie in noncoding regions of the genome. There are several methods for integrating GWAS and eQTL data, including transcriptome-wide association studies (TWAS). With TWAS, gene expression levels for GWAS samples are predicted and then tested for association between the predicted expression and traits.[52]

Researchers using TWAS to study obesity-related complex traits demonstrated a putative causal relationship between body mass index and triglyceride levels. The study integrated gene expression measurements from 45 expression panels with GWAS data and identified 1196 genes whose expression are associated with these traits.[53] By leveraging the growing wealth of publicly available GWAS, gene expression, and phenotypic data, researchers are shedding light on the role of gene expression in complex diseases.

Uncover gene targets and pathways tied to disease

Differential expression analysis can help determine the functional effects of gene variants.

Learn more

Advances in RNA-Seq techniques

Single-cell RNA-Seq

Considerations, trends, and the future of single-cell sequencing

Single-cell transcriptomics is improving researchers' understanding of biological systems in health and disease.

Watch video

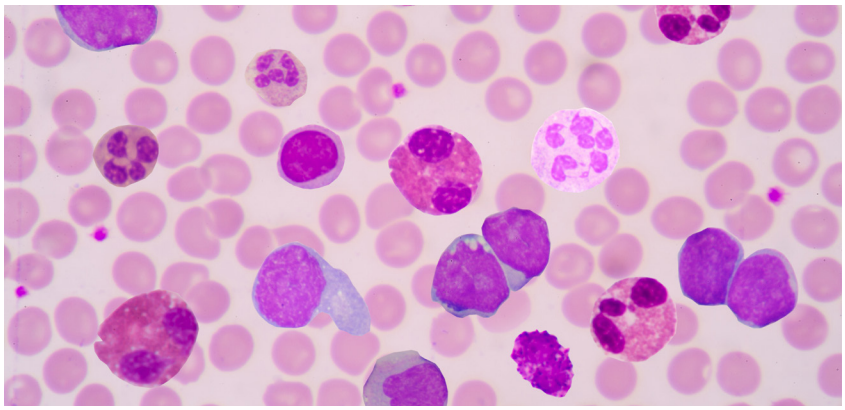
Learn how one of the pioneers in single-cell genomics is using Illumina technology.

Read article

Single-cell sequencing is a popular approach used to characterize hundreds to tens of thousands of individual cells from a tissue. This method reveals cellular heterogeneity and provides a more comprehensive understanding of tissue composition. Significant advances in the area of single-cell characterization include technologies for cell isolation and new methods and applications for single-cell sequencing. These advances have stimulated the launch of accessible commercial solutions for every step of the single-cell sequencing workflow, from tissue preparation through data analysis.

Single-cell RNA-Seq (scRNA-Seq) has become a powerful tool in immunology, cancer research, and developmental biology. As part of the Human Cell Atlas, a large-scale effort to map human development, researchers used single-cell combinatorial indexing to profile the transcriptomes of ~2 million cells derived from 61 embryos staged between 9.5 and 13.5 days of gestation in a single experiment.[54] Cell atlas studies often have implications in genetic disease research, such as cystic fibrosis. scRNA-Seq of human bronchial epithelial cells helped uncover a rare cell type, pulmonary ionocytes, that accounts for the majority of *CFTR* gene expression in the lungs.[55]

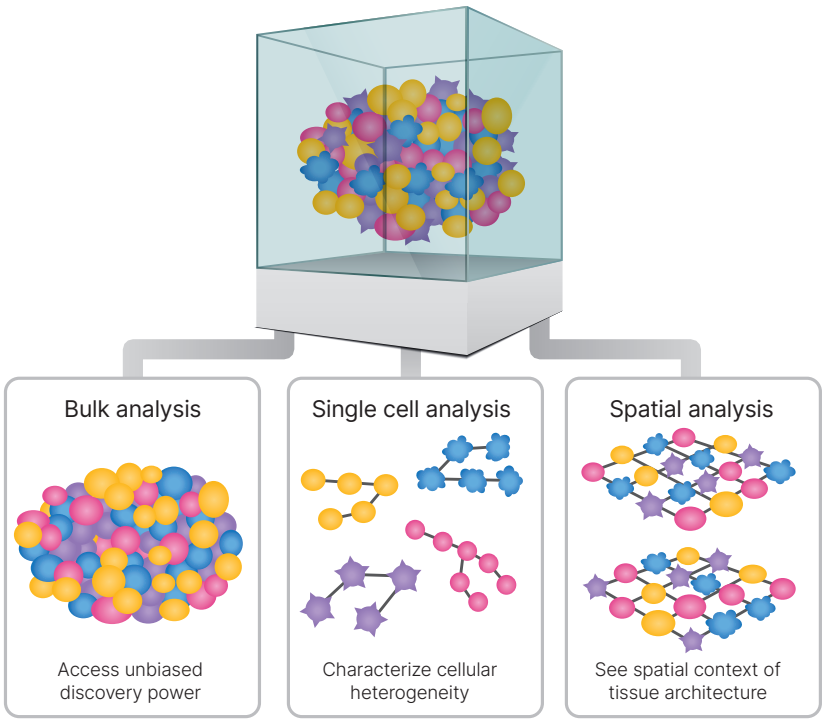
Cancer researchers use scRNA-Seq to better understand cancer biology, as traditional bulk RNA-Seq does not address the heterogeneity within and between tumors. scRNA-Seq has aided the development of targeted therapy and immunotherapy treatments. scRNA-Seq has also been widely used to understand COVID-19 infection and disease with dramatic transcription changes observed in virus-positive cells. Researchers often use scRNA-Seq in conjunction with cell-surface protein expression and immune repertoire sequencing to characterize the inflammatory response.



Spatial RNA-Seq

Typical NGS methods using dissociated samples can lose key spatial information present *in vivo*. Traditionally, IHC and *in situ* hybridization have been the tools of choice to reveal spatial gene expression in tissue sections. But the throughput of these procedures is limited, analyzing only a few genes at a time.

By combining high-throughput imaging and sequencing technologies, spatial RNA-Seq provides a previously inaccessible view of the full transcriptome in morphological context. Spatial RNA-Seq methods that retain the precise location of biological molecules in tissue samples can further our understanding of mechanisms in health and disease.



High-resolution, high-throughput spatial transcriptomics

Learn how to leverage Visium Spatial Gene Expression from 10x Genomics and Illumina sequencing systems for transcriptional profiling of entire tissue sections.

Read Technical Note

See how the Nanostring GeoMX Digital Spatial Profiler and Illumina sequencing can reveal the tissue architecture of kidney disease.

Read Application Note

See how scientists are using scRNA-Seq and spatial RNA-Seq in disease research

Cancer research

Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response[19]
Cell atlases generated using scRNA-Seq helped predict cells of origin for renal cell carcinoma (RCC) subtypes. This study highlights the role of the tumor microenvironment in influencing RCC biology and response to therapy.

Single-cell transcriptomic landscape reveals the differences in cell differentiation and immune microenvironment of papillary thyroid carcinoma between genders[20]
Researchers used scRNA-Seq to map single-cell gene expression in thyroid cancer and identify differences in cell clusters between male and female patients.

Inter- and intra-tumor heterogeneity of metastatic prostate cancer determined by digital spatial gene expression profiling[22]
Metastatic prostate cancer (mPC) comprises a spectrum of diverse phenotypes. However, the extent of inter- and intra-tumor heterogeneity is not established. Researchers used bulk RNA-Seq and spatial RNA-Seq to quantitate transcript and protein abundance in spatially distinct regions of mPCs from 53 FFPE samples.

