Single-Cell Research

An Overview of Recent Single-Cell Research Publications Featuring Illumina® Technology
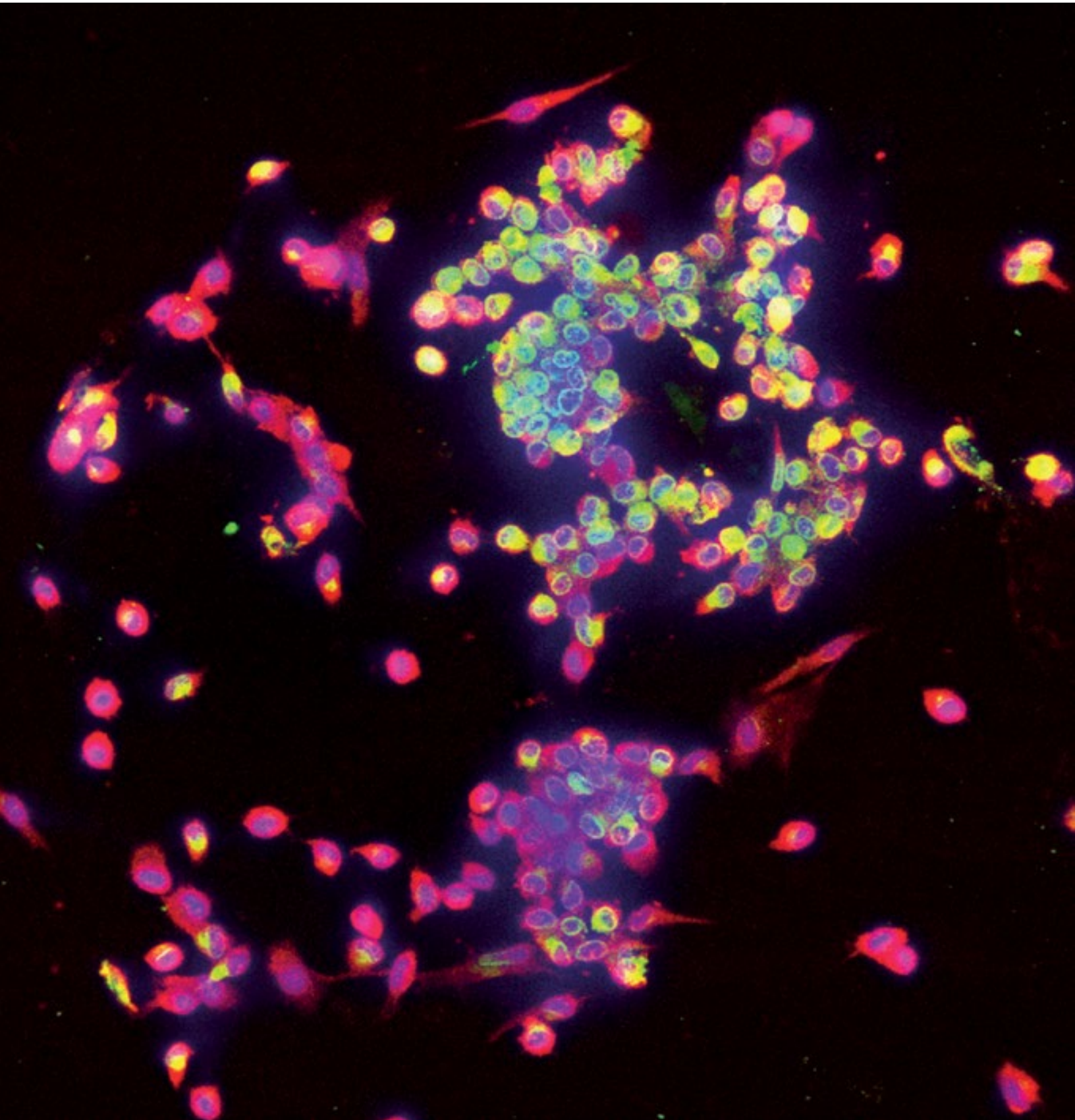
TABLE OF CONTENTS

Quartz-Seq

Smart-Seq

Smart-Seq2

Single-Cell Methylome & Transcriptome Sequencing

Genome & Transcriptome Sequencing

Genomic DNA and mRNA Sequencing

T Cell–Receptor Chain Pairing

Unique Molecular Identifiers

Cell Expression by Linear Amplification Sequencing

Flow Cell–Surface Reverse-Transcription Sequencing

Single-Cell Tagged Reverse-Transcription Sequencing

Fixed and Recovered Intact Single-Cell RNA Sequencing

Cell Labeling via Photobleaching

Indexing Droplets

Drop-Seq

CytoSeq

Single-Cell RNA Barcoding and Sequencing

High-Throughput Single-Cell Labeling

INTRODUCTION

Living tissues are composed of a variety of cell types. Each cell type has a distinct lineage and unique function that contribute to tissue and organ biology and, ultimately, define the biology of the organism as a whole. The lineage and development stage of each cell determine how they respond to other cells and to their native environment. In addition, subpopulations of cells of the same type are often genetically heterogeneous from each other as well as from other cell types.[1] Recently, scientists have launched the ambitious Human Cell Atlas Project, an international collaborative effort to map all cell types in the human body by using single-cell sequencing.[2]

> "Single-cell sequencing has emerged as a revolutionary method that reveals biological processes with unprecedented resolution and scale, and has already greatly impacted biology and medicine."
>
> - Benitez et al. 2017

Much of the initial impetus for single-cell tissue sequencing has come from cancer research, where cell lineage and detection of residual disease is of paramount importance.[3] Currently, single-cell approaches are also used to improve our understanding of other complex biological systems, including the central nervous system (CNS), immune system, and mammalian development.[4]

Single-cell sequencing is also an effective approach to characterize organisms that are difficult to culture *in vitro*.[5] Advances in single-cell sequencing have improved the detection and analysis of infectious diseases, food-borne pathogens, and microbial diversities in the environment or the gut. [6,7,8,9]

The high accuracy and specificity of next-generation sequencing (NGS) makes it ideal for single-cell and low-level DNA/RNA sequencing. The growing collection of published single-cell techniques includes detection of DNA mutations, copy-number variants (CNVs), DNA-protein binding, RNA splicing, and the measurement of mRNA expression.[10] More recently, microfluidics platforms and droplet-based methods have enabled massively parallel sequencing of mRNA in large numbers of individual cells.[11,12] The function of an individual cell is largely governed by interactions with its neighbors. This spatial context is typically lost in single-cell sequencing experiments, but new methods[13,14] and analysis algorithms[15] are combining measurements of single-cell gene expression with spatial localization within tissues.

This review highlights recent publications demonstrating how Illumina technology is being used in single-cell sequencing applications and techniques. To learn more about Illumina sequencing and microarray technologies, visit www.illumina.com.

1. Tanay A and Regev A. Scaling single-cell genomics from phenomenology to mechanism. *Nature.* 2017;541:331-338
2. Regev A, Teichmann SA, Lander ES, et al. The Human Cell Atlas. *Elife.* 2017;6: e27041
3. Zhang X, Marjani SL, Hu Z, Weissman SM, Pan X and Wu S. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Res.* 2016;76:1305-1312
4. Wang Y and Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* 2015;58:598-609
5. Solden L, Lloyd K and Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol.* 2016;31:217-226
6. Bergholz TM, Moreno Switt AI and Wiedmann M. Omics approaches in food safety: fulfilling the promise? *Trends Microbiol.* 2014;22:275-281
7. Woyke T, Doud DFR and Schulz F. The trajectory of microbial single-cell sequencing. *Nat Methods.* 2017;14:1045-1054
8. Mills E and Avraham R. Breaking the population barrier by single cell analysis: one host against one pathogen. *Curr Opin Microbiol.* 2017;36:69-75
9. Hug LA, Baker BJ, Anantharaman K, et al. A new view of the tree of life. *Nature Microbiology.* 2016;1:16048
10. Bacher R and Kendziorski C. Design and computational analysis of single-cell RNA-sequencing experiments. *Genome Biol.* 2016;17:63
11. Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* 2015;161:1202-1214
12. Klein AM, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell.* 2015;161:1187-1201
13. Lee JH, Daugharthy ER, Scheiman J, et al. Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat Protoc.* 2015;10:442-458
14. Lovatt D, Ruble BK, Lee J, et al. Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods.* 2014;11:190-196
15. Satija R, Farrell JA, Gennert D, Schier AF and Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* 2015;33:495-502
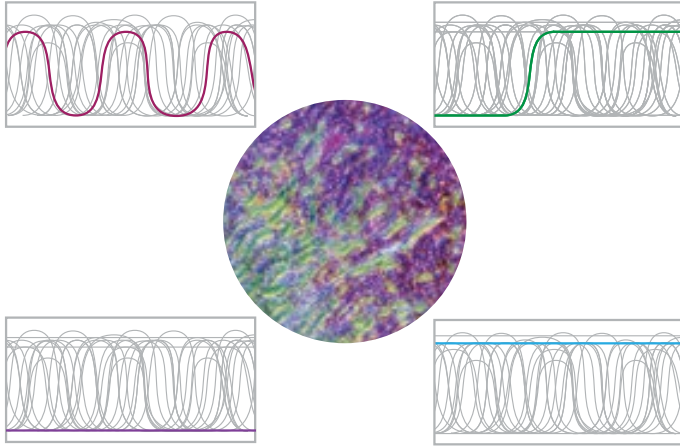
The same gene can be expressed at different levels, and influenced by different control mechanisms, in different cell types within the same tissue.

Reviews

Baslan T and Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer.* 2017;17:557-569

Benitez JA, Cheng S and Deng Q. Revealing allele-specific gene expression by single-cell transcriptomics. *Int J Biochem Cell Biol. 2017;90:155-160*

Haque A, Engel J, Teichmann SA and Lonnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med.* 2017;9:75

Marioni JC and Arendt D. How Single-Cell Genomics Is Changing Evolutionary and Developmental Biology. *Annu Rev Cell Dev Biol.* 2017;33:537-553

Tanay A and Regev A. Scaling single-cell genomics from phenomenology to mechanism. *Nature.* 2017;541:331-338

Gawad C, Koh W and Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet.* 2016;17:175-188

Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges [version 1; referees: 2 approved]. F1000Research. 2016;5(F1000 Faculty Review):182

Grun D and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell.* 2015;163:799-810

Wang Y and Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* 2015;58:598-609

APPLICATIONS

Cancer

Tumor progression occurs through driver mutations that undergo Darwinian selection for successive clonal expansion of tumor subclones. As a result, advanced tumors may contain a number of unique subclones[16,17,18] with different sets of mutations, different histopathology, and different responses to therapy.[19, 20,21] Molecular profiling of all subclones at diagnosis is important, because a subclone that makes up only 0.3% of a primary tumor can become the predominant clone following relapse.[22] Deep sequencing can detect subclone abundance as low as 1% of the total tumor cell population, but single-cell sequencing approaches are required to fully characterize therapeutic efficacy on rare cell populations.[23,24]

> "Understanding single cancer cells at their individual level and as an ensemble is bound to advance our understanding of not only therapeutic resistance but all facets of tumor biology." - Baslan and Hicks 2017

Circulating tumor cells (CTC) can also be used to detect cancer.[25,26] Cell-free detection of cancer nucleic acid markers—so-called liquid biopsies—may prove more sensitive and reproducible.[27,28] Single-cell approaches for the molecular profiling of cancer stem cells (CSCs) and disseminated cancer cells also add to our understanding of tumor development, metastasis, and therapeutic response.[29,30]

Recent clinical data have demonstrated that therapeutic enhancement of immune system function can improve cancer outcomes.[31] Antibodies that block cytotoxic T-lymphocyte–associated protein 4 (CTLA-4) as well as programmed death 1 (PD-1) induce clinical responses in a number of cancers, including melanoma, lung cancer, renal cancer, bladder cancer, and Hodgkin's lymphoma.[32,33] Single-cell sequencing approaches offer the possibility for a deeper understanding of the complex interactions among immune cells and tumor cells, as well as a more thorough characterization of the cellular ecosystem of tumors.[34,35,36]
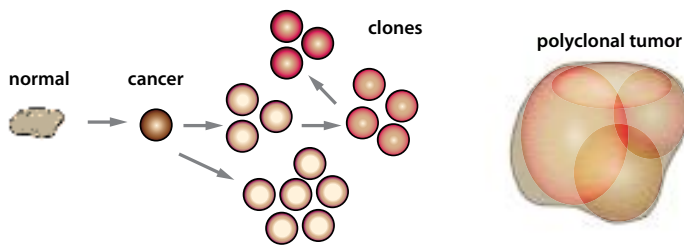
Intratumor heterogeneity: The progressive accumulation of somatic mutations results in a heterogeneous polyclonal tumor, in which different clones may respond differently to treatment.[37]

16. Ellsworth DL, Blackburn HL, Shriver CD, Rabizadeh S, Soon-Shiong P and Ellsworth RE. Single-cell sequencing and tumorigenesis: improved understanding of tumor evolution and metastasis. *Clin Transl Med.* 2017;6:15

17. Baslan T and Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer.* 2017;17:557-569

18. Litzenburger UM, Buenrostro JD, Wu B, et al. Single-cell epigenomic variability reveals functional cancer heterogeneity. *Genome Biol.* 2017;18:15

19. Li H, Courtois ET, Sengupta D, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet.* 2017;49:708-718

20. DiLoreto R, Khush K and De Vlaminck I. Precision monitoring of immunotherapies in solid organ and hematopoietic stem cell transplantation. *Adv Drug Deliv Rev.* 2017;114:272-284

21. Gerber T, Willscher E, Loeffler-Wirth H, et al. Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. *Oncotarget.* 2017;8:846-862

22. Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. *Blood.* 2016;127:2122-2130

23. Navin N and Hicks J. Future medical applications of single-cell sequencing in cancer. *Genome Med.* 2011;3:31

24. Hou Y, Song L, Zhu P, et al. Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell.* 2012;148:873-885

25. Krasnitz A, Kendall J, Alexander J, Levy D and Wigler M. Early Detection of Cancer in Blood Using Single-Cell Analysis: A Proposal. *Trends Mol Med.* 2017;23:594-603

26. Kondo Y, Hayashi K, Kawakami K, Miwa Y, Hayashi H and Yamamoto M. KRAS mutation analysis of single circulating tumor cells from patients with metastatic colorectal cancer. *BMC Cancer.* 2017;17:311

27. Swanton C. Plasma-derived tumor DNA analysis at whole-genome resolution. *Clin Chem.* 2013;59:6-8

28. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med.* 2014;20:548-554

29. Zhang X, Marjani SL, Hu Z, Weissman SM, Pan X and Wu S. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Res.* 2016;76:1305-1312

Reviews

Baslan T and Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer.* 2017;17:557-569

DiLoreto R, Khush K and De Vlaminck I. Precision monitoring of immunotherapies in solid organ and hematopoietic stem cell transplantation. *Adv Drug Deliv Rev.* 2017;114:272-284

Ellsworth DL, Blackburn HL, Shriver CD, Rabizadeh S, Soon-Shiong P and Ellsworth RE. Single-cell sequencing and tumorigenesis: improved understanding of tumor evolution and metastasis. *Clin Transl Med.* 2017;6:15

Zhang C, Guan Y, Sun Y, Ai D and Guo Q. Tumor heterogeneity and circulating tumor cells. *Cancer Lett.* 2016;374:216-223

Zhang X, Marjani SL, Hu Z, Weissman SM, Pan X and Wu S. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Res.* 2016;76:1305-1312

References

Chung W, Eum HH, Lee HO, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun.* 2017;8:15081
Single-cell approaches have demonstrated that tumors are genetically heterogeneous. In this study, the authors used the HiSeq™ 2500 System to perform RNA-Seq of 515 individual cells isolated from breast cancer tumors. Single-cell transcriptomic analyses identified heterogeneous as well as conserved gene expression signatures for subtype-specific breast cancer cells. In addition to breast cancer cells, the authors identified gene expression signatures for T cells, B cells, and macrophages. Individual T cells could be classified as activated or exhausted, suggesting that immune cell interactions within tumors are dynamic and distinct. These data suggest that individual breast cancer tumor microenvironments contain cells with unique patterns of gene expression, and that tumor subtypes are shaped by tumor cells as well as immune cells within the tumor microenvironment.

Illumina Technology: Nextera™ XT DNA Sample Prep Kit, TruSeq™ Rapid SBS Kit, HiSeq 2500 System

Giustacchini A, Thongjuea S, Barkas N, et al. Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. *Nat Med.* 2017;23:692-702
Cancer stem cells are rare cell types present in some tumors that possess the characteristics of normal stem cells and can give rise to tumors. In addition, cancer stem cells can be resistant to chemotherapy and can persist during remission. Because cancer stem cells are vastly outnumbered by tissue cells and are difficult to isolate, genetically characterizing them has been challenging. In this study, the authors used single-cell transcriptomic analyses to characterize cancer stem cells in chromic myeloid leukemia (CML). CML is typically characterized by the presence of the *BCR-ABL* fusion gene and is less genetically complex than other cancers. The authors developed a single-cell protocol for improved detection of *BCR-ABL* combined with RNA sequencing on the HiSeq 2000/4000 systems. They subsequently performed RNA-Seq on over 2000 individual stem cells from CML samples. Their data revealed the heterogeneity of CML cancer stem cells and identified gene expression signatures that correlated with poor treatment response in CML. Finally, their data suggest a gene expression signature for cancer stem cells, already present at CML diagnosis, that confers resistance to chemotherapy.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2000/4000 System

Lavin Y, Kobayashi S, Leader A, et al. Innate Immune Landscape in Early Lung Adenocarcinoma by Paired Single-Cell Analyses. *Cell.* 2017;169:750-765 e7172
Checkpoint blockade cancer immunotherapies act by altering the function of host T cells, and they have made a dramatic impact on the treatment of advanced non–small cell lung cancer (NSCLC). Tumor-infiltrating myeloid cells are host immune cells that present tumor-associated antigens to host T cells, and they contribute to tumor progression and treatment response. However, a detailed understanding of how the immune landscape develops in response to tumor cues is currently lacking. To address this knowledge gap, the authors performed single-cell analysis of immune cells from patient NSCLC tumors, noncancerous lung tissues, and blood. They found a higher number of immune cells in tumor than in nontumor lung tissue. Further, T and B cells were enriched in early-stage lung tumors, whereas natural killer (NK) cells were significantly reduced. In addition, the authors used single-cell RNA sequencing to characterize the genetic diversity of tumor-infiltrating monocytes from early-stage NSCLC and identified gene expression signatures for these monocyte cell types across NSCLC stages.