Complex disease research

High-throughput single-cell functional elucidation of neurodevelopmental disease-associated genes reveals convergent mechanisms altering neuronal differentiation[21]
Functional genomics studies used scRNA-Seq to generate mechanistic insights for a set of 13 autism spectrum disorder (ASD)–associated genes. Two functionally convergent modules of ASD genes were identified: one that delays neuron differentiation and one that accelerates it.

Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex[10]
This study defined the spatial topography of gene expression in the six-layered human dorsolateral prefrontal cortex. By integrating neuropsychiatric disorder gene sets, the authors showed differential layer-enriched expression of genes associated with schizophrenia and autism spectrum disorder.

COVID-19 research

B cell genomics behind cross-neutralization of SARS-CoV-2 variants and SARS-CoV[16]
This study used scRNA-Seq to investigate SARS-CoV-2 spike-specific B cell responses in 14 subjects who had recovered from COVID-19. These results help to design therapeutic treatments against coronaviruses.

SARS-CoV-2 infection of the oral cavity and saliva[17]
Oral cavity scRNA-Seq revealed a positive correlation between salivary viral load and taste loss; and demonstrated persistent salivary antibody responses to SARS-CoV-2 nucleocapsid and spike proteins.

Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions[18]
Researchers used spatial RNA-Seq to characterize the distribution of a set of greater than 1800 genes across tissues from four postmortem COVID-19 patients; compared to three normal lung donors. The autopsy tissue data revealed distinct *ACE2* expression loci, with macrophage and neutrophil infiltration in the lungs.

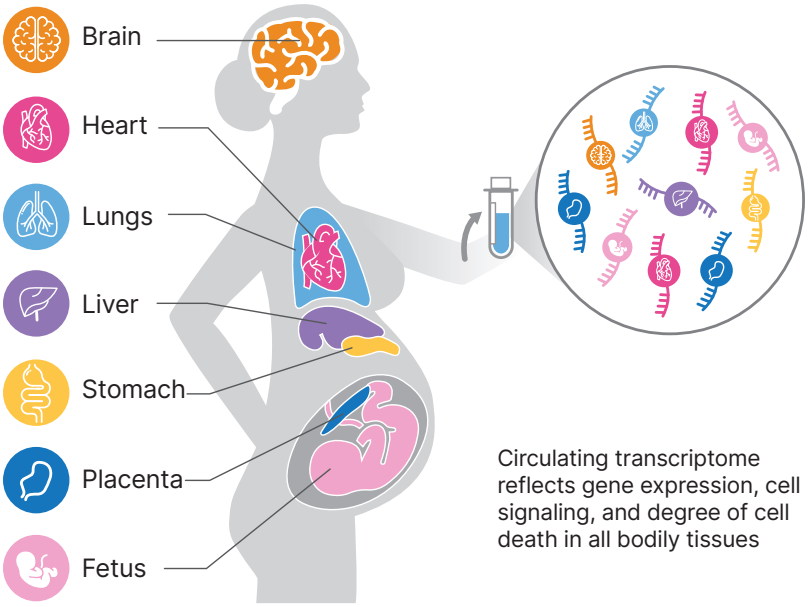
Circulating cell-free RNA analysis

Nucleic acids in the bloodstream can serve as biomarkers to monitor human health.[56] Circulating cell-free RNA (cfRNA) is released by many tissues into the circulation via cellular processes of apoptosis, microvesicle shedding, and exosome signaling. Because of these diverse origins, cfRNA measurements reflect tissue-specific changes in gene expression, intercellular signaling, and the degree of cell death occurring within different tissues throughout the body.[57] Researchers are evaluating the use of cfRNA as a noninvasive biomarker for disease surveillance and treatment monitoring.[58-62]

Improved detection of circulating transcripts

RNA-Seq library preparation using a tagmentation with enrichment protocol enables robust sequencing for low-input samples like circulating RNA.

Read Application Note



Simultaneous bulk protein and gene expression profiling

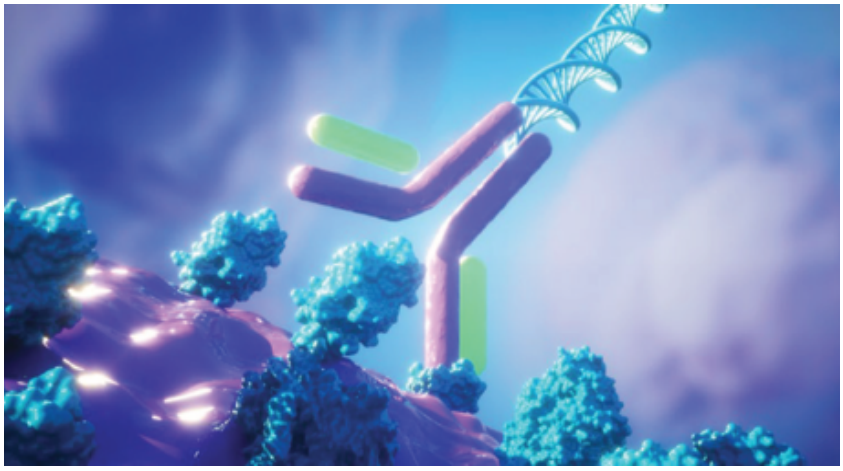
BEN-Seq method uses Illumina NGS and oligo-conjugated antibodies for easy quantification of multiple protein targets.

Read Application Note

Simultaneous protein detection with sequencing

RNA-Seq offers unparalleled discovery power to interrogate the transcriptome without prior knowledge. Incorporating protein detection with RNA-Seq can tie new discoveries back to known canonical markers and historical clinical outcomes. Combining transcriptome and protein detection is especially useful when characterizing the immune repertoire at the single-cell level.[63]

Antibodies tagged with oligonucleotide barcodes enable analysis of cell surface proteins with results read by sequencing, which scales to a much higher number of parameters than flow cytometry or mass cytometry. Methods like cellular indexing of transcriptomes and epitopes by sequencing (CITE-Seq) combine single-cell RNA-Seq with cell surface protein analysis.[64,65] Bulk epitope and nucleic acid sequencing (BEN-Seq) is performed at the bulk level.

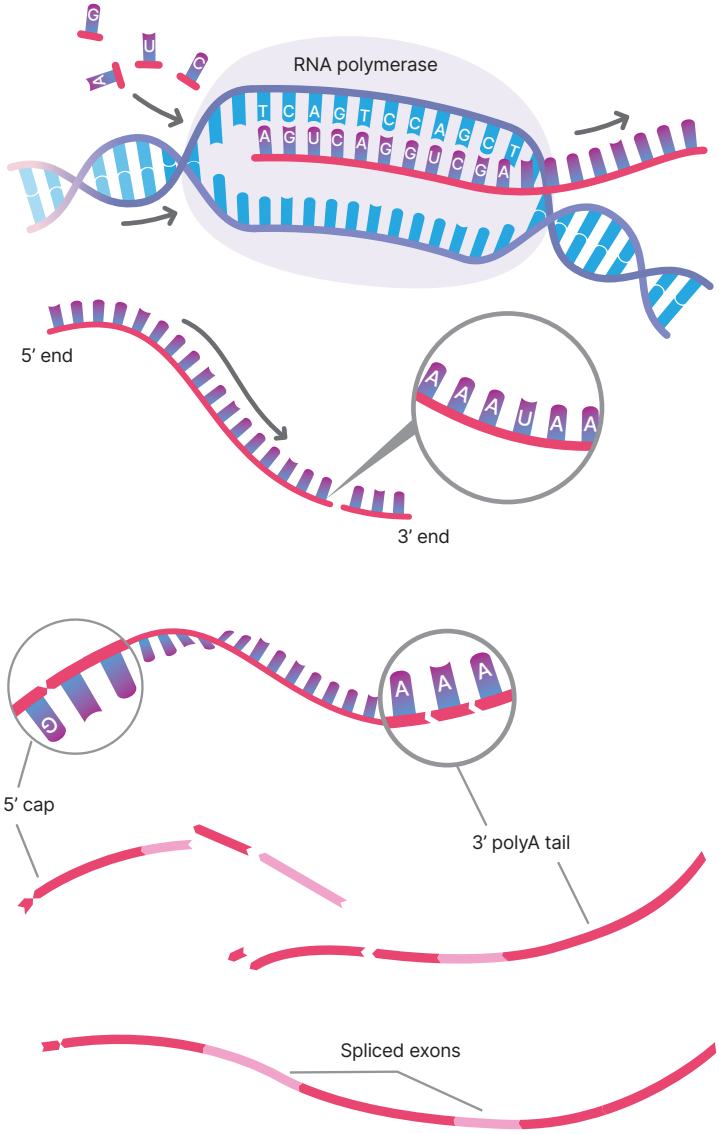


Protein detection with sequencing—DNA-tagged antibodies allow for detection of cell surface proteins with results read by sequencing.

Illumina workflows for gene expression studies

RNA-Seq utilizes well-established and easy-to-execute protocols, validated by thousands of publications. Illumina offers fully integrated, RNA-to-data workflows spanning from initial library preparation to final data analysis. Illumina library prep kits are available for a wide range of RNA-Seq applications including total RNA-Seq, mRNA-Seq, small RNA-Seq, sequencing low-quality samples, and more.

Download RNA-Seq Workflows Guide



Total RNA-Seq

Ribosomal RNA (rRNA) can account for 80% of transcripts. Depletion-based total RNA library preparations remove these transcripts that are not of interest and allow for sequencing whole transcriptomes, including noncoding RNA.[66]

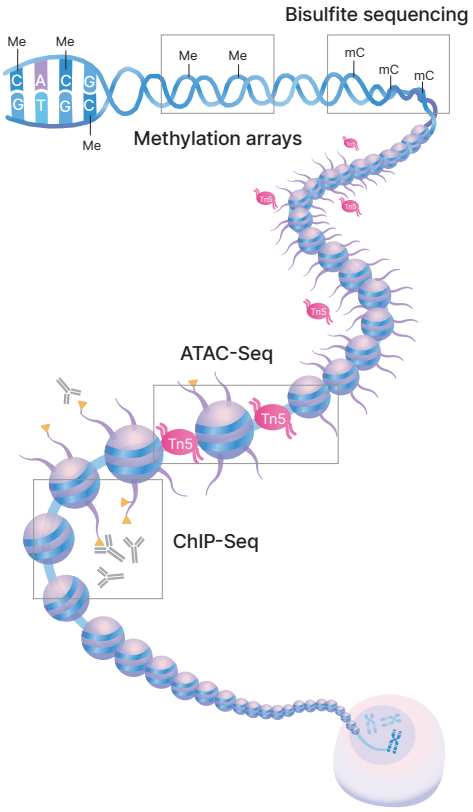
mRNA-Seq

PolyA capture–based mRNA library preparations interrogate mRNA with a 3' polyA tail. Because an intact polyA tail is needed, it is not ideal for degraded samples.[67]

Enrichment-based RNA-Seq

Hybridization-capture–based library preparations utilize a probe-based approach to target transcripts of interest and enable sensitive detection of splicing events and fusions, even in degraded samples.[28]

Gene regulation studies and DNA methylation



What is gene regulation and why is it important?

The regulation of gene expression is a biological process that controls the temporal and spatial expression of gene products, including mRNA and noncoding RNA transcripts. A wide range of mechanisms are used to increase or decrease gene expression, including binding of regulatory proteins to DNA motifs, binding of RNA polymerase to regulatory elements, and modulation of histone chromatin structure.[68] Determining how protein–DNA interactions regulate gene expression is essential for fully understanding many biological processes and disease states.[69] Chromatin immunoprecipitation sequencing (ChIP-Seq) and its variations (CUT&RUN, CUT&Tag) can leverage NGS to efficiently determine the distribution and abundance of DNA-bound protein targets across the genome at base-pair level resolution.[70-74] More advanced epigenetic NGS techniques include chromatin conformation capture (Hi-C, HiChIP) and assay for transposase-accessible chromatin (ATAC-Seq).[73-80]

DNA methylation of cytosine-guanine dinucleotides (CpGs) also plays an important role in connecting the workings of the genetic code to changing environmental factors. It allows cells to acquire and maintain a specialized state and suppresses the expression of viral and nonhost DNA elements. Aberrant DNA methylation and its impact on gene expression have been implicated in many disease processes, including cancer, neurological disorders, aging, and development.[81,82] High-throughput technologies, such as whole-genome bisulfite sequencing (WGBS), targeted bisulfite sequencing, and methylation microarrays, are powerful tools for investigating the dynamic state of DNA methylation across the genome. This epigenetic information is also highly complementary to DNA sequencing, genotyping, gene expression, and other forms of integrated genomic analysis.

Chromatin analysis applications

The chromatin landscape sets the context for the dynamic processes of gene regulation. Chromatin features—including nucleosomes and their modifications, bound transcription factors (TF), and three-dimensional (3D) conformation—mark regions of gene activation and silencing that differ between cell types and change during development and disease.[72] ATAC-Seq is a versatile method to map open chromatin across the genome and track cis-regulatory elements.[74,79,80] Researchers use a hyperactive transposase enzyme that cuts and adds sequencing adapters into exposed DNA as a tool to survey chromatin accessibility. ATAC-Seq can also be used at the single-nucleus level to study gene regulation in heterogeneous cell populations.[83] Chromatin accessibility is shedding light on gene regulation mechanisms in cancer, neuropsychiatric disorders, and other complex diseases.[83-85] Cleavage under targets and tagmentation (CUT&Tag), a modification to ChIP-Seq, is another method that uses transposase to improve the resolution and efficiency of chromatin analysis.[71-73]

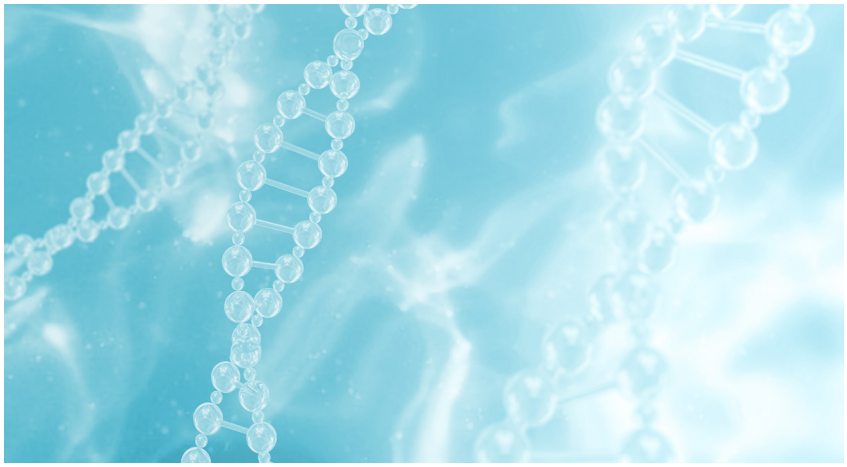
Unify single-cell gene expression and chromatin accessibility

Enable RNA-Seq and ATAC-Seq measurements from the same single cells in the same assay.