Illumina Technology: Unidentified Illumina System

30. Zhang C, Guan Y, Sun Y, Ai D and Guo Q. Tumor heterogeneity and circulating tumor cells. *Cancer Lett.* 2016;374:216-223
31. Wucherpfennig KW and Cartwright AN. Genetic screens to study the immune system in cancer. *Curr Opin Immunol.* 2016;41:55-61
32. Wolchok JD, Hodi FS, Weber JS, et al. Development of ipilimumab: a novel immunotherapeutic approach for the treatment of advanced melanoma. *Ann N Y Acad Sci.* 2013;1291:1-13
33. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med.* 2015;372:311-319
34. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352:189-196
35. Zheng C, Zheng L, Yoo JK, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell.* 2017;169:1342-1356 e1316
36. Chung W, Eum HH, Lee HO, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun.* 2017;8:15081
37. Navin NE. Cancer genomics: one cell at a time. *Genome Biol.* 2014;15:452
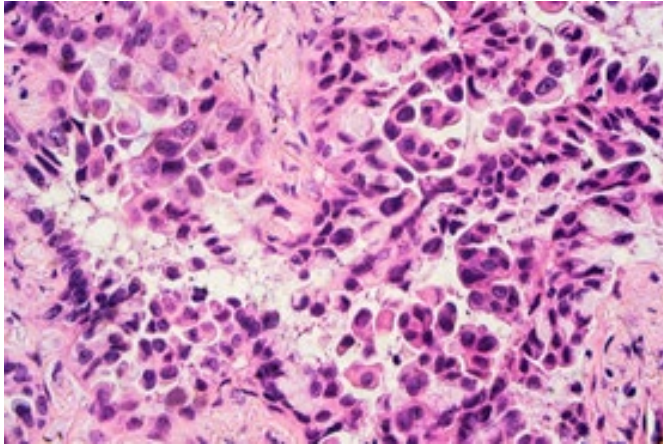
Hematoxylin and eosin (H&E) stained lung specimen containing adenocarcinoma

Levine MS, Bakker B, Boeckx B, et al. Centrosome Amplification Is Sufficient to Promote Spontaneous Tumorigenesis in Mammals. *Dev Cell.* **2017;40:313-322 e315**

The centrosome is the main microtubule-organizing center and regulator of cell-cycle progression in animal cells. In a normal cell, a centrosome is copied only once per cell cycle, with one copy passed on to the daughter cells during mitosis. In cancer cells, however, extra centrosomes are common; they correlate with poor prognosis and abnormal chromosome number, or aneuploidy. It is currently unclear whether extra centrosomes cause cancer or simply arise due to tumorigenesis. To address this question, the authors created transgenic mice overexpressing Plk4, a gene that regulates centrosome duplication. These transgenic mice have a persistent increase in centrosome number in cells from numerous tissues. The authors then karyotyped 99 individual epidermal cells by single-cell genomic copy-number analysis using the HiSeq 2500 System. Their data demonstrated that centrosome amplification was sufficient to promote chromosome segregation errors and aneuploidy in transgenic mouse cells. Further, starting at just 36 weeks, the Plk4 transgenic mice developed spontaneous aneuploid tumors. Taken together, the data suggest that centrosome amplification is sufficient to drive aneuploidy and cancer formation in mice.

Illumina Technology: TruSeq DNA Sample Preparation Kit v2, HiSeq 2500 System

Venteicher AS, Tirosh I, Hebert C, et al. Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science.* **2017;355:eaai8478**

Single-cell sequencing approaches can be useful for interrogating the genetic diversity of tumor cells, as well as characterizing the tumor microenvironment. However, many bulk datasets from large cohorts already exist, such as The Cancer Genome Atlas (TCGA). In this study, the authors combined single-cell RNA-Seq of 14,266 individual cells from 16 isocitrate dehydrogenase (IDH)-mutant gliomas, with 165 TCGA bulk RNA profiles. The TCGA RNA data demonstrated differences in gene expression in the 2 different subtypes of IDH-mutant gliomas, astrocytoma and oligodendroglioma. Single-cell RNA-Seq demonstrated that differences in gene expression were due to signature tumor events and tumor microenvironment composition. Notably, glial cell lineages were similar in both glioma subtypes. The study demonstrates the power of combining single-cell and bulk RNA-Seq approaches, and the data shed new light on the genetic composition of IDH-mutant glioma.

Illumina Technology: Nextera XT Library Prep Kit, NextSeq™ 500 System

Zheng C, Zheng L, Yoo JK, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell.* **2017;169:1342-1356 e1316**

Hepatocellular carcinoma (HCC) is a leading cause of cancer deaths worldwide, with the highest incidence in Asia and Africa. Despite the promise for other tumor types, cancer immunotherapies have not been successful in treating HCC, even though HCC tumors contain significant levels of tumor-infiltrating monocytes. In this study, the authors use single-cell RNA-Seq to understand and characterize tumor-infiltrating monocytes in HCC. Specifically, they isolated more than 5000 individual T cells from peripheral blood, tumor, and adjacent normal tissues from 6 HCC samples. They performed deep RNA-Seq on single T cells as well as T-cell receptor (TCR) sequencing using HiSeq 2500/4000 Systems. The combination of single-cell transcriptional profiles and TCR sequences identified 11 different T-cell subtypes. Notably, the authors found that exhausted CD8+ T cells and T_{regs} were enriched and clonally expanded in HCC. This study demonstrates the utility of single-cell sequencing approaches in characterizing the tumor microenvironment and characteristics of tumor-infiltrating cells in HCC.

Illumina Technology: HiSeq 2500/4000 Systems

Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352:189-196

Single-cell sequencing has the potential to inform treatment response and drug resistance by assessing malignant, microenvironmental, and immunologic states within tumors. In this study, the authors applied scRNA-Seq to 4645 single cells (malignant, stromal, immune, and endothelial) isolated from 19 patients with metastatic melanoma. They found that malignant cells within the same tumor displayed transcriptional heterogeneity associated with cell cycle, spatial context, and drug resistance. The same tumor had cells with high expression levels of microphthalmia-associated transcription factor (MITF), as well as cells with low MITF levels and elevated levels of AXL kinase (cells prone to early drug resistance). Infiltrating T-cell analysis revealed exhaustion programs, connection to T-cell activation/expansion, and patient variability. This study demonstrates how single-cell genomics can unravel the cellular ecosystem of tumors, with implications for targeted and immune therapies.

Illumina Technology: Nextera XT Sample Preparation Kit, NextSeq 500 System

Wei W, Shin YS, Xue M, et al. Single-Cell Phosphoproteomics Resolves Adaptive Signaling Dynamics and Informs Targeted Combination Therapy in Glioblastoma. *Cancer Cell.* 2016;29:563-573

Glioblastoma is one of the deadliest forms of cancer. Glioblastoma tumors have mutations in a number of druggable pathways, but current targeted therapies have proven ineffective due to rapid and universal drug resistance. Specifically, the mechanistic target of rapamycin (mTOR) pathway is a key driver in 90% of glioblastomas, yet tumor cells develop rapid resistance to mTOR-targeted therapies. In this study, the authors used the NextSeq 500 system to obtain single-cell genomic data, which they correlated with single-cell pro-teomic data in tumor cells treated with mTOR inhibitor. Their data showed that resistance to mTOR inhibitor in glioblastoma tumor cells occurred within days of drug therapy. Surprisingly, the correlation with single-cell sequencing data demonstrated that this drug resistance proceeds via nongenetic mechanisms, through upregulation of specific signaling phosphoproteins. This study suggests a novel approach for designing drug combination therapy in treating glioblastoma.

Illumina Technology: NextSeq 500 System

Engblom C, Pfirschke C, Zilionis R, et al. Osteoblasts remotely supply lung tumors with cancer-promoting SiglecF(high) neutrophils. *Science.* 2017;358:eaal5081

Gerber T, Willscher E, Loeffler-Wirth H, et al. Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. *Oncotarget.* 2017;8:846-862

Li H, Courtois ET, Sengupta D, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet.* 2017;49:708-718

Litzenburger UM, Buenrostro JD, Wu B, et al. Single-cell epigenomic variability reveals functional cancer heterogeneity. *Genome Biol.* 2017;18:15

Wang Q, Hu B, Hu X, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell.* 2017;32:42-56 e46

Hou Y, Guo H, Cao C, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcrip-tomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26:304-319

Kimmerling RJ, Lee Szeto G, Li JW, et al. A microfluidic platform enabling single-cell RNA-seq of multigenera-tional lineages. *Nat Commun.* 2016;7:10220

Nguyen A, Yoshida M, Goodarzi H and Tavazoie SF. Highly variable cancer subpopulations that exhibit en-hanced transcriptome variability and metastatic fitness. *Nat Commun.* 2016;7:11246

Piccirillo SG, Colman S, Potter NE, et al. Genetic and functional diversity of propagating cells in glioblastoma. *Stem Cell Reports.* 2015;4:7-15

Metagenomics

Although microorganisms represent the most abundant and diverse life forms on earth, in some environments only 0.1%–1% of the members have been cultivated.[38] Single-cell genomics approaches have led to a renewed understanding of microbial ecology, and they have greatly expanded our view of the tree of life.[39] This understanding has also revealed the extent and importance of viruses in the environment and their role in shaping bacterial populations. When combined with function- or phenotype-based screens, single-cell sequencing can enable the identification of microbial genomes that encode the function or phenotype.[40] Surprisingly, bacterial cell colonies—the paragon of microbial homogeneity—also display complex collective dynamics that allow for adaptation to their local environment.[41]

> "Single-cell genome sequencing of individual archaeal and bacterial cells is a vital approach to decipher the genetic makeup of uncultured microorganisms."
>
> - Bowers et al. 2017

Single-cell sequencing and metagenomics can help us understand the contribution of each microorganism to its surrounding environment.

Reviews

Bowers RM, Doud DFR and Woyke T. Analysis of single-cell genome sequences of bacteria and archaea. *Emerging Topics in Life Sciences.* 2017;1:249-255

Doud DFR and Woyke T. Novel approaches in function-driven single-cell genomics. *FEMS Microbiol Rev.* 2017;41:538-548

Kodzius R and Gojobori T. Single-cell technologies in environmental omics. *Gene.* 2016;576:701-707

Hedlund BP, Murugapiran SK, Alba TW, et al. Uncultivated thermophiles: current status and spotlight on 'Aigarchaeota'. *Curr Opin Microbiol.* 2015;25:136-145

Rashid M and Stingl U. Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. *Biotechnol Adv.* 2015;33:1755-1773

38. Solden L, Lloyd K and Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol.* 2016;31:217-226
39. Hug LA, Baker BJ, Anantharaman K, et al. A new view of the tree of life. *Nature Microbiology.* 2016;1:16048
40. Doud DFR and Woyke T. Novel approaches in function-driven single-cell genomics. *FEMS Microbiol Rev.* 2017;41:538-548
41. Cho H, Jonsson H, Campbell K, et al. Self-organization in high-density bacterial colonies: efficient crowd control. *PLoS Biol.* 2007;5:e302

References

Kashtan N, Roggensack SE, Berta-Thompson JW, Grinberg M, Stepanauskas R and Chisholm SW. Fundamental differences in diversity and genomic population structure between Atlantic and Pacific Prochlorococcus. *ISME J.* 2017;11:1997-2011

Prochlorococcus is a photosynthetic cyanobacterium that is highly abundant in the ocean surface and contributes to marine productivity. Previous studies have shown that *Prochlorococcus* populations from different sites in the Atlantic Ocean can be highly diverse genomically, suggesting that ocean habitats can affect genomic structure and diversity.[42] In this study, the authors sampled wild populations of cells from habitats in the Pacific and Atlantic Oceans. They used the Illumina GAIIx System to perform single-cell sequencing of *Prochlorococcus* from the different sites. They found that *Prochlorococcus* from the 2 different ocean sites were composed of nonoverlapping and distinct subpopulations with different genomic backbones. These data suggest that environmental selection pressures, as well as geographic separation, can shape *Prochlorococcus* populations.

Illumina Technology: GAIIx System

Dyksma S, Bischof K, Fuchs BM, et al. Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. *ISME J.* 2016;8:1939-1953

Marine sediments are the largest carbon sink on the planet, with half of chemosynthetic oceanic carbon fixation occurring in coastal sediments. However, the microbes responsible for this activity are unknown. By surveying bacterial 16S ribosomal RNA (rRNA) gene diversity from 13 coastal sediments across Europe and Australia, the authors identified groups of *Gammaproteobacteria* that were affiliated with sulfur-oxidizing bacteria. ^{14}C-carbon assimilation studies showed that these uncultured *Gammaproteobacteria* accounted for 80% of carbon fixation in coastal sediments. Finally, the authors isolated individual cells from the environmental sample and performed single-cell whole-genome sequencing (WGS) to identify genes that linked hydrogen-oxidizing activity with sulfur-oxidizing *Gammaproteobacteria*.

Illumina Technology: MiSeq™ and HiSeq 2000 System

Eloe-Fadrosh EA, Paez-Espino D, Jarett J, et al. Global metagenomic survey reveals a new bacterial candidate phylum in geothermal springs. *Nat Commun.* 2016;7:10476

Molecular environmental surveys using 16S rRNA sequencing have greatly expanded our knowledge of microbial phylogenetic diversity. However, some bacterial and archaeal clades can be systematically underrepresented in current surveys or missed altogether. In this study, the authors analyzed 5.2 Tb of metagenomic data and discovered a novel bacterium (*Candidatus* Kryptonia) found exclusively in geothermal springs. The lineage had been missed in classical metagenomic surveys, because of mismatches in commonly used 16S rRNA primers. The authors combined metagenomic data with single-cell sequencing to generate high-quality genomes that represented 4 unique genera within this phylum.

Illumina Technology: MiSeq System

Mende DR, Aylward FO, Eppley JM, Nielsen TN and DeLong EF. Improved Environmental Genomes via Integration of Metagenomic and Single-Cell Assemblies. *Front Microbiol.* 2016;7:143

Single-cell genomics has led to a number of individual draft genomes for uncultivated microbes; however, multiple-strand displacement amplification (MDA) artifacts during the amplification step lead to incomplete and uneven coverage. Metagenomic data sets do not suffer the same sequence bias, but the genomic complexity of microbial communities precludes the recovery of draft genomes. In this study, the authors developed a new method for generating population genome assemblies from metagenomic-guided, single-cell amplified genome assembly data. They validated the approach by completing single-cell amplified genomes for Marine Group 1 Thaumarchaeota and SAR324 clade bacterioplankton. The improved method assembly of the SAR324 clade genome revealed the presence of many genes not present in the single-cell amplified genome.

Illumina Technology: TruSeq LT Nano Kit, MiSeq System

Spencer SJ, Tamminen MV, Preheim SP, et al. Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *ISME J.* 2016;10:427-436

In microbial ecology studies, 16S rRNA sequencing can identify microbial community members, whereas shotgun metagenomics can determine the functional diversity of the community. However, combining the 2 approaches is technically challenging. In this study, the authors developed emulsion, paired isolation, and concatenation PCR (epicPCR), a technique that links functional genes and phylogenetic markers. They applied the technique to millions of uncultured individual cells from the freshwater Upper Mystic Lake in Massachusetts. Specifically, they profiled the sulfate-reducing community within the freshwater lake community and were able to identify new putative sulfate reducers. The method is suitable for identifying functional community members, tracing gene transfer, and mapping ecological interactions in microbial cells.
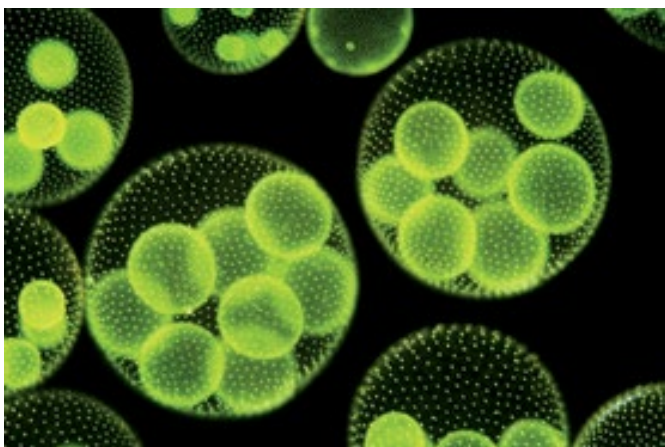
Illumina Technology: MiSeq System

42. Kashtan N, Roggensack SE, Rodrigue S, et al. Single-cell genomics reveals hundreds of coexisting subpopulations in wild Prochlorococcus. *Science.* 2014;344:416-420

Lima-Mendez G, Faust K, Henry N, et al. Ocean plankton. Determinants of community structure in the global plankton interactome. *Science.* 2015;348:1262073

Oceanic plankton is the world's largest ecosystem and is composed of viruses, prokaryotes, microbial eukaryotes, phytoplankton, and zooplankton. This ecosystem structure and composition are influenced by environmental conditions and nutrient availability. In this study, the authors analyzed 313 plankton samples from the *Tara* Oceans expedition and obtained viral, eukaryotic, and prokaryotic abundance profiles from Illumina-sequenced metagenomes and 18S rDNA V9 sequences. They used network inference and machine-learning methods to construct an interactome among plankton groups. In particular, the authors confirmed predicted virus-host interactions by comparing putative host contigs with viral data from single-cell genomes.

Illumina Technology: Illumina-sequenced metagenomes (mitags) and 18S rRNA V9 sequences

Volvox aureus.

Hosokawa M, Nishikawa Y, Kogawa M and Takeyama H. Massively parallel whole genome amplification for single-cell sequencing using droplet microfluidics. *Sci Rep.* 2017;7:5199

Lan F, Demaree B, Ahmed N and Abate AR. Single-cell genome sequencing at ultra-high-throughput with microfluidic droplet barcoding. *Nat Biotechnol.* 2017;35:640-646

Leon-Zayas R, Peoples L, Biddle JF, et al. The metabolic potential of the single cell genomes obtained from the Challenger Deep, Mariana Trench within the candidate superphylum Parcubacteria (OD1). *Environ Microbiol.* 2017;19:2769-2784

Neave MJ, Rachmawati R, Xun L, et al. Differential specificity between closely related corals and abundant Endozoicomonas endosymbionts across global scales. *ISME J.* 2017;11:186-200

Kolinko S, Richter M, Glockner FO, Brachmann A and Schuler D. Single-cell genomics of uncultivated deep-branching magnetotactic bacteria reveals a conserved set of magnetosome genes. *Environ Microbiol.* 2016;18:21-37

Mende DR, Aylward FO, Eppley JM, Nielsen TN and DeLong EF. Improved Environmental Genomes via Integration of Metagenomic and Single-Cell Assemblies. *Front Microbiol.* 2016;7:143

Stem Cells

Human life begins from a single oocyte, which undergoes mitotic divisions to generate a population of cells that make up the human embryo.[43] Embryonic stem cells (ESCs) are pluripotent stem cells derived from the inner cell mass of a blastocyst, an early-stage preimplantation embryo. Each stem cell undergoes a series of cell divisions that results in a specific lineage, which determines its genetic code and response to local environmental factors. This process gives rise to an array of unique, genetically heterogeneous cells.[44] Sequencing these single stem cells during differentiation has helped elucidate the underlying mechanisms.[45,46]

> "Single-cell sequencing provides powerful tools for characterizing the omic-scale features of heterogeneous cell populations including, those of stem cells." - Wen and Tang 2016

Hematopoietic stem cells (HSCs), and the mechanisms regulating their differentiation into erythroid, myeloid, or lymphoid lineages, are a unique example of cellular development.[47] Single-cell sequencing has helped to identify a population of neural stem cells that become activated in response to brain injury,[48] suggesting a possible approach for treating traumatic brain injury. Induced pluripotent stem cells (iPSCs), a type of pluripotent stem cell that can be generated directly from adult cells, also have potential use in cell-replacement therapies. Single-cell sequencing has helped to characterize the genetic heterogeneity of individual iPSCs, mechanisms regulating their differentiation and pluripotency, and self-organization into organoid structures.[49,50,51]
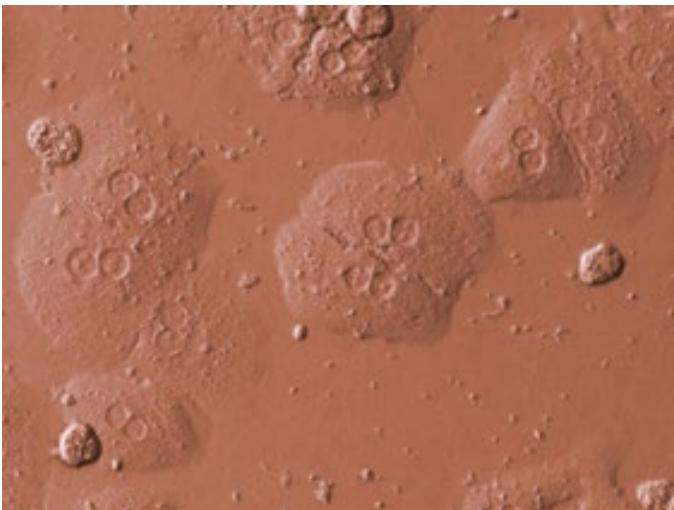
43. Speicher MR. Single-cell analysis: toward the clinic. *Genome Med.* 2013;5:74

44. Voet T, Kumar P, Van Loo P, et al. Single-cell paired-end genome sequencing reveals structural variation per cell cycle. *Nucleic Acids Res.* 2013;41:6119-6138

45. Yao Z, Mich JK, Ku S, et al. A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell.* 2017;20:120-134

46. Kumar P, Tan Y and Cahan P. Understanding development and stem cells using single cell-based analyses of gene expression. *Development.* 2017;144:17-32

47. Paul F, Arkin Y, Giladi A, et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell.* 2015;163:1663-1677

48. Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K and Martin-Villalba A. Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. *Cell Stem Cell.* 2015;17:329-340

49. Camp JG and Treutlein B. Human organomics: a fresh approach to understanding human development using single-cell transcriptomics. *Development.* 2017;144:1584-1587

50. Li C, Klco JM, Helton NM, et al. Genetic heterogeneity of induced pluripotent stem cells: results from 24 clones derived from a single C57BL/6 mouse. *PLoS One.* 2015;10:e0120585

51. Cacchiarelli D, Trapnell C, Ziller MJ, et al. Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. *Cell.* 2015;162:412-424

Embryonic stem cells.

Reviews

Camp JG and Treutlein B. Human organomics: a fresh approach to understanding human development using single-cell transcriptomics. *Development.* 2017;144:1584-1587

Kumar P, Tan Y and Cahan P. Understanding development and stem cells using single cell-based analyses of gene expression. *Development.* 2017;144:17-32

Ye F, Huang W and Guo G. Studying hematopoiesis using single-cell technologies. *J Hematol Oncol.* 2017;10:27

Wen L and Tang F. Single-cell sequencing in stem cell biology. *Genome Biol.* 2016;17:71

References

Cabezas-Wallscheid N, Buettner F, Sommerkamp P, et al. Vitamin A-Retinoic Acid Signaling Regulates Hematopoietic Stem Cell Dormancy. *Cell.* 2017;169:807-823 e819
HSCs are pluripotent stem cells found in the bone marrow, and they give rise to all myeloid and lymphoid cell lineages in the blood. In mice, a small subset of HSCs are dormant, dividing only 5 times per lifetime and maintaining a high long-term reconstitution potential. Dormant HSCs can transition into activated HSCs, which are still quiescent but primed for cell-cycle entry. The mechanisms regulating this transition are not well understood. In this study, the authors used single-cell RNA-Seq to compare gene expression in individual dormant HSCs, activated HSCs, and multipotent progenitor cells. They isolated HSCs using fluorescence-activated cell sorting (FACS) with cells from transgenic mice overexpressing an HSC-specific green fluorescent protein (GFP) marker. The single-cell transcriptome data revealed that dormant HSCs were characterized by low expression levels of the *Myc* gene and high expression levels of retinoic acid genes. Further, dormant HSCs expressed low levels of genes required for essential biosynthetic pathways, and transition to activated HSCs was characterized by a continuous increase in biosynthetic gene expression. Overall, their data shed new light on the biological pathways that regulate activation of HSCs and suggest that retinoic acid-vitamin A may regulate HSC plasticity in mice.

Illumina Technology: Nextera XT Sample Preparation Kit, HiSeq 2000/2500 Systems

Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol.* 2017;19:271-281
Bulk RNA-Seq approaches average gene expression data from individual cells together, destroying crucial information on cell-state transition. Consequently, bulk RNA-Seq is not suitable for characterizing cellular states or their role in development. Single-cell RNA-Seq approaches allow for the measurement of gene expression from large numbers of individual cells, allowing the cataloging and definition of cellular states. In the case of hematopoiesis, current models suggest that HSCs differentiate stepwise into discrete multipotent progenitor cell lineages.[52] In the current study, the authors combined single-cell RNA-Seq with functional analyses to challenge this model. They used the HiSeq 2500 System to perform single-cell sequencing of 1,413 single HSCs isolated from human bone marrow. They also individually cultured *ex vivo* 2038 single HSCs to characterize lineage potential. These 2 data sets were then integrated to quantitatively link transcriptomic and functional single-cell data. Their data suggest that individual HSCs gradually acquire lineage biases across multiple directions, without transitioning into discrete, hierarchically organized progenitor cells.

Illumina Technology: HiSeq 2500 System

Freeman BT, Jung JP and Ogle BM. Single-cell RNA-seq reveals activation of unique gene groups as a consequence of stem cell-parenchymal cell fusion. *Sci Rep.* 2016;6:23270
Bone marrow stem cell transplants are commonly used to treat conditions such as leukemia and lymphoma. Following transplantation, mesenchymal stem cells can fuse with parenchymal cells of the brain, liver, small intestine, and heart. The resultant effects of this hybridization are not known. In this study, the authors used the MiSeq system to perform scRNA-Seq of individual mesenchymal stem cell–cardiomyocyte hybrids to characterize global gene expression. The expression of cell-cycle genes was generally decreased in hybrids. However, for most other gene groups, individual hybrids were genetically distinct. Moreover, 2 hybrids were genetically similar to breast cancer cells, suggesting that monitoring stem cell transplantation for tumor emergence is warranted.

Illumina Technology: Nextera XT Sample Preparation Kit, MiSeq System

52. Chao MP, Seita J and Weissman IL. Establishment of a normal hematopoietic and leukemia stem cell hierarchy. Cold Spring Harb Symp Quant Biol. 2008;73:439-449

For Research Use Only. Not for use in diagnostic procedures.

Kang E, Wang X, Tippner-Hedges R, et al. Age-Related Accumulation of Somatic Mitochondrial DNA Mutations in Adult-Derived Human iPSCs. Cell Stem *Cell.* 2016;18:625-636

iPSCs offer the potential for autologous cell-replacement therapies. Maintaining the genetic integrity of the cultured iPSCs is an important consideration in potential therapy. The authors characterized the frequency of somatic mitochondrial DNA (mtDNA) mutations in cells derived from young or elderly subjects. The data show that the mtDNA mutation frequency in iPSCs increases with the age of the individual. The results highlight the importance of genetically monitoring mtDNA mutations in iPSCs, especially those that are generated from older patients.

Illumina Technology: Nextera XT Sample Preparation Kit, MiSeq System

Luo Y, Coskun V, Liang A, et al. Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell.* 2015;161:1175-1186

The scarcity of tissue-specific stem cells, and the complexity of their surrounding environment, make single-cell sequencing methods imperative for characterizing these cell types. In this study, the authors used single-cell sequencing and weighted gene coexpression network analysis to identify CD133+ ependymal cells from the adult mouse forebrain neurogenic zone as NSCs. These subpopulations of cells were enriched for immune-responsive genes, as well as genes encoding angiogenic factors. Administration of vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (bFGF) enhanced migration and elicited differentiation into neurons and glia.

Illumina Technology: HiSeq 2500 System

Paul F, Arkin Y, Giladi A, et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell.* 2015;163:1663-1677

Hematopoiesis is the developmental process by which bone marrow–derived HSCs differentiate into erythroid, myeloid, or lymphoid lineages. In this study, the authors combined massively parallel scRNA-Seq with fluorescence-activated cell sorting (FACS), chromatin profiling, genetic perturbation, and computational modeling to characterize the transcriptome of myeloid progenitor populations. The single-cell transcription data grouped HSCs into multiple progenitor subgroups with 7 differentiation states. These data provide a new reference model for studying hematopoiesis at the single-cell level.

Illumina Technology: NextSeq 500 and HiSeq 1500 System

Athanasiadis EI, Botthof JG, Andres H, Ferreira L, Lio P and Cvejic A. Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat Commun.* 2017;8:2045

Chen YJ, Friedman BA, Ha C, et al. Single-cell RNA sequencing identifies distinct mouse medial ganglionic eminence cell types. *Sci Rep.* 2017;7:45656

Dulken BW, Leeman DS, Boutet SC, Hebestreit K and Brunet A. Single-Cell Transcriptomic Analysis Defines Heterogeneity and Transcriptional Dynamics in the Adult Neural Stem Cell Lineage. *Cell Rep.* 2017;18:777-790

Kee N, Volakakis N, Kirkeby A, et al. Single-Cell Analysis Reveals a Close Relationship between Differentiating Dopamine and Subthalamic Nucleus Neuronal Lineages. *Cell Stem Cell.* 2017;20:29-40

Li L, Dong J, Yan L, et al. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell.* 2017;20:858-873 e854

Yang J, Ryan DJ, Wang W, et al. Establishment of mouse expanded potential stem cells. *Nature.* 2017;550:393-397

Yao Z, Mich JK, Ku S, et al. A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell.* 2017;20:120-134

Dixit A, Parnas O, Li B, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell.* 2016;167:1853-1866 e1817

Milani P, Escalante-Chong R, Shelley BC, et al. Cell freezing protocol suitable for ATAC-Seq on motor neurons derived from human induced pluripotent stem cells. *Sci Rep.* 2016;6:25474

Zhao X, Han Y, Liang Y, Nie C and Wang J. RNA-Seq Reveals the Angiogenesis Diversity between the Fetal and Adults Bone Mesenchyme Stem Cell. *PLoS One.* 2016;11:e0149171
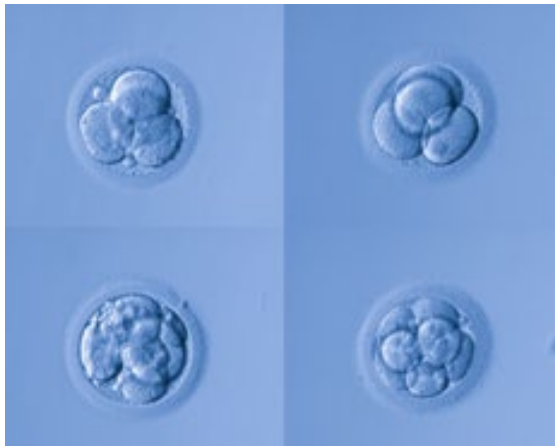
Developmental Biology

Typically, only small numbers of cells are available to researchers studying developmental biology. Minute changes in cellular environment, as well as temporal changes, have profound effects on cell behavior.[53] Single-cell approaches have led to a new understanding of tumor,[54,55] neuronal,[56,57] and immune system,[58,59] development. In turn, this understanding has led to breakthroughs in the understanding of the cellular lineages and regulatory networks underlying hematopoiesis.[60,61,62]

> "Assaying the genomes of multiple cells during development reveals developmental lineage trees whereas cellular transcriptomes inform us about the regulatory state of cells and their gradual restriction in potency." - Marioni and Arendt 2017

The study of human preimplantation development has been based typically on small numbers of samples, often pooled.[63] These studies fail to capture a detailed view of the first days of human preimplantation development. Recent single-cell sequencing studies are beginning to capture a more comprehensive and detailed view.[64]

Single-cell approaches are also providing a window into embryonic development in other model organisms (such as mouse,[65] zebrafish,[66] and *Caenorhabditis elegans*[67]), and they are leading to a greater understanding of the regulatory mechanisms underlying embryonic development in humans.[68,69] Comparative single-cell studies are even beginning to reveal the evolutionary history of germ layers—a fundamental concept in developmental biology for the past 150 years.[70]

During human embryogenesis, the single-cell zygote divides several times to form a morula.

53. Yalcin D, Hakguder ZM and Otu HH. Bioinformatics approaches to single-cell analysis in developmental biology. *Mol Hum Reprod.* 2016;22:182-192

54. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352:189-196

55. Ortmann CA, Kent DG, Nangalia J, et al. Effect of mutation order on myeloproliferative neoplasms. *N Engl J Med.* 2015;372:601-612

56. Lodato MA, Woodworth MB, Lee S, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science.* 2015;350:94-98

57. Quadrato G, Nguyen T, Macosko EZ, et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature.* 2017;545:48-53

58. Stubbington MJT, Rozenblatt-Rosen O, Regev A and Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. *Science.* 2017;358:58-63

59. Villani AC, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science.* 2017;356:eaah4573

60. Perie L and Duffy KR. Retracing the in vivo haematopoietic tree using single-cell methods. *FEBS Lett.* 2016;590:4068-4083

61. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol.* 2017;19:271-281

62. Paul F, Arkin Y, Giladi A, et al. Transcriptional Heterogeneity and Lineage Commitment in Myeloid Progenitors. *Cell.* 2015;163:1663-1677

63. Zhang P, Zucchelli M, Bruce S, et al. Transcriptome profiling of human pre-implantation development. *PLoS One.* 2009;4:e7844

64. Mohammed H, Hernando-Herraez I, Savino A, et al. Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Rep.* 2017;20:1215-1228

65. Bolton H, Graham SJ, Van der Aa N, et al. Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. *Nat Commun.* 2016;7:11165

66. Junker JP, Noel ES, Guryev V, et al. Genome-wide RNA Tomography in the zebrafish embryo. *Cell.* 2014;159:662-675

67. Cao J, Packer JS, Ramani V, et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science.* 2017;357:661-667

68. Li L, Dong J, Yan L, et al. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. Cell Stem *Cell.* 2017;20:858-873 e854

69. Otte J, Wruck W and Adjaye J. New insights into human primordial germ cells and early embryonic development from single-cell analysis. *FEBS Lett.* 2017;591:2226-2240

70. Hashimshony T, Feder M, Levin M, Hall BK and Yanai I. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature.* 2015;519:219-222

Reviews

Marioni JC and Arendt D. How Single-Cell Genomics Is Changing Evolutionary and Developmental Biology. *Annu Rev Cell Dev Biol.* 2017;33:537-553

Woodworth MB, Girskis KM and Walsh CA. Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat Rev Genet.* 2017;18:230-244

Perie L and Duffy KR. Retracing the in vivo haematopoietic tree using single-cell methods. *FEBS Lett.* 2016;590:4068-4083

Yalcin D, Hakguder ZM and Otu HH. Bioinformatics approaches to single-cell analysis in developmental biology. *Mol Hum Reprod.* 2016;22:182-192

References

Bae T, Tomasini L, Mariani J, et al. Different mutational rates and mechanisms in human cells at pre-gastrulation and neurogenesis. *Science.* **2018;359:550-555**
Somatic mutations are permanent changes in the DNA sequence of a genome. They arise because of mistakes in DNA replication in cells or from environmental factors. Previous studies using single-cell sequencing approaches have demonstrated that individual neurons contain unique single-nucleotide variant (SNV) somatic mutations, a phenomenon referred to as genomic mosaicism.[71] Although genomic mosaicism is common among multiple cell types, neurons are postmitotic, and thus the SNVs have the potential to persist and alter neuronal function. In this study, the authors isolated individual neurons from the forebrains of embryonic samples and grew them into small clones in culture. They used the HiSeq platform to analyze the genomes of these samples, and they identified 200–400 SNVs per cell. The SNV patterns in these embryonic neurons resembled those found in cancer cell genomes. By analyzing the genomes of neurons at different stages of embryonic development, they determined that the mutation rate speeds up during neurogenesis. These data suggest that embryonic neurons are protected from mutations, compared to neurons at later stages of neurodevelopment.

Illumina Technology: HiSeq 2500 and HiSeq X Systems

Cao J, Packer JS, Ramani V, et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science.* **2017;357:661-667**
Single-cell sequencing approaches have been instrumental in highlighting the genetic and cellular diversity of biological tissues. In this study, the authors extended this concept in a tour de force single-cell sequencing study of an entire organism. They used the NextSeq 500 System to perform single-cell RNA-Seq of all 50,000 cells of the *Caenorhabditis elegans* L2 larval stage. To accomplish this, they developed a combinatorial indexing method to uniquely label the transcriptomes of large numbers of single cells, a method termed sci-RNA-seq. Their single-cell transcriptomic data indicated that the *C. elegans* L2 larva is composed of 27 different cell types, including some that are extremely rare within the organism. Further, they integrated the sci-RNA-seq data with chromatin immunoprecipitation sequencing (ChIP-Seq) data to determine the effects of transcription factors on specific cell types. Their work demonstrates the ability of single-cell sequencing approaches to reconstruct the cell complement for an entire organism, and it provides a valuable resource for nematode biologists.

Illumina Technology: Nextera XT Sample Preparation Kit, NextSeq 500 System
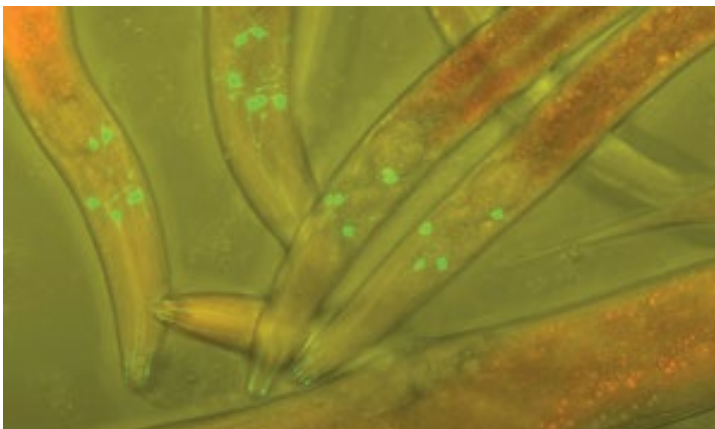
Fluorescently labeled cells in the roundworm, *Caenorhabditis elegans.*

71. Lodato MA, Woodworth MB, Lee S, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science.* 2015;350:94-98

Halpern KB, Shenhav R, Matcovitch-Natan O, et al. Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature.* **2017;542:352-356**

The liver is an organ that is critically important for detoxifying metabolites, producing carrier proteins, and producing biochemicals necessary for digestion. Anatomically, it is composed of hepatic lobules, hexagonal-shaped structures vital for carrying blood out from the liver. The hepatic lobules are composed of 15 radial layers of hepatocytes, the primary cell type within the liver, and each radial layer specializes in a distinct process.[72] Given the highly specialized microenvironment within each hepatic lobule, bulk sequencing approaches lack the resolution to characterize zonal gene expression within each radial layer. In this study, the authors combined single-cell RNA sequencing of dissociated individual hepatocytes with single-molecule fluorescence *in situ* hybridization (smFISH) of liver tissue sections. Single-cell RNA-Seq enabled measurement of gene expression patterns of thousands of cells, while smFISH enabled measurement of *in situ* expression levels for numerous landmark genes. The combined data revealed the zonal expression of all expressed genes within the hepatic lobule, with 50% of all liver genes expressed zonally. This approach is suitable for reconstructing spatial gene expression organization of other organs.