Read Technical Note

Methylation sequencing and microarray applications

DNA methylation is recognized as a key factor for both normal physiology and the development of complex disease. Methylation status has been shown to play a role in obesity, cancer, and other complex diseases.[86-88] External factors, such as the environment, diet, and exercise, can alter DNA methylation patterns to mediate disease progression.[89] Depending on the goals and size of the study, microarray analysis, targeted methylation sequencing, and WGBS can be used to study the effects of DNA methylation on human health.



Bisulfite sequencing

WGBS has enabled the genome-wide mapping of methylation patterns in normal cells compared to cells affected by cancer, cardiovascular disease, and obesity.[87,90,91] During library preparation, bisulfite conversion changes unmethylated cytosines to uracil. Converted bases are identified as thymine in the sequencing data and read counts are used to determine the percentage of methylated cytosines. NGS-based bisulfite sequencing enables researchers to discover the methylation patterns of CpG regions across the entire genome, at single-base resolution.

With targeted methylation sequencing, bisulfite conversion is followed by PCR amplification of specific regions of interest and sequencing. This cost-effective method allows for higher sequencing depth and produces more manageable data sets and faster workflows compared to WGBS.[92,93] Targeted methylation sequencing is ideal for hypothesis-testing studies of target regions of interest as well as confirmation of regions identified in GWAS studies.[93]

Whole-genome methylation mapping of normal and tumor genomes has confirmed that nearly all cancer types show tens to hundreds of genes with abnormal gains in DNA methylation.[87] The Circulating Cell-Free DNA Genome Atlas study has demonstrated the use of targeted methylation sequencing of cell-free DNA as a noninvasive early-detection test for over 50 cancer types.[94,95]

See how scientists are using bisulfite sequencing for complex disease research

Whole-genome DNA hyper-methylation in iPSC-derived dopaminergic neurons from Parkinson's disease patients[96]
This study describes WGBS of induced pluripotent stem cells (iPSC)-derived dopaminergic neurons from Parkinson's Disease (PD) patients. The results indicated that dopaminergic neurons from PD subjects exhibit global DNA hyper-methylation changes.

Rheumatoid arthritis-relevant DNA methylation changes identified in ACPA-positive asymptomatic individuals using methylome capture sequencing[97]
Anti-citrullinated protein antibodies (ACPA) are a key serological marker of rheumatoid arthritis (RA) risk. To investigate the differentially methylated regions between ACPA-positive vs ACPA-negative subjects, researchers performed targeted bisulfite sequencing of 5 million CpGs located in regulatory or hypomethylated regions of DNA.

Peripheral blood DNA methylation differences in twin pairs discordant for Alzheimer's disease[98]
In this study of 120 Swedish twin pairs discordant for AD, targeted methylation sequencing revealed 11 differentially methylated regions associated with AD. Further analysis of the Swedish twin pairs demonstrated that methylation in the *ADARB2* gene is influenced by gender, age, zygosity, and smoking.

Methylation arrays for disease association studies

Advances in methylation array technology are making a big impact on the field of epigenetics and enabling researchers to perform cost-effective epigenome-wide association studies (EWAS) with large sample cohorts.[98-100] Researchers have found significant connections between behavioral epigenetic factors, such as diet, smoking, and exercise, and their effects on complex conditions, including obesity, heart disease, and mental illness.[86,101,102] Methylation arrays have also powered research to identify blood-based epigenetic signatures tied to coronary heart disease.[103]

Methylation array data can be correlated with RNA-Seq and methylation sequencing results.[99,100] For example, researchers compared existing sequence-based methylation data with findings from three large-scale array-based schizophrenia methylation studies that interrogated up to ~450,000 CpGs. The integrated data identified 22 highly significant loci and 852 suggestively significant loci associated with schizophrenia.[100]

Methylation arrays for cancer classification

Methylation arrays are also used to define tumor classes and subclasses for many types of cancer.[104] For central nervous system (CNS) and sarcoma tumors, translational researchers used DNA methylation signatures to build classifiers that can help determine the tissue-of-origin of ambiguous tumor entities.[105,106] Also, researchers have demonstrated the ability to dissect the tumor inflammatory microenvironment using methylation signatures measured by arrays.[107]

Multiomics applications

Multiomics – a holistic view of biology

Advances in genomic technologies are giving researchers the tools to access more molecular data than ever before—enabling transcriptomics, epigenetics, proteomics, and beyond. Find timely examples from the scientific literature of how multiomics can fuel unique discovery power for deeper biological insights.

Download Multiomics eBook

To see the full picture of biology, scientists are increasingly turning towards a multiomics approach that integrates genomics, transcriptomics, epigenetics, and proteomics methods. Each modality is a piece of the puzzle offering important insights into the details of biological and disease mechanisms. Integrating these complementary metrics into multiomic data sets brings a more comprehensive picture of cellular phenotypes and helps pull more high-quality information from each sample.

Multiomics has enabled significant discoveries related to kidney disease.[1,108,109] One study integrated genotype, gene expression, alternative splicing, and methylation profiles of over 400 human kidneys to uncover genetic mechanisms of hypertension.[108] Other researchers used TWAS with single-cell chromatin and gene expression analysis to identify *DACH1* as a kidney disease risk gene.[1] A third study combined single-nucleus ATAC-Seq with scRNA-Seq to generate paired, cell-type-specific chromatin accessibility and transcriptional profiles of the adult human kidney.[109] This multiomics approach improved the ability to detect unique cell states within the kidney.

See how scientists are combining RNA-Seq and epigenetics methods to enhance their research

c-Jun overexpression in CAR T cells induces exhaustion resistance[110]
This study used overexpression of a T cell activating TF to examine mechanisms of exhaustion in chimeric antigen receptor (CAR) T cells. The researchers used RNA-Seq (bulk and single-cell), ChIP-Seq, and ATAC-Seq to track effects of the overexpression experiment.

Cell type-specific epigenetic links to schizophrenia risk in the brain[111]
To evaluate cell-type specific epigenetic variation in schizophrenia, researchers assayed neurons and oligodendrocytes with WGBS and whole-transcriptome sequencing.

Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos[112]
These researchers characterized preimplantation mouse embryos using a single-cell multiomics sequencing approach to profile transcriptomes, methylation, and chromatin accessibility in parallel in the same individual cell.

An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease[113]
This multiomics study examined epigenetic dysregulation in early AD. Total RNA-Seq and ChIP-Seq comparing postmortem brain tissue from AD patients and healthy controls revealed a reconfiguration of the epigenomic landscape with early AD.

Tobacco smoking induces changes in true DNA methylation, hydroxymethylation and gene expression in bronchoalveolar lavage cells[114]
Researchers interrogated changes in DNA methylation and gene expression in healthy smokers using methylation arrays and RNA-Seq. Integrated analysis revealed alteration of genes involved in migration, signaling, and inflammatory response of immune cells.

Summary

Advances in genomics and transcriptomics have led to an improved understanding of complex diseases, cancer biology, and environmental impacts on human health. NGS capabilities have shifted the scope of transcriptomics from the interrogation of a few genes at a time to profiling genome-wide expression levels in a single experiment. EWAS support these findings with an additional layer of information, indicating methylation status and the genetic response to environmental cues. Indeed, one of the most exciting aspects of the genomics revolution is the ability to weave transcriptional, epigenetic, and genetic studies into an integrated view at an unprecedented pace and scale.

While much progress has been made since the advent of NGS, so much more remains to be explored and discovered. Combining a broad library prep portfolio, high-quality data, and user-friendly analysis apps, Illumina RNA-Seq solutions empower researchers to investigate the molecular mechanisms of human health and disease. Since the introduction of the first sequencing system in 2006, Illumina has been committed to accelerating the pace of research through continuous innovation. Together, in collaboration with scientists from around the world, Illumina works to bring the power of NGS toward a deeper understanding of human biology and toward the promise of advanced precision medicine for future generations.

Acronyms

3D: three dimensional

ACPA: Anti-citrullinated protein antibodies

AD: Alzheimer's Disease

ASD: Autism spectrum disorder

ATAC-Seq: assay for transposase-accessible chromatin with sequencing

BEN-Seq: bulk epitope and nucleic acid sequencing

CAR-T: chimeric antigen receptor T cells

cfRNA: circulating cell-free RNA

ChIP-Seq: chromatin immunoprecipitation sequencing

CITE-Seq: cellular indexing of transcriptomes and epitopes by sequencing

CNS: central nervous system

COVID-19: coronavirus disease of 2019

CpGs: cytosine-guanine dinucleotides

CUT&RUN: cleavage under targets and release using nuclease

CUT&Tag: cleavage under targets and tagmentation

EWAS: epigenome-wide association studies

eQTL: expression quantitative trait loci

FFPE: formalin-fixed, paraffin-embedded

GWAS: genome-wide association studies

Hi-C: high-throughput chromatin conformation capture

HiChIP: Hi-C with ChIP-Seq

IHC: immunohistochemistry

iPSC: induced pluripotent stem cells

lncRNA: long non-coding RNA

mPC: metastatic prostate cancer

miRNA: micro RNA

NGS: next-generation sequencing

PD: Parkinson's Disease

scRNA-Seq: single-cell RNA sequencing

TF: transcription factor

Glossary

chromatin: The complex of DNA and protein that organizes and condenses genetic material within a nucleus.

cis-regulatory elements: Regions of noncoding DNA sequence that regulate transcription of genes. Examples include promoters, enhancers, and silencers.

coverage depth: The average number of sequenced bases that align to each base of the reference DNA. For example, a whole genome sequenced at 30× coverage means that, on average, each base in the genome was sequenced 30 times.

CpG site: A cytosine-guanine dinucleotide in DNA, separated by a phosphate. The p represents the phosphate group in the DNA backbone, indicating the 5'-3' directionality of the site. The C is always 5' of the 3'. CpG sites are locations where DNA may be methylated. CpG sites are often found in large groups (called CpG islands) in promoters. The methylation status of CpG islands often correlates with gene expression levels.

discovery power: In genomics, the ability to identify novel variants.

epigenetics: The study of how modifications in gene regulation can affect phenotypes without changing DNA sequence.

gene expression: The process by which instructions encoded in DNA are turned into RNA and proteins. Gene expression is a process susceptible to variation and evolutionary selection, due to variations in the timing, location, and amount of gene expression.

genome-wide association study (GWAS): A study that examines genetic variation across many genomes and compares the DNA of people with a phenotypic trait, such as disease, to the DNA of control subjects. The goal of this type of study is to determine the causal variant for a given phenotypic trait, or a variant in linkage disequilibrium with a causal variant. To do this, researchers investigate the association between certain genetic variants (or alleles) and variations of phenotype. It is assumed that by using a large enough population, non-causative variants will disappear into the noise, above which the signal of true causal variants will be detected.

methylation: Occurs when a methyl group (carbon and hydrogen) is added to DNA, often at CpG sites.

r**ead:** In general terms, a sequence "read" refers to the data string of A, T, C, and G bases corresponding to the sample DNA. With Illumina technology, millions of reads are generated in a single sequencing run. In more specific terms, each cluster on the flow cell produces a single sequencing read. For example, 10,000 clusters on the flow cell would produce 10,000 single reads and 20,000 paired-end reads.

reverse transcription quantitative real-time PCR (RT-qPCR): An application to measure RNA expression levels using qPCR. RNA starting material is reverse transcribed into complementary DNA (cDNA) by a reverse transcriptase enzyme. Expression levels are usually expressed as a relative value, in comparison to the expression of a reference gene.

sensitivity: In genomics, the ability to detect low-frequency variants or low-abundance transcripts.

target region: A specific sequence of the genome, identified as a region of interest, due to possible involvement in or association with biological development, pathogenesis, or other area of study of interest to the investigator. The sequence can be a gene, a gene segment, a gene fusion, a promotor region, part of an intron or exon, or any stretch of sequence of interest to the investigator.

transcriptome: Set of all RNA molecules, or transcripts, produced in one or a population of cells.