Illumina Technology: HiSeq 2500 System

Bolton H, Graham SJ, Van der Aa N, et al. Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. *Nat Commun.* **2016;7:11165**

Human preimplantation embryos often display chromosome mosaicism, commonly euploid-aneuploid mosaicism with complements of both normal and abnormal cells. This mosaicism occurs early in development, within the first few cell divisions following fertilization. Although preimplantation mosaicism is common, and results in high rates of early human pregnancy failures, the fate of aneuploid cells in the embryo is still unclear. In this study, the authors developed a mouse model for preimplantation chromosome mosaicism by treating developing mouse embryos with reversine during the 4- to 8-cell stage. The developing mosaic embryos were then characterized by using a combination of live-cell imaging and single-cell sequencing. The data show that aneuploid cells were eliminated from the embryo through apoptosis, starting just before implantation. Mosaic euploid-aneuploid embryos had comparable developmental potential to normal embryos, as long as they contained a sufficient proportion of euploid cells.

Illumina Technology: Nextera XT Sample Preparation Kit, HiSeq 2000/2500 Systems

Adrian J, Chang J, Ballenger CE, et al. Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Dev Cell.* **2015;33:107-118**

Plant stomata facilitate plant gas exchange with the atmosphere. In *Arabidopsis*, the production and pattern of stomata proceeds from a discrete lineage that can be parsed into intermediate steps. Despite the biological significance of ribonuclease L (RNase L), the RNAs cleaved by this enzyme are poorly defined. In this study, the authors used Illumina sequencing to reveal the frequency and location of RNase L cleavage sites within host and viral RNAs. The method was optimized and validated using viral RNAs cleaved with RNase L and RNase A, and RNA from infected and noninfected HeLa cells. The authors identified discrete genomic regions susceptible to RNase L and other single-strand–specific endoribonucleases. Monitoring the frequency and location of these cleavage sites within host and viral RNAs may reveal how these enzymes contribute to health and disease.

Illumina Technology: TruSeq SBS Kit v3–HS, HiSeq 2000 System

Burns JC, Kelly MC, Hoa M, Morell RJ and Kelley MW. Single-cell RNA-Seq resolves cellular complexity in sensory organs from the neonatal inner ear. *Nat Commun.* **2015;6:8557**

Cochlear and vestibular sensory epithelia in the inner ear use similar cell types to transduce 2 types of stimuli: sound and acceleration. However, each individual sensory epithelium is composed of anatomically and physiologically heterogeneous cell types, which have eluded transcriptional characterization due to the limited numbers of each cell type. In this study, the authors performed RNA-Seq on 301 individual cells from the utricular and cochlear sensory epithelia of newborn mice. Cluster analysis determined distinct transcriptional profiles for each cell type. Comparison of expression data from cell types within utricles and cochleae demonstrated divergence between auditory and vestibular cells.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, Nextera XT DNA Sample Preparation Index Kit, HiSeq 1000 System

72. Jungermann K and Kietzmann T. Zonation of parenchymal and nonparenchymal metabolism in liver. *Annu Rev Nutr.* 1996;16:179-203

Hashimshony T, Feder M, Levin M, Hall BK and Yanai I. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. *Nature.* 2015;519:219-222

Germ layers give rise to all of the tissues and organs in an animal, and serve as an organizing principle of developmental biology. The mesoderm is present in complex bilaterian animals but not in phyla Cnidaria and Ctenophora (comb jellies), suggesting that the mesoderm was the final germ layer to evolve. The authors used the HiSeq 2000 system to analyze the transcriptome of individual *C. elegans* blastomeres (AB, MS, E, C, and P3) that collectively account for the entire embryo. They also generated a whole-embryo time course using cell expression by linear amplification sequencing (CEL-Seq), spanning the single-cell stage to the free-living larva, at 10-minute resolution. They found that the gene expression program of *C. elegans* mesoderm was induced after those of the ectoderm and endoderm. Further, the endoderm expression program activated earlier than the endoderm program. This result was also observed for expression of endoderm orthologs in frog (*Xenopus tropicales*), sea anemone (*Nematostella vactensis*), and sponge (*Amphimedon queenslandica*). Taken together, these observations suggest that the endoderm program dates back to the origin of multicellularity, whereas the ectoderm originated as a secondary germ layer.

Illumina Technology: HiSeq 2000 System

Lodato MA, Woodworth MB, Lee S, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science.* 2015;350:94-98

Neurons are postmitotic cells, so their genomes are particularly susceptible to DNA damage. In this study, the authors surveyed the landscape of somatic single-nucleotide variants (SNVs) in the human brain by performing single-cell WGS of 36 individual cortical neurons. The most abundant SNVs included noncoding, noncoding RNA, intronic, and intergenic SNVs. Coding, truncating, splice, and silent SNVs were much less abundant. Moreover, the data showed that each cortical neuron had a distinctive genome that harbored up to 1580 somatic SNVs. Finally, the somatic SNVs created nested linkage trees, demonstrating that somatic mutations could be used to reconstruct the developmental lineage of neurons.

Illumina Technology: TruSeq Nano LT Sample Preparation Kit, MiSeq, HiSeq 2000 and HiSeq X Ten Systems

Satija R, Farrell JA, Gennert D, Schier AF and Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* 2015;33:495-502

scRNA-Seq is an established method for discovering novel cell types, understanding regulatory networks, and reconstructing developmental processes. However, scRNA-Seq typically involves dissociating cells from tissues and thus disrupting their native spatial context. To capture spatial context in scRNA-Seq data, the authors developed Seurat, a computational strategy that combines scRNA-Seq with complementary *in situ* hybridization data for a smaller set of "landmark" genes that guides spatial assignment. They validated Seurat by spatially mapping 851 individual cells from dissociated zebrafish embryos and creating a transcriptome-wide map of spatial patterning. Seurat was able to localize rare subpopulations of cells correctly, and it could map spatially restricted cells as well as those with a more scattered pattern of expression.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System

Zebrafish (*Danio rerio*).

Tohonen V, Katayama S, Vesterlund L, et al. Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development. *Nat Commun.* **2015;6:8207**

Human preimplantation development requires embryonic genome activation and maternal transcript degradation during the first 3 days after fertilization. Studies of the process at this stage are challenging technically, given the sparse samples and lack of suitable methods. To overcome these hurdles, the authors performed scRNA-Seq of 348 oocytes, zygotes, and individual blastomeres from 2- to 3-day old human embryos. They showed that 32 and 129 genes are transcribed during the transition from oocyte to 4-cell stage and from 4- to 8-cell stage, respectively. Several of the transcribed genes were unannotated PRD-like homeobox genes, including *ARGFX, CPHX1, CPHX2, DPRX, DUXA, DUXB,* and *LEUTX*.

Illumina Technology: GA$_{IIx}$, HiSeq 1000, HiSeq 2000 and MiSeq System

Zhan J, Thakare D, Ma C, et al. RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *Plant Cell.* **2015;27:513-531**

Cereal endosperm is a main source of food, feed, and raw material worldwide, yet genetic control of endosperm cell differentiation is not well defined. In this study, the authors coupled laser-capture microdissection (LCM) and Illumina sequencing to profile mRNAs in 5 major cell types of differentiating endosperms and 4 compartments of maize (*Zea mays*) kernels. They identified mRNAs that specifically accumulate in each compartment, as well as genes predominantly expressed in 1 or multiple compartments. Their results demonstrate that the MRP-1 transcription factor can activate gene regulatory networks within the basal endosperm transfer layer. These data provide a high-resolution gene activity atlas of the compartments of the maize kernel. The study also uncovers the regulatory modules associated with differentiation of the major endosperm cell types.

Illumina Technology: TruSeq DNA Sample Preparation v2 Kit, HiSeq 2000 System

Li L, Dong J, Yan L, Yong J, Liu X, et al. Single-Cell RNA-Seq Analysis Maps Development of Human Germline Cells and Gonadal Niche Interactions. *Cell Stem Cell.* 2017;20:858-873 e854

Mohammed H, Hernando-Herraez I, Savino A, et al. Single-Cell Landscape of Transcriptional Heterogeneity and Cell Fate Decisions during Mouse Early Gastrulation. *Cell Rep.* 2017;20:1215-1228

Poran A, Notzel C, Aly O, et al. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature.* 2017;551:95-99

Rizvi AH, Camara PG, Kandror EK, et al. Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nat Biotechnol.* 2017;35:551-560

Wurtzel O, Oderberg IM and Reddien PW. Planarian Epidermal Stem Cells Respond to Positional Cues to Promote Cell-Type Diversity. *Dev Cell.* 2017;40:491-504 e495

Chan SS, Chan HH and Kyba M. Heterogeneity of Mesp1+ mesoderm revealed by single-cell RNA-seq. *Biochem Biophys Res Commun.* 2016;474:469-475

Blakeley P, Fogarty NM, Del Valle I, et al. Defining the three cell lineages of the human blastocyst by single-cell RNA-seq. *Development.* 2015;142:3613

Boroviak T, Loos R, Lombard P, et al. Lineage-Specific Profiling Delineates the Emergence and Progression of Naive Pluripotency in Mammalian Embryogenesis. *Dev Cell.* 2015;35:366-382

Guo F, Yan L, Guo H, et al. The Transcriptome and DNA Methylome Landscapes of Human Primordial Germ Cells. *Cell.* 2015;161:1437-1452

Moignard V, Woodhouse S, Haghverdi L, et al. Decoding the regulatory network of early blood development from single-cell gene expression measurements. *Nat Biotechnol.* 2015;33:269-276

Pollen AA, Nowakowski TJ, Chen J, et al. Molecular identity of human outer radial glia during cortical development. *Cell.* 2015;163:55-67

Wurtzel O, Cote LE, Poirier A, Satija R, Regev A and Reddien PW. A Generic and Cell-Type-Specific Wound Response Precedes Regeneration in Planarians. *Dev Cell.* 2015;35:632-645

Immunology

The immune system consists of a number of specialized cell types that play unique roles in the host immune response. In the adaptive immune system, T and B lymphocytes (T and B cells) express specific surface receptors (T-cell receptors [TCRs] and B-cell receptors [BCRs]) that recognize and engage specific antigens presented on the surface of antigen-presenting cells via the major histocompatibility complex (MHC). Individual immune cell types are typically isolated by FACS, based on specific surface molecular markers.[73] Due to the technical limitations of FACS, FACS-isolated cells can still consist of mixed populations at various stages of development or activation, and FACS approaches are limited to available markers.[74] In contrast, single-cell sequencing approaches are not limited by specific molecular markers, and they can identify unique gene expression patterns and splice variants in T cells and B cells.[75, 76]

> "Advances in computer science and single-cell sequencing technology have created a data-driven revolution in immunology." - Neu et al. 2017

Single-cell sequencing approaches are refining our understanding of the role played by allergen-specific B cells in food allergies,[77, 78] as well as the B cell–mediated neutralizing antibody response to infectious agents.[79] Single-cell sequencing approaches are also elucidating novel mechanisms that regulate T cell differentiation and biology in human autoimmune diseases, including rheumatoid arthritis,[80] systemic lupus erythematosus (SLE),[81] multiple sclerosis,[82] type 1 diabetes mellitus,[83] and autoimmune encephalomyelitis.[84]

Antibodies that block CTLA-4 as well as PD-1 induce clinical responses in a number of cancers, including melanoma, lung cancer, renal cancer, bladder cancer, and Hodgkin's lymphoma.[85, 86] Single-cell approaches offer the possibility for a deeper understanding of the complex interactions among immune cells and tumor cells, as well as a more thorough characterization of the cellular ecosystem of tumors.[87,88,89]

73. Hackl H, Charoentong P, Finotello F and Trajanoski Z. Computational genomics tools for dissecting tumour-immune cell interactions. *Nat Rev Genet.* 2016;17:441-458

74. Newman AM and Alizadeh AA. High-throughput genomic profiling of tumor-infiltrating leukocytes. *Curr Opin Immunol.* 2016;41:77-84

75. Stubbington MJ, Lonnberg T, Proserpio V, et al. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods.* 2016;13:329-332

76. Villani AC, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science.* 2017;356:eaah4573

77. Patil SU, Ogunniyi AO, Calatroni A, et al. Peanut oral immunotherapy transiently expands circulating Ara h 2-specific B cells with a homologous repertoire in unrelated subjects. *J Allergy Clin Immunol.* 2015;136:125-134 e112

78. Hoh RA, Joshi SA, Liu Y, et al. Single B-cell deconvolution of peanut-specific antibody responses in allergic patients. *J Allergy Clin Immunol.* 2016;137:157-167

79. Tsioris K, Gupta NT, Ogunniyi AO, et al. Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. *Integr Biol (Camb).* 2015;7:1587-1597

80. Ishigaki K, Shoda H, Kochi Y, et al. Quantitative and qualitative characterization of expanded CD4+ T cell clones in rheumatoid arthritis patients. *Sci Rep.* 2015;5:12937

81. Tipton CM, Fucile CF, Darce J, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol.* 2015;16:755-765

82. Held K, Bhonsle-Deeng L, Siewert K, et al. alphabeta T-cell receptors from multiple sclerosis brain lesions show MAIT cell-related features. *Neurol Neuroimmunol Neuroinflamm.* 2015;2:e107

83. Eugster A, Lindner A, Catani M, et al. High diversity in the TCR repertoire of GAD65 autoantigen-specific human CD4+ T cells. *J Immunol.* 2015;194:2531-2538

84. Gaublomme JT, Yosef N, Lee Y, et al. Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. *Cell.* 2015;163:1400-1412

85. Wolchok JD, Hodi FS, Weber JS, et al. Development of ipilimumab: a novel immunotherapeutic approach for the treatment of advanced melanoma. *Ann N Y Acad Sci.* 2013;1291:1-13

86. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med.* 2015;372:311-319

87. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352:189-196

88. Chung W, Eum HH, Lee HO, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun.* 2017;8:15081

89. Zheng C, Zheng L, Yoo JK, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell.* 2017;169:1342-1356 e1316
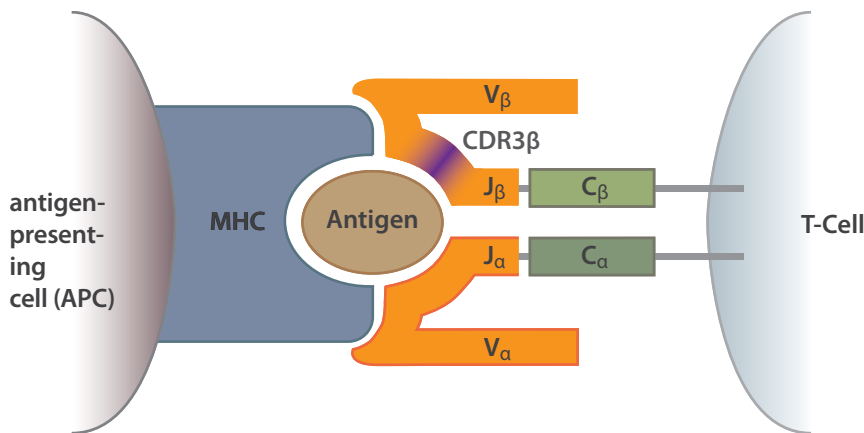
B-Cell Repertoire

Antibodies are produced by B cells in a developmentally ordered series of somatic gene rearrangements that continue throughout the life of an organism. Antibodies are composed of disulfide-linked heavy (V_H) and light (V_L) chains, which determine antigen-binding specificity. Each B cell contains a unique pair of V_H and V_L that are encoded by several distinct gene loci. To predict antibody-antigen binding specificity accurately, V_H and V_L genes should be analyzed from the same single B cell.[90]
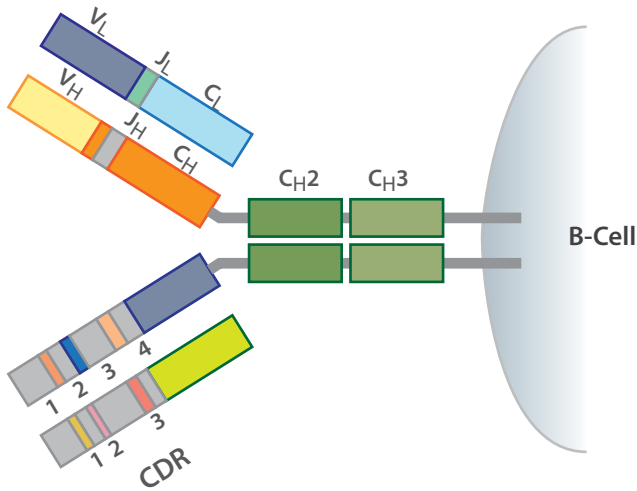
The antibody V_H repertoire is generated through somatic recombination of variable (V), diversity (D), and joining (J) gene segments. VJ recombination of V_L, V_H, and V_L heterodimeric pairing completes the antibody.

T-Cell Repertoire

Every T cell expresses unique TCRs, which are heterodimeric proteins composed of a unique combination of α and β chains. TCRs engage with peptide antigens presented by MHCs on the surface of antigen-presenting cells.[91,92] Several single-cell sequencing methods are available to sequence TCRs without disrupting α and β chain pairing through cell lysis.[93,94]

TCR-antigen-MHC interaction and TCR gene recombination. An antigen-presenting cell presents peptide antigen bound to MHC (blue). Similar to mechanisms generating antibody diversity, the TCR (orange) repertoire is generated through V(D)J recombination. TCRs bind to MHC-presented antigens. The complementarity-determining region 3 (CDR3) domain is shown in purple.[95]

90. Georgiou G, Ippolito GC, Beausang J, Busse CE, Wardemann H and Quake SR. The promise and challenge of high-throughput sequencing of the antibody repertoire. *Nat Biotechnol.* 2014;32:158-168
91. Birnbaum ME, Mendoza JL, Sethi DK, et al. Deconstructing the peptide-MHC specificity of T cell recognition. *Cell.* 2014;157:1073-1087
92. Mandl JN and Germain RN. Focusing in on T cell cross-reactivity. *Cell.* 2014;157:1006-1008
93. Woodsworth DJ, Castellarin M and Holt RA. Sequence analysis of T-cell repertoires in health and disease. *Genome Med.* 2013;5:98
94. Turchaninova MA, Britanova OV, Bolotin DA, et al. Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol.* 2013;43:2507-2515
95. Woodsworth DJ, Castellarin M and Holt RA. Sequence analysis of T-cell repertoires in health and disease. *Genome Med.* 2013;5:98

Reviews

Neu KE, Tang Q, Wilson PC and Khan AA. Single-Cell Genomics: Approaches and Utility in Immunology. *Trends Immunol.* 2017;38:140-149

Stubbington MJT, Rozenblatt-Rosen O, Regev A and Teichmann SA. Single-cell transcriptomics to explore the immune system in health and disease. *Science.* 2017;358:58-63

Lossius A, Johansen JN, Vartdal F and Holmoy T. High-throughput sequencing of immune repertoires in multiple sclerosis. *Ann Clin Transl Neurol.* 2016;3:295-306

Proserpio V and Mahata B. Single-cell technologies to study the immune system. *Immunology.* 2016;147:133-140

Schober K and Busch DH. A synergistic combination: using RNAseq to decipher both T-cell receptor sequence and transcriptional profile of individual T cells. *Immunol Cell Biol.* 2016;94:529-530

Wucherpfennig KW and Cartwright AN. Genetic screens to study the immune system in cancer. *Curr Opin Immunol.* 2016;41:55-61

References

Chung W, Eum HH, Lee HO, Lee KM, Lee HB, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun.* 2017;8:15081
Single-cell approaches have demonstrated that tumors are genetically heterogeneous. In this study, the authors used the HiSeq 2500 System to perform RNA-Seq of 515 individual cells isolated from breast cancer tumors. Single-cell transcriptomic analyses identified heterogeneous as well as conserved gene expression signatures for subtype-specific breast cancer cells. In addition, the authors identified gene expression signatures for T cells, B cells, and macrophages. Individual T cells could be classified as activated or exhausted, suggesting that immune cell interactions within tumors are dynamic and distinct. These data suggest that individual breast cancer tumor microenvironments contain cells with unique patterns of gene expression. Further, tumor subtypes are shaped by tumor cells, as well as immune cells, within the tumor microenvironment.