References

1. Doke T, Huang S, Qiu C, et al. Transcriptome-wide association analysis identifies *DACH1* as a kidney disease risk gene that contributes to fibrosis. *J Clin Invest.* 2021;131(10):e141801. doi:10.1172/JCI141801
2. Ozsolak F, Milos PM. RNA-Sequencing: advances, challenges and opportunities. *Nat Rev Genet.* 2011;12:87-98. doi:10.1038/nrg2934
3. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet.* 2009;10:57-63. doi:10.1038/nrg2484
4. Wilhelm BT, Landry JR. RNA-Seq—quantitative measurement of expression through massively parallel RNA-Sequencing. *Methods.* 2009;48:249-57. doi:10.1016/j.ymeth.2009.03.016
5. Geraci F, Saha I, Bianchini M. Editorial: RNA-Seq Analysis: Methods, Applications and Challenges. *Front Genet.* 2020;11:220. doi:10.3389/fgene.2020.00220
6. Corchete LA, Rojas EA, Alonso-López D, De Las Rivas J, Gutiérrez NC, Burguillo FJ. Systematic comparison and assessment of RNA-seq procedures for gene expression quantitative analysis. *Sci Rep.* 2020;10(1):19737. doi:10.1038/s41598-020-76881-x
7. Hong M, Tao S, Zhang L, et al. RNA sequencing: new technologies and applications in cancer research. *J Hematol Oncol.* 2020;13(1):166. doi:10.1186/s13045-020-01005-x
8. Sorokin M, Ignatev K, Poddubskaya E, et al. RNA Sequencing in Comparison to Immunohistochemistry for Measuring Cancer Biomarkers in Breast Cancer and Lung Cancer Specimens. *Biomedicines.* 2020;8(5):114. doi:10.3390/biomedicines8050114
9. Digital Science. Dimensions [Software] available from app.dimensions.ai. Accessed June 21, 2021, under licence agreement.
10. Maynard KR, Collado-Torres L, Weber LM, et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat Neurosci.* 2021;24(3):425-436. doi:10.1038/s41593-020-00787-0
11. Fatica A, Bozzoni I. Long noncoding RNAs: new players in cell differentiation and development. *Nat Rev Genet.* 2014;1:7-21. doi:10.1038/nrg3606
12. de Goede OM, Nachun DC, Ferraro NM, et al. Population-scale tissue transcriptomics maps long non-coding RNAs to complex disease. *Cell.* 2021;184(10):2633-2648.e19. doi:10.1016/j.cell.2021.03.050
13. Longo SK, Guo MG, Ji AL, Khavari PA. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet.* 2021;10.1038/s41576-021-00370-8. doi:10.1038/s41576-021-00370-8
14. Delorey TM, Ziegler CGK, Heimberg G, et al. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature.* 2021;595(7865):107-113. doi:10.1038/s41586-021-03570-8
15. Desai N, Neyaz A, Szabolcs A, et al. Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. *Nat Commun.* 2020;11(1):6319. doi:10.1038/s41467-020-20139-7
16. Scheid JF, Barnes CO, Eraslan B, et al. B cell genomics behind cross-neutralization of SARS-CoV-2 variants and SARS-CoV. *Cell.* 2021;184(12):3205-3221.e24. doi:10.1016/j.cell.2021.04.032
17. Huang N, Pérez P, Kato T, et al. SARS-CoV-2 infection of the oral cavity and saliva. *Nat Med.* 2021;27(5):892-903. doi:10.1038/s41591-021-01296-8
18. Butler D, Mozsary C, Meydan C, et al. Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. *Nat Commun.* 2021;12(1):1660. doi:10.1038/s41467-021-21361-7
19. Zhang Y, Narayanan SP, Mannan R, et al. Single-cell analyses of renal cell cancers reveal insights into tumor microenvironment, cell of origin, and therapy response. *Proc Natl Acad Sci U S A.* 2021;118(24):e2103240118. doi:10.1073/pnas.2103240118
20. Peng M, Wei G, Zhang Y, et al. Single-cell transcriptomic landscape reveals the differences in cell differentiation and immune microenvironment of papillary thyroid carcinoma between genders. *Cell Biosci.* 2021;11(1):39. doi:10.1186/s13578-021-00549-w
21. Lalli MA, Avey D, Dougherty JD, Milbrandt J, Mitra RD. High-throughput single-cell functional elucidation of neurodevelopmental disease-associated genes reveals convergent mechanisms altering neuronal differentiation. *Genome Res.* 2020;30(9):1317-1331. doi:10.1101/gr.262295.120
22. Brady L, Kriner M, Coleman I, et al. Inter- and intra-tumor heterogeneity of metastatic prostate cancer determined by digital spatial gene expression profiling. *Nat Commun.* 2021;12(1):1426. doi:10.1038/s41467-021-21615-4
23. Ji AL, Rubin AJ, Thrane K, et al. Multimodal Analysis of Composition and Spatial Architecture in Human Squamous Cell Carcinoma. *Cell.* 2020;182(6):1661-1662. doi:10.1016/j.cell.2020.08.043
24. Fawkner-Corbett D, Antanaviciute A, Parikh K, et al. Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell.* 2021;184(3):810-826.e23. doi:10.1016/j.cell.2020.12.016
25. Zhang X, van Rooij JGJ, Wakabayashi Y, et al. Genome-wide transcriptome study using deep RNA sequencing for myocardial infarction and coronary artery calcification. *BMC Med Genomics.* 2021;14(1):45. doi:10.1186/s12920-020-00838-2
26. Zhao S, Fung-Leung WP, Bittner A, Ngo K, Liu X. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PLoS One.* 2014;16;9(1):e78644. doi:10.1371/journal.pone.0078644
27. Wang C, Gong B, Bushel PR, et al. The concordance between RNA-Seq and microarray data depends on chemical treatment and transcript abundance. *Nat Biotechnol.* 2014;32:926-932. doi:10.1038/nbt.3001
28. Illumina. Illumina RNA Prep with Enrichment Data Sheet. illumina.com/products/by-type/sequencing-kits/library-prep-kits/rna-prep-enrichment.html. Published 2020. Accessed August 20, 2021.
29. Xu J, Gong B, Wu L, Thakkar S, Hong H, Tong W. Comprehensive assessments of RNA-Seq by the SEQC consortium: FDA-led efforts advance precision medicine. *Pharmaceutics.* 2016;8(1):8. doi:10.3390/pharmaceutics8010008
30. Crist AM, Hinkle KM, Wang X, et al. Transcriptomic analysis to identify genes associated with selective hippocampal vulnerability in Alzheimer's disease. *Nat Commun.* 2021;12(1):2311. doi:10.1038/s41467-021-22399-3
31. Allen M, Wang X, Burgess JD, et al. Conserved brain myelination networks are altered in Alzheimer's and other neurodegenerative diseases. *Alzheimers Dement.* 2018;14(3):352-366. doi:10.1016/j.jalz.2017.09.012

32. Guennewig B, Lim J, Marshall L, et al. Defining early changes in Alzheimer's disease from RNA sequencing of brain regions differentially affected by pathology. *Sci Rep.* 2021;11(1):4865. doi:10.1038/s41598-021-83872-z
33. Luo X, Yin J, Dwyer D, et al. Chamber-enriched gene expression profiles in failing human hearts with reduced ejection fraction. *Sci Rep.* 2021;11(1):11839. doi:10.1038/s41598-021-91214-2
34. Jain PN, Robertson M, Lasa JJ, et al. Altered metabolic and inflammatory transcriptomics after cardiac surgery in neonates with congenital heart disease. *Sci Rep.* 2021;11(1):4965. doi:10.1038/s41598-021-83882-x
35. Meng Z, Chen C, Cao H, Wang J, Shen E. Whole transcriptome sequencing reveals biologically significant RNA markers and related regulating biological pathways in cardiomyocyte hypertrophy induced by high glucose. *J Cell Biochem.* 2019;120(1):1018-1027. doi:10.1002/jcb.27546
36. Andersen JD, Jacobsen SB, Trudsø LC, Kampmann ML, Banner J, Morling N. Whole genome and transcriptome sequencing of post-mortem cardiac tissues from sudden cardiac death victims identifies a gene regulatory variant in NEXN. *Int J Legal Med.* 2019;133(6):1699-1709. doi:10.1007/s00414-019-02127-9
37. Srivastava SP, Goodwin JE, Tripathi P, Kanasaki K, Koya D. Interactions among Long Non-Coding RNAs and microRNAs Influence Disease Phenotype in Diabetes and Diabetic Kidney Disease. *Int J Mol Sci.* 2021;22(11):6027. doi:10.3390/ijms22116027
38. An M, Zang X, Wang J, Kang J, Tan X, Fu B. Comprehensive analysis of differentially expressed long noncoding RNAs, miRNAs and mRNAs in breast cancer brain metastasis. *Epigenomics.* 2021;10.2217/epi-2021-0152. doi:10.2217/epi-2021-0152
39. Pon JR, Marra MA. Driver and passenger mutations in cancer. *Annu Rev Pathol.* 2015;10:25-50. doi:10.1146/annurev-pathol-012414-040312
40. Mertens F, Tayebwa J. Evolving techniques for gene fusion detection in soft tissue tumours. *Histopathology.* 2014;64(1):151-162. doi:10.1111/his.12272
41. Chen X, Wang F, Zhang Y, et al. Fusion gene map of acute leukemia revealed by transcriptome sequencing of a consecutive cohort of 1000 cases in a single center. *Blood Cancer J.* 2021;11(6):112. doi:10.1038/s41408-021-00504-5
42. Zubovic L, Piazza S, Tebaldi T, et al. The altered transcriptome of pediatric myelodysplastic syndrome revealed by RNA sequencing. *J Hematol Oncol.* 2020;13(1):135. doi:10.1186/s13045-020-00974-3
43. Guo X, Wang S, Godwood A, et al. Pharmacodynamic biomarkers and differential effects of TNF- and GM-CSF-targeting biologics in rheumatoid arthritis. *Int J Rheum Dis.* 2019;22(4):646-653. doi:10.1111/1756-185X.13395
44. Malehmir M, Pfister D, Gallage S, et al. Platelet GPIbα is a mediator and potential interventional target for NASH and subsequent liver cancer. *Nat Med.* 2019;25(4):641-655. doi:10.1038/s41591-019-0379-5
45. Ferrero E, Agarwal P. Connecting genetics and gene expression data for target prioritisation and drug repositioning. *BioData Min.* 2018;11:7. doi:10.1186/s13040-018-0171-y
46. Zemek RM, De Jong E, Chin WL, et al. Sensitization to immune checkpoint blockade through activation of a STAT1/NK axis in the tumor microenvironment. *Sci Transl Med.* 2019;11(501):eaav7816. doi:10.1126/scitranslmed.aav7816
47. Bagaev A, Kotlov N, Nomie K, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell.* 2021;39(6):845-865.e7. doi:10.1016/j.ccell.2021.04.014
48. Kremer LS, Bader DM, Mertes C, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun.* 2017;8:15824. doi:10.1038/ncomms15824
49. Cummings BB, Marshall JL, Tukiainen T, et al. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med.* 2017;9(386):eaal5209. doi:10.1126/scitranslmed.aal5209
50. Frésard L, Smail C, Ferraro NM, et al. Identification of rare-disease genes using blood transcriptome sequencing and large control cohorts. *Nat Med.* 2019;25(6):911-919. doi:10.1038/s41591-019-0457-8
51. Mancuso N, Shi H, Goddard P, Kichaev G, Gusev A, Pasaniuc B. Integrating Gene Expression with Summary Association Statistics to Identify Genes Associated with 30 Complex Traits. *Am J Hum Genet.* 2017;100(3):473-487. doi:10.1016/j.ajhg.2017.01.031
52. Odhams CA, Cunninghame Graham DS, Vyse TJ. Profiling RNA-Seq at multiple resolutions markedly increases the number of causal eQTLs in autoimmune disease. *PLoS Genet.* 2017;13(10):e1007071. doi:10.1371/journal.pgen.1007071
53. Gusev A, Ko A, Shi H, et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat Genet.* 2016;48(3):245-252. doi:10.1038/ng.3506
54. Cao J, Spielmann M, Qiu X, et al. The single-cell transcriptional landscape of mammalian organogenesis. *Nature.* 2019;566(7745):496-502. doi:10.1038/s41586-019-0969-x
55. Plasschaert LW, Žilionis R, Choo-Wing R, et al. A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature.* 2018;560(7718):377-381. doi:10.1038/s41586-018-0394-6
56. Munchel S, Rohrback S, Randise-Hinchliff C, et al. Circulating transcripts in maternal blood reflect a molecular signature of early-onset preeclampsia. *Sci Transl Med.* 2020;12(550):eaaz0131. doi:10.1126/scitranslmed.aaz0131
57. Turchinovich A, Drapkina O, Tonevitsky A. Transcriptome of extracellular vesicles: State-of-the-art. *Front Immunol.* 2019;10:202. doi:10.3389/fimmu.2019.00202
58. Metzenmacher M, Váraljai R, Hegedüs B, et al. Plasma Next Generation Sequencing and Droplet Digital-qPCR-Based Quantification of Circulating Cell-Free RNA for Noninvasive Early Detection of Cancer. *Cancers (Basel).* 2020;12(2):353. doi:10.3390/cancers12020353
59. Ibarra A, Zhuang J, Zhao Y, et al. Non-invasive characterization of human bone marrow stimulation and reconstitution by cell-free messenger RNA sequencing. *Nat Commun.* 2020;11(1):400. doi:10.1038/s41467-019-14253-4
60. Toden S, Zhuang J, Acosta AD, et al. Noninvasive characterization of Alzheimer's disease by circulating, cell-free messenger RNA next-generation sequencing. *Sci Adv.* 2020;6(50):eabb1654. doi:10.1126/sciadv.abb1654
61. Hulstaert E, Morlion A, Avila Cobos F, et al. Charting Extracellular Transcriptomes in The Human Biofluid RNA Atlas. *Cell Rep.* 2020;33(13):108552. doi:10.1016/j.celrep.2020.108552
62. Larson MH, Pan W, Kim HJ, et al. A comprehensive characterization of the cell-free transcriptome reveals tissue- and subtype-specific biomarkers for cancer detection. *Nat Commun.* 2021;12(1):2357. doi:10.1038/s41467-021-22444-1
63. Stephenson E, Reynolds G, Botting RA, et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat Med.* 2021;27(5):904-916. doi:10.1038/s41591-021-01329-2

64. Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*. 2017;14(9):865-868. doi:10.1038/nmeth.4380
65. Mair F, Erickson JR, Voillet V, et al. A Targeted Multi-omic Analysis Approach Measures Protein Expression and Low-Abundance Transcripts on the Single-Cell Level. *Cell Rep*. 2020;31(1):107499. doi:10.1016/j.celrep.2020.03.063
66. Illumina. Illumina Stranded Total RNA Prep, Ligation with Ribo-Zero™ Plus Data Sheet. illumina.com/products/by-type/sequencing-kits/library-prep-kits/stranded-total-rna-prep.html. Published 2020. Accessed August 20, 2021.
67. Illumina. Illumina Stranded mRNA Prep, Ligation Data Sheet. illumina.com/products/by-type/sequencing-kits/library-prep-kits/stranded-mrna-prep.html. Published 2020. Accessed August 20, 2021.
68. Yilmaz A, Grotewold E. Components and mechanisms of regulation of gene expression. *Methods Mol Biol*. 2010;674:23-32. doi:10.1007/978-1-60761-854-6_2
69. Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science*. 2007;316(5830):1497-1502. doi:10.1126/science.1141319
70. Illumina. TruSeq ChIP Library Prep Kit Data Sheet. illumina.com/products/by-type/sequencing-kits/library-prep-kits/truseq-chip.html. Published 2014. Accessed August 20, 2021.
71. Kaya-Okur HS, Wu SJ, Codomo CA, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun*. 2019;10(1):1930. doi:10.1038/s41467-019-09982-5
72. Kaya-Okur HS, Janssens DH, Henikoff JG, Ahmad K, Henikoff S. Efficient low-cost chromatin profiling with CUT&Tag. *Nat Protoc*. 2020;15(10):3264-3283. doi:10.1038/s41596-020-0373-x
73. Henikoff S, Henikoff JG, Kaya-Okur HS, Ahmad K. Efficient chromatin accessibility mapping in situ by nucleosome-tethered tagmentation. *Elife*. 2020;9:e63274. doi:10.7554/eLife.63274
74. Li N, Jin K, Bai Y, Fu H, Liu L, Liu B. Tn5 Transposase Applied in Genomics Research. *Int J Mol Sci*. 2020;21(21):8329. doi:10.3390/ijms21218329
75. Jin F, Li Y, Dixon JR, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. 2013;503(7475):290-294. doi:10.1038/nature12644
76. Mumbach MR, Rubin AJ, Flynn RA, et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods*. 2016;13(11):919-922. doi:10.1038/nmeth.3999
77. Mumbach MR, Satpathy AT, Boyle EA, et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat Genet*. 2017;49(11):1602-1612. doi:10.1038/ng.3963
78. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol*. 2015;109:21.29.1-21.29.9. doi:10.1002/0471142727.mb2129s109
79. Marinov GK, Shipony Z. Interrogating the Accessible Chromatin Landscape of Eukaryote Genomes Using ATAC-seq. *Methods Mol Biol*. 2021;2243:183-226. doi:10.1007/978-1-0716-1103-6_10
80. Sun Y, Miao N, Sun T. Detect accessible chromatin using ATAC-sequencing, from principle to applications. *Hereditas*. 2019;156:29. doi:10.1186/s41065-019-0105-9
81. Portela A, Esteller M. Epigenetic modifications and human disease. *Nat Biotechnol*. 2010;28(10):1057-1068. doi:10.1038/nbt.1685
82. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nat Rev Genet*. 2011;12(8):529-541. doi:10.1038/nrg3000
83. Pierce SE, Granja JM, Greenleaf WJ. High-throughput single-cell chromatin accessibility CRISPR screens enable unbiased identification of regulatory networks in cancer. *Nat Commun*. 2021;12(1):2969. doi:10.1038/s41467-021-23213-w
84. Corces MR, Granja JM, Shams S, et al. The chromatin accessibility landscape of primary human cancers. *Science*. 2018;362(6413):eaav1898. doi:10.1126/science.aav1898
85. Rizzardi LF, Hickey PF, Rodriguez DiBlasi V, et al. Neuronal brain-region-specific DNA methylation and chromatin accessibility are associated with neuropsychiatric trait heritability. *Nat Neurosci*. 2019;22(2):307-316. doi:10.1038/s41593-018-0297-8
86. Dhana K, Braun K, Nano J, et al. An Epigenome-Wide Association Study of Obesity-Related Traits. *Am J Epidemiol*. 2018;187(8):1662-1669. doi:10.1093/aje/kwy025
87. Skvortsova K, Masle-Farquhar E, Luu PL, et al. DNA Hypermethylation Encroachment at CpG Island Borders in Cancer Is Predisposed by H3K4 Monomethylation Patterns. *Cancer Cell*. 2019;35(2):297-314.e8. doi:10.1016/j.ccell.2019.01.004
88. Konki M, Malonzo M, Karlsson IK, et al. Peripheral blood DNA methylation differences in twin pairs discordant for Alzheimer's disease. *Clin Epigenetics*. 2019;11:130. doi:10.1186/s13148-019-0729-7
89. Vijayakumar NT, Judy MV. Autism spectrum disorders: Integration of the genome, transcriptome and the environment. *J Neurol Sci*. 2016;364:167-76. doi:10.1016/j.jns.2016.03.026
90. Huan T, Joehanes R, Song C, et al. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nat Commun*. 2019;10:4267. doi:10.1038/s41467-019-12228-z
91. Smail HO. The epigenetics of diabetes, obesity, and cardiovascular disease. *AIMS Genet*. 2019;6:36-45. doi:10.3934/genet.2019.3.36
92. Illumina. TruSeq Methyl Capture EPIC Library Prep Kit Data Sheet. illumina.com/products/by-type/sequencing-kits/library-prep-kits/truseq-methyl-capture-epic.html. Published 2016. Accessed August 20, 2021.
93. Masser DR, Stanford DR, Freeman WM. Targeted DNA methylation analysis by next-generation sequencing. *J Vis Exp*. 2015;(96):52488. doi:10.3791/52488
94. Chen X, Dong Z, Hubbell E, et al. Prognostic Significance of Blood-Based Multi-cancer Detection in Plasma Cell-Free DNA. *Clin Cancer Res*. 2021;10.1158/1078-0432.CCR-21-0417. doi:10.1158/1078-0432.CCR-21-0417
95. Klein EA, Richards D, Cohn A, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. *Ann Oncol*. 2021;S0923-7534(21)02046-9. doi:10.1016/j.annonc.2021.05.806