Illumina Technology: Nextera XT DNA Sample Prep Kit, TruSeq Rapid SBS Kit, HiSeq 2500 System

Martinez-Jimenez CP, Eling N, Chen HC, et al. Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science.* 2017;355:1433-1436
As cells age, they undergo a progressive decline in physiology and function. However, the molecular mechanisms underlying this age-dependent decline are not currently understood. In this study, the authors used single-cell RNA-Seq to uncover the effects of aging on the immune system. They used the HiSeq 2500 System to sequence the RNA of naïve and memory CD4[+] T cells in young and old mice, in both stimulated and unstimulated states. They found that older T cells exhibited much more variability in gene expression compared with younger T cells. This heightened variability was also associated with a lack of coordination of the immune response in older mice. These data suggest that age-related transcriptional variability among individual cells may be a biomarker for aging cells.

Illumina Technology: Nextera XT DNA Sample Kit, HiSeq 2500 System

Villani AC, Satija R, Reynolds G, et al. Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science.* 2017;356:eaah4573
All cellular blood components are derived from HSCs in a biological process called hematopoiesis. Among the numerous cell types that are formed, monocytes and dendritic cells are key components of the immune system. Dendritic cells process antigens and present these to T cells, thus activating the adaptive immune response. Both cell types are composed of several subtypes that have been determined through a combination of morphology, localization, and expression of surface markers. In this study, the authors performed single-cell sequencing of 2400 individual cells isolated from human blood and enriched for dendritic cells and monocytes by FACS. This unbiased strategy resulted in a reclassification of 6 dendritic cell and 4 monocyte subtypes. These data provide a more refined understanding of dendritic cells and monocytes, and also suggest a new approach to map the entire taxonomy of the immune system.

Illumina Technology: Nextera XT DNA Sample Kit, HiSeq 2500 System

Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell.* **2017;169:1342-1356 e1316**

HCC is a leading cause of cancer deaths worldwide, with the highest incidence in Asia and Africa. Despite the promise for other tumor types, cancer immunotherapies have not been successful in treating HCC, even though HCC tumors contain significant levels of tumor-infiltrating monocytes. In this study, the authors use single-cell RNA-Seq to understand and characterize tumor-infiltrating monocytes in HCC. Specifically, they isolated more than 5000 individual T cells from peripheral blood, tumor, and adjacent normal tissues from 6 HCC samples. They performed deep RNA-Seq on single T cells as well as TCR sequencing using HiSeq 2500/4000 Systems. The combination of single-cell transcriptional profiles and TCR sequences identified 11 different T-cell subtypes. Notably, the authors found that exhausted CD8$^+$ T cells and T$_{regs}$ were enriched and clonally expanded in HCC. This study demonstrates the utility of single-cell sequencing approaches in characterizing the tumor microenvironment and characteristics of tumor-infiltrating cells in HCC.

Illumina Technology: HiSeq 2500/4000 Systems

T lymphocytes (T cells) attacking cancer cell.

Kimmerling RJ, Lee Szeto G, Li JW, Genshaft AS, Kazer SW, et al. A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages. *Nat Commun.* **2016;7:10220**

Single-cell sequencing has led to an enhanced resolution in the characterization of transcriptional heterogeneity in cancer, the immune system, and in pluripotent stem cells. scRNA-Seq measurements typically rely on single time-point measurements. They provide a snapshot of tissue heterogeneity, but extending these measurements to serial time points could potentially enhance our understanding of the mechanisms for generating tissue heterogeneity. In this study, the authors developed a microfluidic platform that enabled off-chip scRNA-Seq after multigenerational lineage tracking. They used the platform to collect single-cell transcriptional data for lineages of mouse lymphocytic leukemia cells, as well as primary murine CD8+ T cells. Their results reveal transcriptional signatures for each cell type that depend on lineage and cell cycle.

Illumina Technology: NextSeq 500 System

Stubbington MJ, Lonnberg T, Proserpio V, et al. T cell fate and clonality inference from single-cell transcriptomes. *Nat Methods.* **2016;13:329-332**

TCRs are able to recognize antigen with a high degree of specificity. This is due, in part, to the high degree of genetic heterogeneity of TCRs produced by the recombination of V(D)J loci during T-cell development. In this study, the authors developed TraCeR, a computational method that allowed them to reconstruct full-length paired TCR sequences from scRNA-Seq data. The ability to interrogate recombined TCR sequences in the context of scRNA-Seq data allowed the authors to link T-cell specificity with functional response, by revealing clonal relationships between cells and their transcriptional profiles.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System

---

For Research Use Only. Not for use in diagnostic procedures.

Tirosh I, Izar B, Prakadan SM, Wadsworth MH, 2nd, Treacy D, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352:189-196

Single-cell sequencing enables the detailed evaluation of genetic and transcriptional features present in individual cells within tumors. The technology has the potential to inform treatment response and drug resistance by assessing malignant, microenvironmental, and immunologic states within tumors. In this study, the authors applied scRNA-Seq to 4645 single cells (malignant, stromal, immune, and endothelial) isolated from 19 patients with metastatic melanoma. They found that malignant cells within the same tumor displayed transcriptional heterogeneity associated with cell cycle, spatial context, and drug resistance. The same tumor had cells with high expression levels of MITF, as well as cells with low MITF levels and elevated levels of AXL kinase (cells prone to early drug resistance). Infiltrating T-cell analysis revealed exhaustion programs, connection to T-cell activation/expansion, and patient variability. This study demonstrates how single-cell genomics can unravel the cellular ecosystem of tumors, with implications for targeted and immune therapies.

Illumina Technology: Nextera XT Sample Preparation Kit, NextSeq 500 System

Avraham R, Haseley N, Brown D, et al. Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses. *Cell.* 2015;162:1309-1321

The interactions between immune cells and invading pathogens determine the course of infection. Bulk cell-sequencing approaches can mask the heterogeneous, stochastic, and dynamic nature of the cell-pathogen interactions. In this study, the authors used the HiSeq 2500 system to perform scRNA-Seq with fluorescent markers to probe the responses of 150 individual macrophages to invading strains of *Salmonella* bacteria. Their data showed that variable PhoPQ activity, which upregulates *Salmonella* virulence factors in infecting *Salmonella*, drove variable host cell type I interferon (IFN) responses by modifying lipopolysaccharides (LPS) in a subset of bacteria. The results suggest that functional heterogeneity in the host cell response to infection is linked to cell-to-cell variations in the population of infecting pathogens.

Illumina Technology: HiSeq 2500 System

Fan HC, Fu GK and Fodor SP. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science.* 2015;347:1258367

Combinatorial labeling of single cells is rapid and relatively inexpensive, and it can boost the throughput of massively parallel single-cell sequencing approaches dramatically. In this study, the authors developed CytoSeq, a method to label large numbers of individual cells combinatorially. Individual cells are placed in single wells, along with combinatorial libraries of beads containing cell- and transcript-barcoding probes. The authors performed CytoSeq on human peripheral blood mononuclear cells (PBMCs) and used the MiSeq system to sequence amplified cDNAs. They analyzed several genes and were able to identify major subsets of PBMCs. In addition, by comparing cellular heterogeneity in naïve and cytomegalovirus (CMV)-activated CD8+ T cells, they identified rare cells specific to the CMV antigen. CytoSeq can be applied to complex mixtures of cells of varying size and shape, as well as to other biomolecules.

Illumina Technology: MiSeq System

Gaublomme JT, Yosef N, Lee Y, et al. Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. *Cell.* 2015;163:1400-1412

Autoimmune encephalomyelitis (EAE) in mice is widely studied as an animal model of human CNS demyelinating diseases, including multiple sclerosis and acute disseminated encephalomyelitis. IL-17-producing Th17 cells are a critically important part of the adaptive immune system but are also implicated in the pathogenesis of autoimmunity. In this study, the authors used the HiSeq 2000/2500 system to perform RNA-Seq of 976 individual Th17 cells isolated from CNS and lymph nodes of mice with EAE. Computational analysis of scRNA-Seq data revealed the marked genetic heterogeneity of Th17 cells and related them to a spectrum of Th17 cells spanning regulatory to pathogenic functional states. The authors identified and validated[96] four genes (*Grp65, Plzp, Toso*, and *Cd5l*) that regulate Th17 pathogenicity, suggesting possible new drug targets in autoimmunity.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2000/2500 Systems

96. Wang C, Yosef N, Gaublomme J, et al. CD5L/AIM Regulates Lipid Biosynthesis and Restrains Th17 Cell Pathogenicity. *Cell.* 2015;163:1413-1427
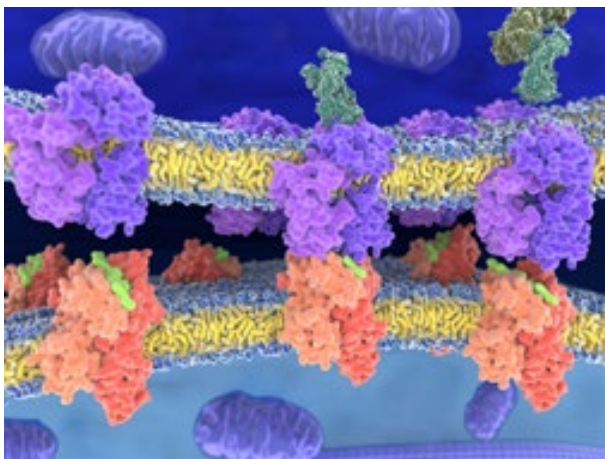
TCRs interact with MHC class II antigen complexes.

Tipton CM, Fucile CF, Darce J, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol.* **2015;16:755-765**
Acute SLE is a recurrent autoimmune disease that attacks various tissues and has no cure. Autoimmune activity is associated with surges in B cells. The only SLE therapy approved by the US Food and Drug Administration (FDA), belimumab, targets B-cell activating factor. In this study, the authors isolated B cells from patients experiencing SLE flares and used deep sequencing and proteomic approaches to analyze the diversity of B cells. They demonstrated that B cells from patients with SLE flares were polyclonal. By sequencing single B cells, they also identified a subpopulation of newly activated naïve B cells that provided an important source of autoantibodies during SLE flares, suggesting that SLE autoreactivities occur during polyclonal activation. These results may guide patient treatment options and facilitate the design of future SLE therapies.

Illumina Technology: MiSeq System

Carmona SJ, Teichmann SA, Ferreira L, et al. Single-cell transcriptome analysis of fish immune cells provides insight into the evolution of vertebrate immune cell types. *Genome Res.* 2017;27:451-461

Engblom C, Pfirschke C, Zilionis R, Da Silva Martins J, Bos SA, et al. Osteoblasts remotely supply lung tumors with cancer-promoting SiglecF(high) neutrophils. *Science.* 2017;358:eaal5081

Haber AL, Biton M, Rogel N, et al. A single-cell survey of the small intestinal epithelium. *Nature.* 2017;551:333-339

Medaglia C, Giladi A, Stoler-Barak L, et al. Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. *Science.* 2017;358:1622-1626

Mildner A, Schonheit J, Giladi A, et al. Genomic Characterization of Murine Monocytes Reveals C/EBPbeta Transcription Factor Dependence of Ly6C- Cells. *Immunity.* 2017;46:849-862 e847

Wallrapp A, Riesenfeld SJ, Burkett PR, et al. The neuropeptide NMU amplifies ILC2-driven allergic lung inflammation. *Nature.* 2017;549:351-356

Bjorklund AK, Forkel M, Picelli S, et al. The heterogeneity of human CD127 innate lymphoid cells revealed by single-cell RNA sequencing. *Nat Immunol.* 2016;17:451-460

Brennecke P, Reyes A, Pinto S, et al. Single-cell transcriptome analysis reveals coordinated ectopic gene-expression patterns in medullary thymic epithelial cells. *Nat Immunol.* 2015;16:933-941

DeKosky BJ, Kojima T, Rodin A, et al. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med.* 2015;21:86-91

Eugster A, Lindner A, Catani M, et al. High diversity in the TCR repertoire of GAD65 autoantigen-specific human CD4+ T cells. *J Immunol.* 2015;194:2531-2538

For Research Use Only. Not for use in diagnostic procedures.

Neurobiology

Among cells comprising specific brain regions, single-cell sequencing approaches have discovered genomic mosaicism in individual neurons, including CNVs and somatic SNVs.[97, 98] The genetic variations underlying this genomic mosaicism in the brain arise during fetal development,[99] and while their functional relevance is not yet fully clear, recent data suggest that aging and neurodegeneration are associated with somatic mutations in neurons.[100] Recent single-cell sequencing studies have also identified high rates of somatic LINE-1 element (L1) retrotransposition in the hippocampus and cerebral cortex that could have major implications for normal brain function;[101] however, other studies have determined that rates of L1 retrotransposition in the brain are lower than first reported.[102]

> "Single-cell RNA sequencing is an exciting new technology allowing the analysis of transcriptomes from individual cells, and is ideally suited to address the inherent complexity and dynamics of the central nervous system." - Ofengeim et al. 2017

In addition to CNS genomic diversity, recent studies have extended our understanding of CNS transcriptome diversity at the single-cell level from both adult brain as well as 3-dimensional brain organoids.[103,104,105,106] Single-cell transcriptomics has identified mechanisms regulating neurodevelopment,[107] and scRNA-Seq studies have recently begun to unravel new biological details of sensory neurons,[108] glial cells,[109] interneurons,[110] and other cell types in the brain.[111] New technical achievements in single-cell sequencing combine scRNA-Seq with electrophysiological recording of individual neurons,[112,113] as well as characterizing gene expression patterns associated with experience-driven induction of activity in individual hippocampal neurons.[114] Finally, single-cell analysis of 6000 methylomes from individual neuronal nuclei recently identified neuronal subtypes in the human frontal cortex.[115]
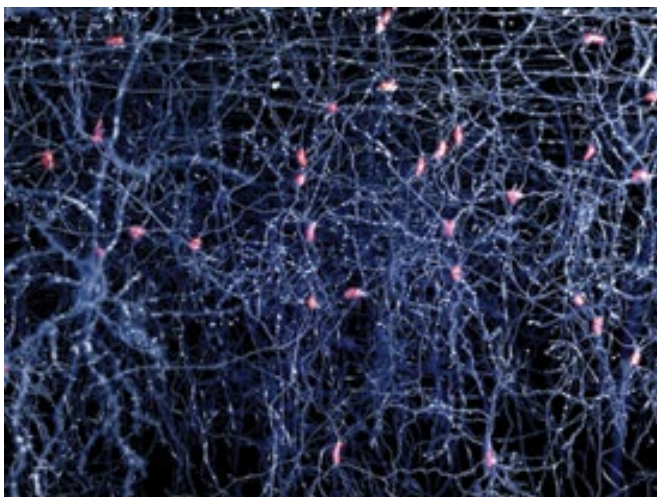
Single-cell sequencing approaches uncover genetic mosaicism in neurons.

97. McConnell MJ, Lindberg MR, Brennand KJ, et al. Mosaic copy number variation in human neurons. *Science.* 2013;342:632-637

98. Lodato MA, Woodworth MB, Lee S, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science.* 2015;350:94-98

99. Bae T, Tomasini L, Mariani J, et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science.* 2018;359:550-555

100. Lodato MA, Rodin RE, Bohrson CL, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science.* 2018;359:555-559

101. Upton KR, Gerhardt DJ, Jesuadian JS, et al. Ubiquitous L1 mosaicism in hippocampal neurons. *Cell.* 2015;161:228-239

102. Evrony GD, Lee E, Park PJ and Walsh CA. Resolving rates of mutation in the brain using single-neuron genomics. *Elife.* 2016;5:e12966

103. Spaethling JM, Na YJ, Lee J, et al. Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics. *Cell Rep.* 2017;18:791-803

104. Hanchate NK, Kondoh K, Lu Z, et al. Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science.* 2015;350:1251-1255

105. Quadrato G, Nguyen T, Macosko EZ, et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature.* 2017;545:48-53

106. Bershteyn M, Nowakowski TJ, Pollen AA, et al. Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell Stem Cell.* 2017;20:435-449 e434

107. Luo Y, Coskun V, Liang A, et al. Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell.* 2015;161:1175-1186

108. Saraiva LR, Ibarra-Soria X, Khan M, et al. Hierarchical deconstruction of mouse olfactory sensory neurons: from whole mucosa to single-cell RNA-seq. *Sci Rep.* 2015;5:18178

109. Pollen AA, Nowakowski TJ, Chen J, et al. Molecular identity of human outer radial glia during cortical development. *Cell.* 2015;163:55-67

110. Frazer S, Prados J, Niquille M, et al. Transcriptomic and anatomic parcellation of 5-HT3AR expressing cortical interneuron subtypes revealed by single-cell RNA sequencing. *Nat Commun.* 2017;8:14219

111. Zeisel A, Munoz-Manchado AB, Codeluppi S, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015;347:1138-1142

112. Fuzik J, Zeisel A, Mate Z, et al. Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat Biotechnol.* 2016;34:175-183

113. Cadwell CR, Palasantza A, Jiang X, et al. Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat Biotechnol.* 2016;34:199-203

114. Lacar B, Linker SB, Jaeger BN, et al. Corrigendum: Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat Commun.* 2016;7:12020

115. Luo C, Keown CL, Kurihara L, et al. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science.* 2017;357:600-604

Reviews

Cuevas-Diaz Duran R, Wei H and Wu JQ. Single-cell RNA-sequencing of the brain. *Clin Transl Med.* 2017;6:20

Lein E, Borm LE and Linnarsson S. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science.* 2017;358:64-69

Ofengeim D, Giagtzoglou N, Huh D, Zou C and Yuan J. Single-Cell RNA Sequencing: Unraveling the Brain One Cell at a Time. *Trends Mol Med.* 2017;23:563-576

Harbom LJ, Chronister WD and McConnell MJ. Single neuron transcriptome analysis can reveal more than cell type classification: Does it matter if every neuron is unique? *Bioessays.* 2016;38:157-161

Bae BI, Jayaraman D and Walsh CA. Genetic changes shaping the human brain. *Dev Cell.* 2015;32:423-434

116. Lodato MA, Woodworth MB, Lee S, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science.* 2015;350:94-98
117. Tasic B, Menon V, Nguyen TN, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci.* 2016;19:335-346
118. Zeisel A, Munoz-Manchado AB, Codeluppi S, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015;347:1138-1142

References

Bae T, Tomasini L, Mariani J, Zhou B, Roychowdhury T, et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science.* **2018;359:550-555**
Somatic mutations are permanent changes in the DNA sequence of a genome. They arise because of mistakes in DNA replication in cells or from environmental factors. Previous studies using single-cell sequencing approaches have demonstrated that individual neurons contain unique SNV somatic mutations, a phenomenon referred to as genomic mosaicism.[116] Although genomic mosaicism is common among multiple cell types, neurons are postmitotic, and thus the SNVs have the potential to persist and alter neuronal function. In this study, the authors isolated individual neurons from the forebrains of embryonic samples and grew them into small clones in culture. They used the HiSeq platform to analyze the genomes of these samples, and they identified 200–400 SNVs per cell. The SNV patterns in these embryonic neurons resembled those found in cancer cell genomes. By analyzing the genomes of neurons at different stages of embryonic development, they determined that the mutation rate speeds up during neurogenesis. These data suggest that embryonic neurons are protected from mutations, compared to neurons at later stages of neurodevelopment.

Illumina Technology: HiSeq 2500 and HiSeq X Systems

Luo C, Keown CL, Kurihara L, et al. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science.* **2017;357:600-604**
Single-cell RNA-Seq approaches have been useful in defining diverse neuronal subtypes within the brain.[117,118] In this study, the authors developed a technique to characterize the epigenomic diversity within individual neuronal cells. This technique, single-nucleotide methylcytosine sequencing (snmC-seq), takes advantage of the significant amount of 5-methylcytosine found in neurons. The authors isolated 3400 individual neurons from mouse cortex and collected nuclei by FACS. They treated the nuclear DNA treated with bisulfite to convert unmethylated cytosine into uracil, while 5-methyl- and 5-hydroxymethylcytosines remained unaffected. Finally, they pooled, amplified, and sequenced the samples using the HiSeq 4000 System. The authors used snmC-seq to generate 6000 methylomes from the isolated neuronal nuclei, and their data identified 16 mouse neuronal subtypes in mouse brain. They also performed snmC-seq on isolated nuclei from human cortex samples, and they identified 21 human neuronal subtypes. The single-cell methylome technique thus expanded the taxonomy of brain cell diversity. Moreover, it identified regulatory elements associated with neuronal diversity.

Illumina Technology: TruSeq Methylation Kit, HiSeq 4000 System

Lodato MA, Rodin RE, Bohrson CL, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science.* **2018;359:555-559**
Aging is associated with increased somatic mutations in cells, as errors in replication accrue. It is unclear whether this is true for postmitotic neurons that do not divide. In this study the authors isolated 161 individual postmitotic neurons from individuals aged 4 months to 82 years. They performed single-cell WGS of each individual neuron to identify SNVs. Their data demonstrated that SNVs increased linearly with increasing age. They also isolated individual neurons from 8 individuals with early-onset neurodegenerative diseases. Single-cell WGS of these neurons indicated SNVs were more abundant than in normal neurons. Overall, these data suggest that age-related accumulation of SNVs in neurons may be important in neuronal diseases.

Illumina Technology: TruSeq Nano LT Sample Preparation Kit, HiSeq 2000 and HiSeq X10 Systems

Quadrato G, Nguyen T, Macosko EZ, et al. Cell diversity and network dynamics in photosensitive human brain organoids. *Nature.* **2017;545:48-53**

The brain is an enormously complex organ, and 2-dimensional tissue cultures have limited utility in understanding brain function. Recently, 3-dimensional tissue culture brain organoids have been developed in the hope of understanding brain biology, disease, and development.[119] In this study, the authors attempted to characterize the extent to which brain organoids mimicked the complexity and cellular diversity of whole brains. They performed single-cell RNA-Seq of over 80,000 individual cells isolated from 31 human brain organoids. The singe-cell gene expression signatures demonstrated that human brain organoids developed and generated a wide diversity of cell types, including subtypes related to adult cortical neurons as well as retinal cells. Moreover, the neuronal connectivity and synapses were present within cells of the organoid. Finally, the organoids developed neuronal networks spontaneously, and photosensitive cells were even sensitive to light stimulation. Overall, the data suggest that human brain organoids have enormous potential as *in vitro* models of the human brain.

Illumina Technology: Nextera XT DNA Library Preparation Kit, NextSeq 500 System

Human brain organoids spontaneously develop neuronal networks.

Cadwell CR, Palasantza A, Jiang X, et al. Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nat Biotechnol.* **2016;34:199-203**

In this study the authors describe Patch-Seq, a method that combines whole-cell electrophysiological characterization, morphological characterization, and scRNA-Seq. They performed Patch-Seq on 58 neuronal cells from layer 1 of the mouse neocortex. After they characterized individual neurons electrophysiologically, they aspirated the cell contents through the patch-clamp pipette and prepared them for RNA-Seq. The authors classified cells based on electrophysiology and morphology, as well as by patterns of gene expression. Their data show that gene expression patterns could infer axonal arborization and the action potential amplitude of individual neurons.

Illumina Technology: HiSeq 2000 System

Lacar B, Linker SB, Jaeger BN, et al. Corrigendum: Nuclear RNA-seq of single neurons reveals molecular signatures of activation. *Nat Commun.* **2016;7:12020**

scRNA-Seq has been a key method for the subclassification of cells that would otherwise be indistinguishable based solely on morphology or anatomy. Profiling the transcriptome of individual neurons in response to activation is important for characterizing brain function. In this study, the authors performed RNA-Seq on isolated nuclei (snRNA-Seq) from individual mouse neurons taken from the dentate gyrus of the hippocampus and stimulated by pentylenetetrazole. There were large-scale changes in the activated neuronal transcriptome, including induction of mitogen-activated protein kinase (MAPK) pathway genes, after brief, novel environment exposure. The data show that snRNA-Seq of activated neurons allows for analysis of gene expression beyond immediate early genes.

Illumina Technology: HiSeq 2500 System

119.   Kelava I and Lancaster MA. Stem Cell Models of Human Brain Development. *Cell Stem Cell.* 2016;18:736-748

Tasic B, Menon V, Nguyen TN, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci.* **2016;19:335-346**
Given the complexity of the human brain, understanding the genetic and functional diversity of individual cells is of paramount importance. In this study, the authors performed scRNA-Seq on more than 1600 individual cells from the mouse primary visual cortex. Their data analysis identified 49 unique transcriptomic cell types. A subset of these transcriptomic cell types displayed specific and differential electrophysiological and axon projection properties, which confirms that single-cell transcriptomic signatures can be linked to specific cellular phenotypes.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq and HiSeq 2000/2500 Systems

Evrony GD, Lee E, Mehta BK, et al. Cell lineage analysis in human brain using endogenous retroelements. *Neuron.* **2015;85:49-59**
Postmitotic, somatic mutations are known to cause cancer, but these mutations may also lead to diverse neurological diseases, including cortical malformations, epilepsy, intellectual disability, and neurodegeneration. Studying pathogenic somatic mutations is challenging due to the variety of ways that their effects are shaped by normal development. The overall pattern of somatic mutation distribution in the human brain is not well characterized. In this study, the authors used the HiSeq 2000 system to perform high-coverage WGS of individual neurons. Somatic mutation analyses of individual neurons from several CNS locations identified multiple cell lineages in the brain, marked by different L1 retrotransposition events. The patterns of somatic mutations mirrored known somatic mutation disorders of brain development, suggesting that focally distributed somatic mutations are also present in normal brains.

Illumina Technology: HiSeq 2000 System

Hanchate NK, Kondoh K, Lu Z, et al. Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science.* **2015;350:1251-1255**
In mammals, odor detection is mediated by G protein–coupled olfactory receptors on neurons in the nasal olfactory epithelium. Mature neurons typically express a single olfactory receptor per neuron. In this study, the authors used the HiSeq 2500 system to perform scRNA-Seq of single epithelial neurons during mouse development, with multiple cells from each stage of development sequenced. The single-cell data confirmed that most neurons expressed high levels of only 1 olfactory receptor. However, many immature neurons expressed multiple olfactory receptors at low levels, with a single neuron capable of expressing olfactory receptors from up to 7 different chromosomes. The data show that developmental pathways ultimately restrict olfactory receptor expression in mature neurons.

Illumina Technology: TruSeq DNA Sample Preparation Kit, HiSeq 2500 System

Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K and Martin-Villalba A. Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. *Cell Stem Cell.* **2015;17:329-340**
Within the human brain, pools of adult NSCs participate in brain maintenance and regeneration following injury. The balance of activation and quiescence of NSCs depends on the induction of specific transcription factors. In this study, the authors used the HiSeq 2000 system to perform scRNA-Seq of NSCs isolated from the brain subventricular zone. They identified the expression of lineage-specific transcription factors in a specific subpopulation of dormant NSCs. They also discovered that brain ischemic injury induced interferon signaling in dormant NSCs, promoting their entry into a primed-quiescent state. This study unveils the general molecular principles underlying NSC activation and suggests potential avenues for brain regenerative medicine.

Illumina Technology: Nextera XT Sample Preparation Kit, HiSeq 2000 System

Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science.* **2015;350:94-98**
Neurons are postmitotic cells, so their genomes are particularly susceptible to DNA damage. In this study, the authors surveyed the landscape of somatic SNVs in the human brain by performing single-cell WGS of 36 individual cortical neurons. The most abundant SNVs included noncoding, noncoding RNA, intronic, and intergenic SNVs. Coding, truncating, splice, and silent SNVs were much less abundant. Moreover, the data showed that each cortical neuron had a distinctive genome that harbored up to 1580 somatic SNVs. Finally, the somatic SNVs created nested linkage trees, demonstrating that somatic mutations could be used to reconstruct the developmental lineage of neurons.

Illumina Technology: TruSeq Nano LT Sample Preparation Kit, MiSeq, HiSeq 2000 and HiSeq X Ten Systems

Luo Y, Coskun V, Liang A, Yu J, Cheng L, et al. Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. *Cell.* 2015;161:1175-1186

The scarcity of tissue-specific stem cells, and the complexity of their surrounding environment, make single-cell sequencing methods imperative for characterizing these cell types. In this study, the authors used single-cell sequencing and weighted gene coexpression network analysis to identify CD133+ ependymal cells from the adult mouse forebrain neurogenic zone as NSCs. These subpopulations of cells were enriched for immune-responsive genes, as well as genes encoding angiogenic factors. Administration of VEGF and bFGF enhanced migration and elicited differentiation into neurons and glia.

Illumina Technology: HiSeq 2500 System

Upton KR, Gerhardt DJ, Jesuadian JS, et al. Ubiquitous L1 mosaicism in hippocampal neurons. *Cell.* 2015;161:228-239

Somatic L1 retrotransposition occurs during neurogenesis and serves as a potential source of genotypic variation among neurons. Pronounced L1 activity is present in the hippocampus, but its biological consequence is unclear. In this study, the authors used the MiSeq system to perform single-cell retrotransposon capture sequencing (RC-Seq) on individual hippocampal neurons, hippocampal glial cells, and cortical neurons. These experiments established that L1-driven mosaicism is abundant in the hippocampus with 13.7 somatic L1 insertions per hippocampal neuron, compared to previous estimates of < 0.1 L1 insertions per cortical neuron. Given the abundance of L1 insertions, it is reasonable to assume that L1-driven mosaicism may alter brain biology and function.

Illumina Technology: MiSeq System

Zeisel A, Munoz-Manchado AB, Codeluppi S, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015;347:1138-1142

The brain is built from a large number of cell types, of which neurons are a subset. To understand brain function, it is important to characterize this complex collection of highly specialized cells. In this study, the authors performed quantitative scRNA-Seq based on 3005 single-cell transcriptomes from the primary somatosensory cortex and hippocampal CA1 regions. RNA molecules were counted using unique molecular identifiers (UMIs). Biclustering methods revealed 9 major classes of cells, including pyramidal neurons, interneurons, oligodendrocytes, astrocytes, microglia, vascular endothelial cells, mural cells, and ependymal cells. The researchers identified 47 molecularly distinct subclasses within the 9 cell types. They also identified unique marker genes that aligned with cell type, morphology, and location.

Illumina Technology: HiSeq 2000 System
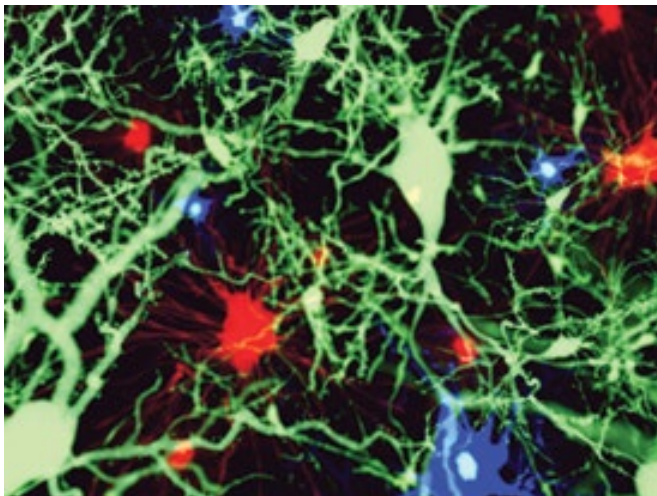
Diversity of brain cells.

Bershteyn M, Nowakowski TJ, Pollen AA, et al. Human iPSC-Derived Cerebral Organoids Model Cellular Features of Lissencephaly and Reveal Prolonged Mitosis of Outer Radial Glia. *Cell Stem Cell.* 2017;20:435-449 e434

Campbell JN, Macosko EZ, Fenselau H, et al. A molecular census of arcuate hypothalamus and median eminence cell types. *Nat Neurosci.* 2017;20:484-496

Frazer S, Prados J, Niquille M, et al. Transcriptomic and anatomic parcellation of 5-HT3AR expressing cortical interneuron subtypes revealed by single-cell RNA sequencing. *Nat Commun.* 2017;8:14219

Karlsson K and Linnarsson S. Single-cell mRNA isoform diversity in the mouse brain. *BMC Genomics.* 2017;18:126

Nowakowski TJ, Bhaduri A, Pollen AA, et al. Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science.* 2017;358:1318-1323

Spaethling JM, Na YJ, Lee J, et al. Primary Cell Culture of Live Neurosurgically Resected Aged Adult Human Brain Cells and Single Cell Transcriptomics. *Cell Rep.* 2017;18:791-803

Crocker A, Guan XJ, Murphy CT and Murthy M. Cell-Type-Specific Transcriptome Analysis in the Drosophila Mushroom Body Reveals Memory-Related Changes in Gene Expression. *Cell Rep.* 2016;15:1580-1596

Evrony GD, Lee E, Park PJ and Walsh CA. Resolving rates of mutation in the brain using single-neuron genomics. *Elife.* 2016;5:e12966

Finegersh A and Homanics GE. Chromatin immunoprecipitation and gene expression analysis of neuronal subtypes after fluorescence activated cell sorting. *J Neurosci Methods.* 2016;263:81-88

Fuzik J, Zeisel A, Mate Z, et al. Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nat Biotechnol.* 2016;34:175-183

Habib N, Li Y, Heidenreich M, et al. Div-Seq: A single nucleus RNA-Seq method reveals dynamics of rare adult newborn neurons in the CNS. *Science.* 2016;353:925-928

Krishnaswami SR, Grindberg RV, Novotny M, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524

Li CL, Li KC, Wu D, et al. Somatosensory neuron types identified by high-coverage single-cell RNA-sequencing and functional heterogeneity. *Cell Res.* 2016;26:83-102

Milani P, Escalante-Chong R, Shelley BC, Patel-Murray NL, Xin X, et al. Cell freezing protocol suitable for ATAC-Seq on motor neurons derived from human induced pluripotent stem cells. *Sci Rep.* 2016;6:25474

Thomsen ER, Mich JK, Yao Z, et al. Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat Methods.* 2016;13:87-93

Drug Discovery

Modern drug discovery involves the identification of novel drug targets that are highly relevant to a disease process, followed by the development of highly specific, target-based therapeutic interventions. In a given tissue, cells may express up to 400 druggable genes, suggesting a relative abundance of drug targets.[120] However, tissues are composed of a variety of genetically diverse cell types, and the resultant biological noise in bulk sequencing approaches can drown out the specific genes of interest. Single-cell sequencing approaches show great promise in discovering and characterizing drug targets, and in improving pharmaceutical screens.[121,122,123]

> "The emergence of single-cell omics tools over the past 5 years has happened at a lightning pace, and the potential for their use in discovery and development of broad classes of therapies and therapeutic strategies is high." - Health et al. 2016

Single-cell sequencing of tumor cells is enhancing our understanding of tumor heterogeneity, drug response, and resistance mechanisms in cancer.[124,125] Similarly, single-cell analysis of circulating tumor cells highlights their genetic heterogeneity, therapeutic sensitivity, and metastatic potential.[126,127, 128]

Recent clinical data have demonstrated that therapeutic enhancement of immune system function can benefit patients with cancer.[129] Antibodies that block CTLA-4 as well as PD-1 induce clinical responses in a number of cancers, including melanoma, lung cancer, renal cancer, bladder cancer, and Hodgkin's lymphoma.[130,131] Single-cell sequencing approaches offer the possibility for a deeper understanding of the complex interactions among immune cells and tumor cells, as well as a more thorough characterization of the cellular ecosystem of tumors.[132,133,134]

In addition to uncovering mechanisms of tumor-host immune interactions, single-cell approaches also uncover host immune responses to infectious, pathogenic microorganisms.[135,136] Our understanding and characterization of microbial diversity continues to expand, and single-cell sequencing can help in characterizing the abundance of infectious pathogens in our environment.[137] Finally, genomic characterization of uncultivatable microbes has the potential for the discovery of novel natural products that could be used as new chemical scaffolds in pharmaceutical development.[138]

Single-cell sequencing approaches are refining our understanding of allergen-specific B cells in food allergies,[139,140] as well as the B cell–mediated neutralizing antibody response to infectious agents.[141] Single-cell sequencing approaches are also elucidating novel mechanisms regulating T-cell differentiation and biology in human autoimmune diseases, including rheumatoid arthritis,[142] SLE,[143] multiple sclerosis,[144] type 1 diabetes mellitus,[145] and neuroinflammation.[146,147]

120. Eberwine J, Sul J-Y, Bartfai T and Kim J. The promise of single-cell sequencing. *Nature Methods.* 2013;11:25-27

121. Bartfai T, Buckley PT and Eberwine J. Drug targets: single-cell transcriptomics hastens unbiased discovery. *Trends Pharmacol Sci.* 2012;33:9-16

122. Weaver WM, Tseng P, Kunze A, et al. Advances in high-throughput single-cell microtechnologies. *Curr Opin Biotechnol.* 2014;25:114-123

123. Heath JR, Ribas A and Mischel PS. Single-cell analysis tools for drug discovery and development. *Nat Rev Drug Discov.* 2016;15:204-216

124. Gerber T, Willscher E, Loeffler-Wirth H, et al. Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. *Oncotarget.* 2017;8:846-862

125. Zhang X, Marjani SL, Hu Z, Weissman SM, Pan X and Wu S. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Res.* 2016;76:1305-1312

126. Shaffer SM, Dunagin MC, Torborg SR, et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature.* 2017;546:431-435

127. Zhang C, Guan Y, Sun Y, Ai D and Guo Q. Tumor heterogeneity and circulating tumor cells. *Cancer Lett.* 2016;374:216-223

128. Chung W, Eum HH, Lee HO, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun.* 2017;8:15081

129. Wucherpfennig KW and Cartwright AN. Genetic screens to study the immune system in cancer. *Curr Opin Immunol.* 2016;41:55-61

130. Wolchok JD, Hodi FS, Weber JS, et al. Development of ipilimumab: a novel immunotherapeutic approach for the treatment of advanced melanoma. *Ann N Y Acad Sci.* 2013;1291:1-13

131. Ansell SM, Lesokhin AM, Borrello I, et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med.* 2015;372:311-319

132. Tirosh I, Izar B, Prakadan SM, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science.* 2016;352:189-196

133. Chung W, Eum HH, Lee HO, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun.* 2017;8:15081

134. Zheng C, Zheng L, Yoo JK, et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell.* 2017;169:1342-1356 e1316

135. Combe M, Garijo R, Geller R, Cuevas JM and Sanjuan R. Single-Cell Analysis of RNA Virus Infection Identifies Multiple Genetically Diverse Viral Genomes within Single Infectious Units. *Cell Host Microbe.* 2015;18:424-432

136. Lonnberg T, Svensson V, James KR, et al. Single-cell RNA-seq and computational analysis using temporal mixture modelling resolves Th1/Tfh fate bifurcation in malaria. *Sci Immunol.* 2017;2:eaal2192

137. McLean JS and Lasken RS. Single cell genomics of bacterial pathogens: outlook for infectious disease research. *Genome Med.* 2014;6:108

138. Rashid M and Stingl U. Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. *Biotechnol Adv.* 2015;33:1755-1773
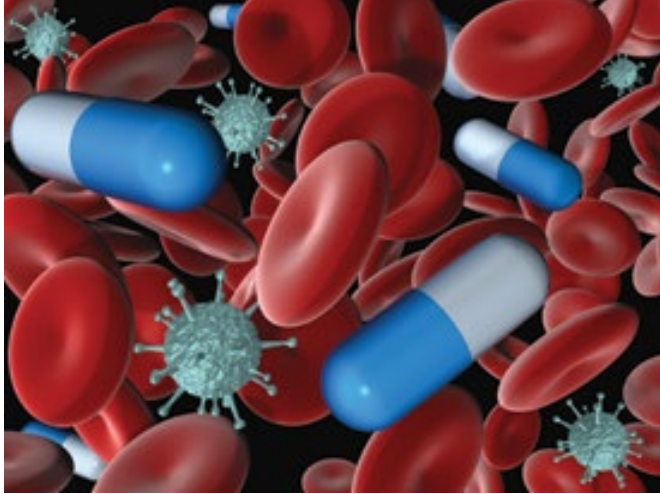
Single-cell genomics promises much greater efficiency in drug development and design.