96. Fernández-Santiago R, Merkel A, Castellano G, et al. Whole-genome DNA hypermethylation in iPSC-derived dopaminergic neurons from Parkinson's disease patients. *Clin Epigenetics.* 2019;11(1):108. doi:10.1186/s13148-019-0701-6
97. Shao X, Hudson M, Colmegna I, et al. Rheumatoid arthritis-relevant DNA methylation changes identified in ACPA-positive asymptomatic individuals using methylome capture sequencing. *Clin Epigenetics.* 2019;11(1):110. doi:10.1186/s13148-019-0699-9
98. Woo HD, Herceg Z. A Method to Investigate the Helicobacter pylori-Associated DNA Methylome. *Methods Mol Biol.* 2021;2283:75-81. doi:10.1007/978-1-0716-1302-3_9
99. Miyakuni K, Nishida J, Koinuma D, et al. Genome-wide analysis of DNA methylation identifies the apoptosis-related gene UQCRH as a tumor suppressor in renal cancer. *Mol Oncol.* 2021;10.1002/1878-0261.13040. doi:10.1002/1878-0261.13040
100. Chan RF, Shabalin AA, Montano C, et al. Independent Methylome-Wide Association Studies of Schizophrenia Detect Consistent Case-Control Differences. *Schizophr Bull.* 2020;46(2):319-327. doi:10.1093/schbul/sbz056
101. Xie T, Gorenjak V, Stathopoulou MG, et al. Epigenome-Wide Association Study (EWAS) of Blood Lipids in Healthy Population from STANISLAS Family Study. *Int J Mol Sci.* 2019;20(5):1014. doi:10.3390/ijms20051014
102. Zhu Y, Strachan E, Fowler E, et al. Genome-wide profiling of DNA methylome and transcriptome in peripheral blood monocytes for major depression: A Monozygotic Discordant Twin Study. *Transl Psychiatry.* 2019;9:215. doi:10.1038/s41398-019-0550-2
103. Navas-Acien A, Domingo-Relloso A, Subedi P, et al. Blood DNA Methylation and Incident Coronary Heart Disease: Evidence From the Strong Heart Study. *JAMA Cardiol.* 2021;10.1001/jamacardio.2021.2704. doi:10.1001/jamacardio.2021.2704
104. Gao H, He X, Li Q, et al. Genome-wide DNA methylome analysis reveals methylation subtypes with different clinical outcomes for acute myeloid leukemia patients. *Cancer Med.* 2020;9(17):6296-6305. doi:10.1002/cam4.3291
105. Koelsche C, Schrimpf D, Stichel D, et al. Sarcoma classification by DNA methylation profiling. *Nat Commun.* 2021;12(1):498. doi:10.1038/s41467-020-20603-4
106. Capper D, Jones DTW, Sill M, et al. DNA methylation-based classification of central nervous system tumours. *Nature.* 2018;555(7697):469-474. doi:10.1038/nature26000
107. Grabovska Y, Mackay A, O'Hare P, et al. Pediatric pan-central nervous system tumor analysis of immune-cell infiltration identifies correlates of antitumor immunity. *Nat Commun.* 2020;11(1):4324. doi:10.1038/s41467-020-18070-y.
108. Eales JM, Jiang X, Xu X, et al. Uncovering genetic mechanisms of hypertension through multi-omic analysis of the kidney. *Nat Genet.* 2021;53(5):630-637. doi:10.1038/s41588-021-00835-w
109. Muto Y, Wilson PC, Ledru N, et al. Single cell transcriptional and chromatin accessibility profiling redefine cellular heterogeneity in the adult human kidney. *Nat Commun.* 2021;12(1):2190. doi:10.1038/s41467-021-22368-w
110. Lynn RC, Weber EW, Sotillo E, et al. c-Jun overexpression in CAR T cells induces exhaustion resistance. *Nature.* 2019;576(7786):293-300. doi:10.1038/s41586-019-1805-z
111. Mendizabal I, Berto S, Usui N, et al. Cell type-specific epigenetic links to schizophrenia risk in the brain. *Genome Biol.* 2019;20(1):135. doi:10.1186/s13059-019-1747-7
112. Wang Y, Yuan P, Yan Z, et al. Single-cell multiomics sequencing reveals the functional regulatory landscape of early embryos. *Nat Commun.* 2021;12(1):1247. doi:10.1038/s41467-021-21409-8
113. Nativio R, Lan Y, Donahue G, et al. An integrated multi-omics approach identifies epigenetic alterations associated with Alzheimer's disease. *Nat Genet.* 2020;52(10):1024-1035. doi:10.1038/s41588-020-0696-0
114. Ringh MV, Hagemann-Jensen M, Needhamsen M, et al. Tobacco smoking induces changes in true DNA methylation, hydroxymethylation and gene expression in bronchoalveolar lavage cells. *EBioMedicine.* 2019;46:290-304. doi:10.1016/j.ebiom.2019.07.006

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-00258 v1.0