Reviews

DiLoreto R, Khush K and De Vlaminck I. Precision monitoring of immunotherapies in solid organ and hematopoietic stem cell transplantation. *Adv Drug Deliv Rev.* 2017;114:272-284

McConnell MJ, Moran JV, Abyzov A, et al. Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. *Science.* 2017;356:eaal1641

Heath JR, Ribas A and Mischel PS. Single-cell analysis tools for drug discovery and development. *Nat Rev Drug Discov.* 2016;15:204-216

Wucherpfennig KW and Cartwright AN. Genetic screens to study the immune system in cancer. *Curr Opin Immunol.* 2016;41:55-61

Zhang X, Marjani SL, Hu Z, Weissman SM, Pan X, et al. Single-Cell Sequencing for Precise Cancer Research: Progress and Prospects. *Cancer Res.* 2016;76:1305-1312

Alizadeh AA, Aranda V, Bardelli A, et al. Toward understanding and exploiting tumor heterogeneity. *Nat Med.* 2015;21:846-853

Issigonis M and Newmark PA. Heal Thy Cell(f): A Single-Cell View of Regeneration. *Dev Cell.* 2015;35:527-528

Robinson WH. Sequencing the functional antibody repertoire--diagnostic and therapeutic discovery. *Nat Rev Rheumatol.* 2015;11:171-182

References

Keren-Shaul H, Spinrad A, Weiner A, et al. A Unique Microglia Type Associated with Restricting Development of Alzheimer's Disease. *Cell.* 2017;169:1276-1290 e1217
Alzheimer's disease is a chronic neurodegenerative disease with no known cure. Its incidence increases with age, with approximately 6% of people age 65 and older affected by the disease.[148] To better understand the cellular etiology of Alzheimer's disease, the authors performed single-cell RNA-Seq of approximately 8000 individual immune cells from healthy transgenic mice as well as those with Alzheimer's disease. The data revealed a novel microglial cell that was associated with neurodegenerative disease. These microglia were found to contain Aβ particles, a molecular signature of Alzheimer's disease. The data suggest that this novel immune cell type has the potential to restrict neurodegeneration, with potential implications for treatment of Alzheimer's disease.

Illumina Technology: NextSeq 500 System

Shaffer SM, Dunagin MC, Torborg SR, et al. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature.* 2017;546:431-435
Many tumors can be successfully treated using chemotherapy, but they eventually develop resistance to cancer drugs. Consequently, understanding the mechanisms of drug resistance is an active area of research. About 60% of melanoma tumors carry a V600E mutation in the BRAF kinase gene, rendering them sensitive

139. Patil SU, Ogunniyi AO, Calatroni A, et al. Peanut oral immunotherapy transiently expands circulating Ara h 2-specific B cells with a homologous repertoire in unrelated subjects. *J Allergy Clin Immunol.* 2015;136:125-134 e112

140. Hoh RA, Joshi SA, Liu Y, et al. Single B-cell deconvolution of peanut-specific antibody responses in allergic patients. *J Allergy Clin Immunol.* 2016;137:157-167

141. Tsioris K, Gupta NT, Ogunniyi AO, et al. Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. *Integr Biol (Camb).* 2015;7:1587-1597

142. Ishigaki K, Shoda H, Kochi Y, et al. Quantitative and qualitative characterization of expanded CD4+ T cell clones in rheumatoid arthritis patients. *Sci Rep.* 2015;5:12937

143. Tipton CM, Fucile CF, Darce J, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol.* 2015;16:755-765

144. Held K, Bhonsle-Deeng L, Siewert K, et al. alphabeta T-cell receptors from multiple sclerosis brain lesions show MAIT cell-related features. *Neurol Neuroimmunol Neuroinflamm.* 2015;2:e107

145. Eugster A, Lindner A, Catani M, et al. High diversity in the TCR repertoire of GAD65 autoantigen-specific human CD4+ T cells. *J Immunol.* 2015;194:2531-2538

146. Gaublomme JT, Yosef N, Lee Y, et al. Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. *Cell.* 2015;163:1400-1412

147. Tischner D, Grimm M, Kaur H, et al. Single-cell profiling reveals GPCR heterogeneity and functional patterning during neuroinflammation. *JCI Insight.* 2017;2:e95063

148. Burns A and Iliffe S. Alzheimer's disease. *BMJ.* 2009;338:

For Research Use Only. Not for use in diagnostic procedures.

to the BRAF inhibitor, vemurafenib; however, a subset of tumor cells will ultimately develop resistance.[149] In this study, the authors isolated single tumor cells from melanoma patient samples, which they then cultured in the presence of vemurafenib. Targeted DNA sequencing of 119 cancer-related genes in parental cells as well as vemurafenib-resistant subclones demonstrated that no new mutations developed in resistant subclones. They then performed population-based RNA-Seq of parental cells and vemurafenib-resistant cells, identifying increased expression of several known resistance marker genes in resistant cells. Moreover, drug-exposed resistant cells showed increased expression of many of the resistance marker genes. Overall, these data suggest a novel mechanism for drug resistance in single tumor cells.

Illumina Technology: HiSeq 2500 System

Nguyen A, Yoshida M, Goodarzi H and Tavazoie SF. Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. *Nat Commun.* **2016;7:11246**
Tumor cell genetic heterogeneity plays a role in drug resistance; however, in this study, the authors focused on understanding how cancer phenotypic heterogeneity might contribute to cancer progression. They isolated isogenic clonal cancer cells displaying high morphological variation. These highly variable subpopulations displayed increased survival and metastatic potential in response to numerous anticancer drugs. The authors performed scRNA-Seq of these clones, and they found that morphological variants expressed enhanced transcriptomic variability, despite their genetic stability. Finally, they correlated high variability in expression of the spliceosome-machinery genes with transcriptomic variability and promotion of metastasis. These data demonstrate a novel aspect of intraclonal tumor heterogeneity, and further suggest a mechanism for increased drug resistance, survival, and metastasis.

Illumina Technology: Nextera Extended Exome Sequencing Kit, HiSeq 2500 System

Wei W, Shin YS, Xue M, Matsutani T, Masui K, et al. Single-Cell Phosphoproteomics Resolves Adaptive Signaling Dynamics and Informs Targeted Combination Therapy in Glioblastoma. *Cancer Cell.* **2016;29:563-573**
Glioblastoma is one of the deadliest forms of cancer. Glioblastoma tumors have mutations in several druggable pathways, but current targeted therapies have proven ineffective due to rapid and universal drug resistance. Specifically, the mTOR pathway is a key driver in 90% of glioblastoma, yet tumor cells develop rapid resistance to mTOR-targeted therapies. In this study, the authors used the NextSeq 500 system to obtain single-cell genomic data, which they correlated with single-cell proteomic data in tumor cells treated with mTOR inhibitor. Their data showed that resistance to mTOR inhibitor in glioblastoma tumor cells occurred within days of drug therapy. Surprisingly, the correlation with single-cell sequencing data demonstrated that this drug resistance proceeds via nongenetic mechanisms, through upregulation of specific signaling phosphoproteins. This study suggests a novel approach for designing drug combination therapy in treating glioblastoma.

Illumina Technology: NextSeq 500 System

Gaublomme JT, Yosef N, Lee Y, Gertner RS, Yang LV, et al. Single-Cell Genomics Unveils Critical Regulators of Th17 Cell Pathogenicity. *Cell.* **2015;163:1400-1412**
In the adaptive immune system, proinflammatory IL-17–producing Th17 cells contribute to pathogen clearance, but they have also been implicated in autoimmune and inflammatory diseases. Although Th17 cells display extensive cellular heterogeneity, the genetic basis of this heterogeneity is not well characterized. In this study, the authors used the HiSeq 2000/2500 system to perform scRNA-Seq of Th17 cells isolated from patients with EAE. They compared this data set with scRNA-Seq data from normal Th17 cells differentiated *in vitro*. Their data showed genetic heterogeneity in Th17 cells from autoimmune patients. Further, they used transgenic mice to validate 4 genes (*Gpr65, Plzp, Toso,* and *Cd5l*) involved in developing EAE. These findings may lead to the discovery of novel drug targets in autoimmune disease.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2000/2500 Systems

Kim KT, Lee HW, Lee HO, et al. Single-cell mRNA sequencing identifies subclonal heterogeneity in anticancer drug responses of lung adenocarcinoma cells. *Genome Biol.* **2015;16:127**
Intratumor heterogeneity correlates with poor cancer clinical prognosis, but the mechanism for this correlation is not entirely understood. In this study, the authors isolated 34 PDX lung adenocarcinoma tumor cells and performed scRNA-Seq using the HiSeq 2000/2500 system. By clustering 69 lung adenocarcinoma–prognostic genes, including *KRAS*, they could classify the PDX cells into 4 distinct subgroups. scRNA-Seq of the PDX cells that survived anticancer drug treatment demonstrated that tumor cells with activated *KRAS* variants were targeted by anticancer drugs, even though *KRAS* itself was not the target. Their scRNA-Seq data also suggest that the tumor cells responsible for drug resistance can be masked by the genomics of the bulk tumor.

Illumina Technology: Nextera XT DNA Sample Prep Kit, HiSeq 2000/2500 Systems

149. Nazarian R, Shi H, Wang Q, et al. Melanomas acquire resistance to B-RAF(V600E) inhibition by RTK or N-RAS upregulation. *Nature.* 2010;468:973-977

Luna JM, Scheel TK, Danino T, et al. Hepatitis C virus RNA functionally sequesters miR-122. *Cell.* **2015;160:1099-1110**

Hepatitis C virus (HCV) is an RNA virus of the Flaviviridae family and is the leading cause of liver disease world-wide. HCV requires the liver-specific tumor suppressor microRNA(miR)-122 for replication, but the regulation of endogenous miRNAs by HCV infection has not been characterized. In this study, the authors generated global miRNA:target interaction maps during HCV infection and demonstrated that HCV viral RNA sequesters miR-122 to derepress normal host targets. They then validated and quantified the RNA sequestration at the single-cell level by analyzing expression data from reporters containing miR-122 sites, confirming significant derepression during HCV infection. The findings describe a model for HCV-induced miR-122 sequestration and suggest that this mechanism may enhance the long-term oncogenic potential of HCV.

Illumina Technology: MiSeq and HiSeq 2000 Systems

Suzuki A, Matsushima K, Makinoshima H, et al. Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment. *Genome Biol.* **2015;16:66**

Single-cell sequencing of individual tumor cells holds the potential of better correlating genetic heterogeneity with the mechanism of drug response and resistance. In this study, the authors characterized the heterogeneity in single-cell gene expression across 336 lung adenocarcinoma cells derived from cell lines. They also analyzed lung adenocarcinoma cells before and after treatment with the multi–tyrosine kinase inhibitor, vandetanib. They found that relative expression diversity of cellular housekeeping genes was reduced in cancer cells exposed to vandetanib. In contrast, the expression diversity of genes targeted by vandetanib (including *EGFR* and *RET*) remained constant. Their data demonstrate that patterns in gene expression divergence play important roles in tumor cells acquiring drug resistance; further, this genetic diversity is not revealed by RNA-Seq of bulk tumors.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System

Tipton CM, Fucile CF, Darce J, Chida A, Ichikawa T, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol.* **2015;16:755-765**

Acute SLE is a recurrent autoimmune disease that attacks various tissues and has no cure. Autoimmune activity is associated with surges in B cells. The only SLE therapy approved by the US FDA, belimumab, targets B-cell activating factor. In this study, the authors isolated B cells from patients experiencing SLE flares and used deep sequencing and proteomic approaches to analyze the diversity of B cells. They demonstrated that B cells from patients with SLE flares were polyclonal. By sequencing single B cells, they also identified a subpopulation of newly activated naïve B cells that provided an important source of autoantibodies during SLE flares, suggesting that SLE autoreactivities occur during polyclonal activation. These results may guide patient treatment options and facilitate the design of future SLE therapies.

Illumina Technology: MiSeq System
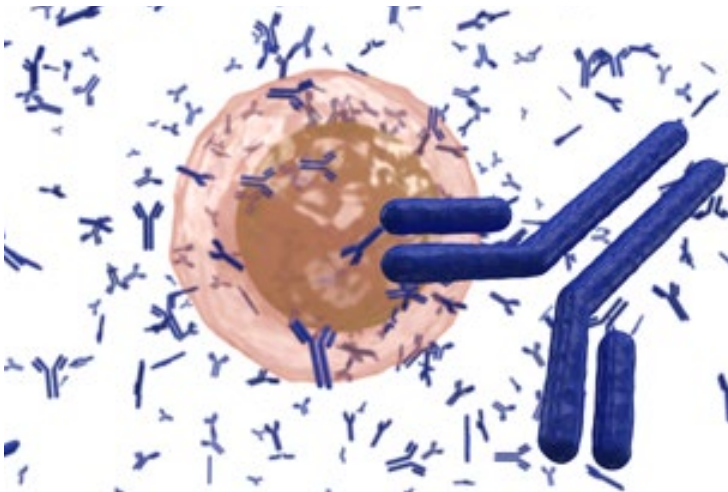
B lymphocytes (B cells) secrete antibodies.

For Research Use Only. Not for use in diagnostic procedures.

Milholland B, Dong X, Zhang L, Hao X, Suh Y and Vijg J. Differences between germline and somatic mutation rates in humans and mice. *Nat Commun.* 2017;8:15183

Tischner D, Grimm M, Kaur H, et al. Single-cell profiling reveals GPCR heterogeneity and functional patterning during neuroinflammation. *JCI Insight.* 2017;2:e95063

Spaethling JM, Sanchez-Alavez M, Lee J, et al. Single-cell transcriptomics and functional target validation of brown adipocytes show their complex roles in metabolic homeostasis. *FASEB J.* 2016;30:81-92

Avraham R, Haseley N, Brown D, Penaranda C, Jijon HB, et al. Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses. *Cell.* 2015;162:1309-1321

Combe M, Garijo R, Geller R, Cuevas JM and Sanjuan R. Single-Cell Analysis of RNA Virus Infection Identifies Multiple Genetically Diverse Viral Genomes within Single Infectious Units. *Cell Host Microbe.* 2015;18:424-432

Eirew P, Steif A, Khattra J, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature.* 2015;518:422-426

Ortmann CA, Kent DG, Nangalia J, et al. Effect of mutation order on myeloproliferative neoplasms. *N Engl J Med.* 2015;372:601-612

Paguirigan AL, Smith J, Meshinchi S, Carroll M, Maley C and Radich JP. Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. *Sci Transl Med.* 2015;7:281re282

Sekar S, McDonald J, Cuyugan L, et al. Alzheimer's disease is associated with altered expression of genes involved in immune response and mitochondrial processes in astrocytes. *Neurobiol Aging.* 2015;36:583-591

Tsioris K, Gupta NT, Ogunniyi AO, et al. Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. *Integr Biol (Camb).* 2015;7:1587-1597

For Research Use Only. Not for use in diagnostic procedures.

**38**   Single-cell Research

Reproductive Health

Prenatal diagnosis refers to testing for diseases or genetic conditions in a fetus or embryo before birth. Birth defects can be diagnosed by detecting detrimental chromosomal or DNA aberrations. Traditionally, this diagnosis has used invasive procedures, such as amniocentesis. Noninvasive prenatal testing (NIPT) offers greater confidence over traditional testing methods and only requires a blood draw.[150] NIPT through cell-free DNA (cfDNA) sequencing is used clinically and is improving prenatal diagnoses.[151]

> "Single-cell sequencing enables global sequencing of the genome, exome, transcriptome, or epigenome for a comprehensive analysis of the early embryo."
>
> - Zhu et al. 2017

Single-cell sequencing can also be useful in the realm of PGS/PGD, where embryos can be screened for chromosomal abnormalities prior to implantation.[152,153] Research shows how sequencing aids genetic testing of in vitro–generated gametes (IVG) either for research or for implantation.[154] Single-cell sequencing diagnostics can accelerate research, improve the early detection of embryo status, and assure transplantation of a healthy *in vitro* fertilized (IVF) embryo.[155,156]
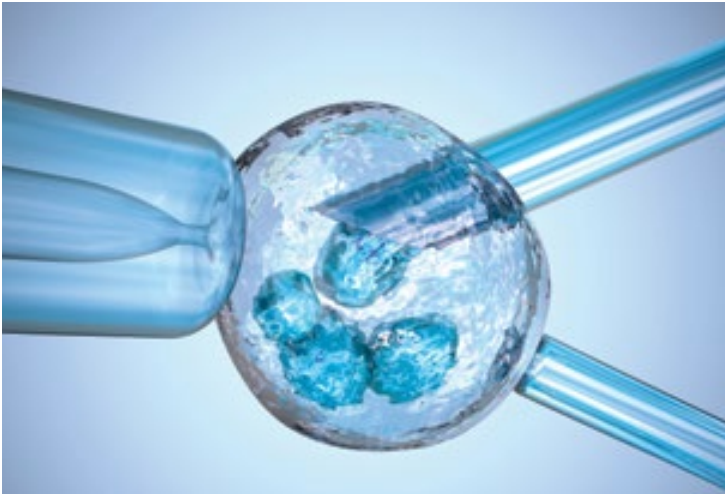
Genetic testing of an embryo can provide early insights into chromosomal abnormalities.

Reviews

Forsberg LA, Gisselsson D and Dumanski JP. Mosaicism in health and disease - clones picking up speed. *Nat Rev Genet.* 2017;18:128-142

Otte J, Wruck W and Adjaye J. New insights into human primordial germ cells and early embryonic development from single-cell analysis. *FEBS Lett.* 2017;591:2226-2240

Zhu W, Zhang XY, Marjani SL, et al. Next-generation molecular diagnosis: single-cell sequencing from bench to bedside. *Cell Mol Life Sci.* 2017;74:869-880

150. Chandrasekharan S, Minear MA, Hung A and Allyse M. Noninvasive prenatal testing goes global. *Sci Transl Med.* 2014;6:231fs215

151. Bianchi DW, Parker RL, Wentworth J, et al. DNA sequencing versus standard prenatal aneuploidy screening. N Engl J Med. 2014;370:799-808

152. Yan L, Yang M, Guo H, et al. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol.* 2013;20:1131-1139

153. Wang J, Fan HC, Behr B and Quake SR. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell.* 2012;150:402-412

154. Palacios-González C, Harris J and Testa G. Multiplex parenting: IVG and the generations to come. *Journal of Medical Ethics.* 2014;40:752-758

155. Ong FS, Lin JC, Das K, Grosu DS and Fan JB. Translational utility of next-generation sequencing. *Genomics.* 2013;102:137-139

156. Tan Y, Yin X, Zhang S, et al. Clinical outcome of preimplantation genetic diagnosis and screening using next generation sequencing. *Gigascience.* 2014;3:30

References

Krjutskov K, Katayama S, Saare M, et al. Single-cell transcriptome analysis of endometrial tissue. *Hum Reprod.* **2016;31:844-853**
Single-cell transcriptome analysis has been performed previously for a number of biopsy tissues, but those for endometrium have not been published. In this study, the authors present a complete pipeline for endometrial single-cell gene expression profiling, from sampling to data analysis. They dissociated individual stromal and epithelial cells from endometrium and isolated the relevant cell types using FACS followed by scRNA-Seq. In parallel, they cultured endometrial stromal cells and compared global expression profiles with uncultured cells. The authors found that 2661 genes (out of 8622 detected genes) were more active in cultured stromal cells than in biopsy cells. Further, 241 genes were differentially expressed in biopsy vs cultured individual cells, at significant levels.

Illumina Technology: HiSeq System

Petropoulos S, Edsgard D, Reinius B, et al. Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell.* **2016;165:1012-1026**
During human development, the zygote undergoes cellular division to establish 3 cell types of the mature blastocyst: trophectoderm (TE), primitive endoderm (PE), and epiblast (EPI). In this study, the authors used the HiSeq 2000 system to sequence the transcriptomes of 1529 individual cells from 88 human preimplantation embryos. They developed a comprehensive transcriptional map of human preimplantation development, demonstrating that lineages of TE, PE, and EPI all occur simultaneously and coincide with blastocyst formation. Their data also identified novel genes, such as ARGFX, FOXA2, and LINC00261, that may be important for preimplantation development. Finally, their transcriptional map identified unique features of X-chromosome dosage compensation in humans.

Illumina Technology: HiSeq 2000 System

Kumar A, Ryan A, Kitzman JO, et al. Whole genome prediction for preimplantation genetic diagnosis. *Genome Med.* **2015;7:35**
PGD allows for genetic profiling of human embryos to screen for genetic disorders prior to implantation. Although recent technical advances in single-cell sequencing highlight the promise of this approach in PGD, DNA amplification bias and relatively long turnaround time limit this approach. In this study, the authors combined haplotype-resolved parental genome sequencing with rapid embryo genotyping to predict the whole-genome sequence of a 5-day human embryo. The authors predicted inheritance at 3 million paternal/maternal heterozygous sites with > 99% accuracy, as well as the transmission of a parental HBA1/HBA2 deletion. The data suggest that PGD may facilitate the diagnosis of genetic diseases in embryos.

Illumina Technology: HiSeq 2000 System, CytoSNP-12 Chip

Yan L, Huang L, Xu L, et al. Live births after simultaneous avoidance of monogenic diseases and chromosome abnormality by next-generation sequencing with linkage analyses. *Proc Natl Acad Sci U S A.* **2015;112:15964-15969**
NGS methods have improved the precision of PGS/PGD. Although the precision has been limited by false-positive and false-negative SNVs, linkage analysis can overcome this challenge. In this study, the authors developed MARSALA, a method that combines NGS using the HiSeq platform with single-cell WGA. The method allows for embryo diagnosis with a single-molecule precision and significantly reduces false-positive and false-negative errors. This is the first integrated NGS-based PGD procedure that simultaneously detects disease-causing mutations and chromosome abnormalities, and performs linkage analyses.

Illumina Technology: HiSeq 2500 System

Zamani Esteki M, Dimitriadou E, Mateiu L, et al. Concurrent whole-genome haplotyping and copy-number profiling of single cells. *Am J Hum Genet.* **2015;96:894-912**
Before analyzing single-cell DNA-Seq data, DNA copy-number aberrations must be differentiated from WGA artifacts. This requirement makes DNA copy-number profiling and haplotyping of single-cell sequencing data challenging. In this study, the authors developed a single-cell genome analysis method that determined haplotypes and copy number across the genome of a single cell—a process called, haplarithmisis. The method deciphered SNP allele fractions of single cells and integrated these data into a computational workflow for imputation of linked disease variants (siCHILD). The authors validated the method by determination of haplotypes carrying disease alleles in single-cell genomes from individual lymphocytes and human blastomeres derived from human IVF embryos.

Illumina Technology: TruSeq DNA LT Sample Preparation Kit, HumanCytoSNP-12v2.1 BeadChips, HiSeq 2000/2500 Systems

Moreira de Mello JC, Fernandes GR, Vibranovski MD and Pereira LV. Early X chromosome inactivation during human preimplantation development revealed by single-cell RNA-sequencing. *Sci Rep.* 2017;7:10794

Pavlicev M, Wagner GP, Chavan AR, et al. Single-cell transcriptomics of the human placenta: inferring the cell communication network of the maternal-fetal interface. *Genome Res.* 2017;27:349-361

Tukiainen T, Villani AC, Yen A, et al. Landscape of X chromosome inactivation across human tissues. *Nature.* 2017;550:244-248

Gerovska D and Arauzo-Bravo MJ. Does mouse embryo primordial germ cell activation start before implantation as suggested by single-cell transcriptomics dynamics? *Mol Hum Reprod.* 2016;22:208-225

Li N, Wang L, Wang H, et al. The Performance of Whole Genome Amplification Methods and Next-Generation Sequencing for Pre-Implantation Genetic Diagnosis of Chromosomal Abnormalities. *J Genet Genomics.* 2015;42:151-159

Microbial Ecology and Evolution

Not only do we discover new species and microorganisms continually,[158,159] current NGS technology can help us to understand the dynamics of microbial ecology and evolution. These discoveries include host-species interactions that generate selection pressures,[160] which can lead to the evolution of a species.[158,160] Such insight is critical to understanding complex ecosystems and the many unique microbes that comprise them.[161] Sequencing can also facilitate the detection of these new species and their evolved roles in their respective environments.[162,163,164,165]

"New genomic information from metagenomics and single cell genomics has provided insights into microbial metabolic cooperation and dependence, generating new avenues for cultivation efforts."

- Solden et al. 2016

Single-cell analysis provides a better assessment of how different organisms pressure selection and the evolution of cohabitants, as well as host-pathogen interactions.

158. Kashtan N, Roggensack SE, Rodrigue S, et al. Single-cell genomics reveals hundreds of coexisting subpopulations in wild Prochlorococcus. *Science.* 2014;344:416-420

159. Hug LA, Baker BJ, Anantharaman K, et al. A new view of the tree of life. *Nature Microbiology.* 2016;1:16048

160. Martiny JB, Riemann L, Marston MF and Middelboe M. Antagonistic coevolution of marine planktonic viruses and their hosts. *Ann Rev Mar Sci.* 2014;6:393-414

161. Stepanauskas R. Single cell genomics: an individual look at microbes. *Curr Opin Microbiol.* 2012;15:613-620

162. Kaster A-K, Mayer-Blackwell K, Pasarelli B and Spormann AM. Single cell genomic study of Dehalococcoidetes species from deep-sea sediments of the Peruvian Margin. *The Isme Journal.* 2014;8:1831

163. Wasmund K, Schreiber L, Lloyd KG, et al. Genome sequencing of a single cell of the widely distributed marine subsurface Dehalococcoidia, phylum Chloroflexi. *ISME J.* 2014;8:383-397

164. Wang FP, Zhang Y, Chen Y, et al. Methanotrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. *ISME J.* 2014;8:1069-1078

165. Mangot JF, Logares R, Sanchez P, et al. Accessing the genomic information of unculturable oceanic picoeukaryotes by combining multiple single cells. *Sci Rep.* 2017;7:41498

Reviews

Howard-Varona C, Hargreaves KR, Abedon ST and Sullivan MB. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* 2017;11:1511-1520

Hug LA, Baker BJ, Anantharaman K, et al. A new view of the tree of life. *Nature Microbiology.* 2016;1:16048

Kodzius R and Gojobori T. Single-cell technologies in environmental omics. *Gene.* 2016;576:701-707

Solden L, Lloyd K and Wrighton K. The bright side of microbial dark matter: lessons learned from the uncultivated majority. *Curr Opin Microbiol.* 2016;31:217-226

References

Dyksma S, Bischof K, Fuchs BM, Hoffmann K, Meier D, et al. Ubiquitous Gammaproteobacteria dominate dark carbon fixation in coastal sediments. *ISME J.* 2016;8:1939-1953
Marine sediments are the largest carbon sink on the planet, with half of chemosynthetic oceanic carbon fixation occurring in coastal sediments. However, the microbes responsible for this activity are unknown. By surveying bacterial 16S rRNA gene diversity from 13 coastal sediments across Europe and Australia, the authors identified groups of *Gammaproteobacteria* that were affiliated with sulfur-oxidizing bacteria. ^{14}C-carbon assimilation studies showed that these uncultured *Gammaproteobacteria* accounted for 80% of carbon fixation in coastal sediments. Finally, the authors isolated individual cells from the environmental sample and performed single-cell WGS to identify genes that linked hydrogen-oxidizing activity with sulfur-oxidizing *Gammaproteobacteria*.

Illumina Technology: MiSeq and HiSeq 2000 Systems

Spencer SJ, Tamminen MV, Preheim SP, Guo MT, Briggs AW, et al. Massively parallel sequencing of single cells by epicPCR links functional genes with phylogenetic markers. *ISME J.* 2016;10:427-436
In microbial ecology studies, 16S rRNA sequencing can identify microbial community members, whereas shotgun metagenomics can determine the functional diversity of the community. However, combining the 2 approaches is technically challenging. In this study, the authors developed emulsion, paired isolation, and concatenation PCR (epicPCR), a technique that links functional genes and phylogenetic markers. They applied the technique to millions of uncultured individual cells from the freshwater Upper Mystic Lake in Massachusetts. Specifically, they profiled the sulfate-reducing community within the freshwater lake community and were able to identify new putative sulfate reducers. The method is suitable for identifying functional community members, tracing gene transfer, and mapping ecological interactions in microbial cells.

Illumina Technology: MiSeq System

Tsementzi D, Wu J, Deutsch S, et al. SAR11 bacteria linked to ocean anoxia and nitrogen loss. *Nature.* 2016;536:179-183
SAR11 bacteria are the most abundant microbes in the earth's oceans, constituting half of all microbial cells in the oxygen-rich surface ocean. Although considered aerobic, SAR11 are also abundant in marine environments where oxygen levels are low. In this study, the authors used the MiSeq system to sequence 19 single-cell amplified genomes from a subpopulation of SAR11 bacteria isolated from ocean oxygen-minimum zones. They found that the SAR11 bacteria that had adapted to their low-oxygen environment encoded abundant respiratory nitrate reductases. These enzymes perform the first step in denitrification, a microbially facilitated process of nitrate reduction that may ultimately produce molecular nitrogen (N_2). These data redefine the ecological niche of earth's most abundant organismal group and suggest that SAR11 bacteria contribute to nitrogen loss in oxygen-minimum zones.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq and HiSeq Systems

Combe M, Garijo R, Geller R, Cuevas JM and Sanjuan R. Single-Cell Analysis of RNA Virus Infection Identifies Multiple Genetically Diverse Viral Genomes within Single Infectious Units. *Cell Host Microbe.* 2015;18:424-432
Genetic diversity is a key determinant in the ability of viruses to escape immunity and vaccination, develop drug resistance, and cause disease. It is assumed that single virions constitute viral infectious units. However, the authors performed single-cell sequencing of 881 VSV plaques derived from 90 individual infected cells and showed that individual virus infectious units were comprised of multiple genetically diverse viral genomes. They also found that several genome viral variants could be delivered simultaneously to the same individual cells, and the rate of spontaneous virus mutation varied across individual cells, with implications for viral yield. This study at the single-cell level has implications for our understanding of viral diversity and evolution.

Illumina Technology: MiSeq System

For Research Use Only. Not for use in diagnostic procedures.

42 Single-cell Research

Labonte JM, Swan BK, Poulos B, et al. Single-cell genomics-based analysis of virus-host interactions in marine surface bacterioplankton. *ISME J.* **2015;9:2386-2399**
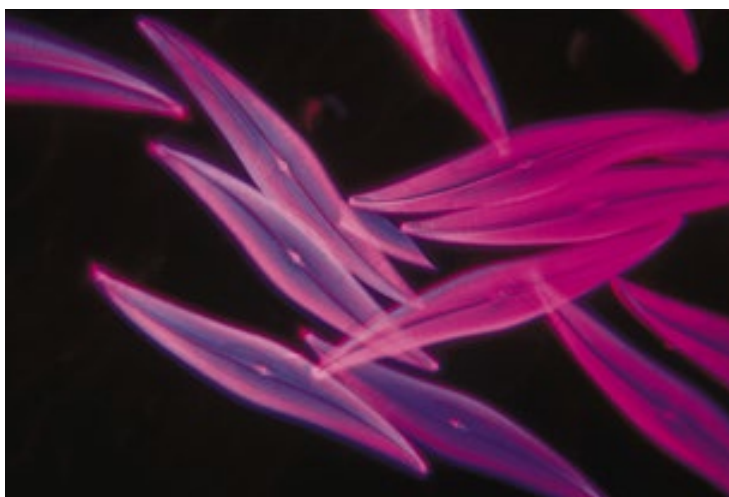
Viral infections can alter the composition and metabolic potential of marine communities, as well as the evolution of host populations. All oceanic microbes are potentially impacted by viral infections; however, our understanding of host-virus interactions is limited. In this study, the authors used single-cell WGS of 58 isolated oceanic microbes to identify genomic blueprints of viruses inside or attached to individual bacterial and archaeal cells. The data include the first known viruses of Thaumarchaeota, Marinimicrobia, Verrucomicrobia, and Gammaproteobacteria. They demonstrate that single-cell genomics approaches can provide insight into host-virus interactions in complex environments.

Illumina Technology: NextSeq 500 System

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, et al. Ocean plankton. Determinants of community structure in the global plankton interactome. *Science.* **2015;348:1262073**

Oceanic plankton is the world's largest ecosystem and is composed of viruses, prokaryotes, microbial eukaryotes, phytoplankton, and zooplankton. This ecosystem structure and composition are influenced by environmental conditions and nutrient availability. In this study, the authors analyzed 313 plankton samples from the *Tara* Oceans expedition and obtained viral, eukaryotic, and prokaryotic abundance profiles from Illumina-sequenced metagenomes and 18S rDNA V9 sequences. They used network inference and machine-learning methods to construct an interactome among plankton groups. In particular, the authors confirmed predicted virus-host interactions by comparing putative host contigs with viral data from single-cell genomes.

Illumina Technology: Illumina-sequenced metagenomes (mitags) and 18S rRNA V9 sequences

Phytoplankton accounts for half of all photosynthetic activity on Earth.

Martijn J, Schulz F, Zaremba-Niedzwiedzka K, et al. Single-cell genomics of a rare environmental alphaproteobacterium provides unique insights into Rickettsiaceae evolution. *ISME J.* **2015;9:2373-2385**

The bacterial family Rickettsiaceae includes the epidemic typhus-causing pathogen *Rickettsia prowazekii*, and thus Rickettsiaceae host-pathogen interactions are of great interest. In this study, the authors discovered *Candidatus* Arcanobacter lacustris, a Rickettsiaceae sister lineage alphaproteobacterium isolated from Damariscotta Lake. They used the HiSeq 2000 system to perform single-cell WGS of *Candidatus* Arcanobacter lacustris. Phylogenetic and comparative analysis of its genome revealed the presence of chemotaxis and flagellar genes. These genes are unique in the Rickettsiaceae family and suggest that the ancestor of Rickettsiaceae may have had a facultative lifestyle.

Illumina Technology: HiSeq 2000 System

Haliburton JR, Shao W, Deutschbauer A, Arkin A and Abate AR. Genetic interaction mapping with microfluidic-based single cell sequencing. *PLoS One.* 2017;12:e0171302

Ionescu D, Bizic-Ionescu M, De Maio N, Cypionka H and Grossart HP. Community-like genome in single cells of the sulfur bacterium Achromatium oxaliferum. *Nat Commun.* 2017;8:455

Mangot JF, Logares R, Sanchez P, et al. Accessing the genomic information of unculturable oceanic picoeukaryotes by combining multiple single cells. *Sci Rep.* 2017;7:41498

Beam JP, Jay ZJ, Schmid MC, et al. Ecophysiology of an uncultivated lineage of Aigarchaeota from an oxic, hot spring filamentous 'streamer' community. *ISME J.* 2016;10:210-224

Cottinet D, Condamine F, Bremond N, et al. Lineage Tracking for Probing Heritable Phenotypes at Single-Cell Resolution. *PLoS One.* 2016;11:e0152395

Mwirichia R, Alam I, Rashid M, et al. Metabolic traits of an uncultured archaeal lineage--MSBL1--from brine pools of the Red Sea. *Sci Rep.* 2016;6:19181

Ngugi DK, Blom J, Stepanauskas R and Stingl U. Diversification and niche adaptations of Nitrospina-like bacteria in the polyextreme interfaces of Red Sea brines. *ISME J.* 2016;10:1383-1399

Gavelis GS, White RA, Suttle CA, Keeling PJ and Leander BS. Single-cell transcriptomics using spliced leader PCR: Evidence for multiple losses of photosynthesis in polykrikoid dinoflagellates. *BMC Genomics.* 2015;16:528

Nobu MK, Narihiro T, Rinke C, et al. Microbial dark matter ecogenomics reveals complex synergistic networks in a methanogenic bioreactor. *ISME J.* 2015;9:1710-1722

Youssef NH, Rinke C, Stepanauskas R, Farag I, Woyke T and Elshahed MS. Insights into the metabolism, lifestyle and putative evolutionary history of the novel archaeal phylum 'Diapherotrites'. *ISME J.* 2015;9:447-460

Plant Biology

Although transcriptomics has advanced our understanding of plant development, single-cell transcriptomics has not yet been employed as widely in plants.[166,167] Single-cell sequencing approaches have great potential to further our understanding of plant biology.[168,169] However, plant cells are enclosed in a rigid cell-wall matrix, and isolating individual plant cells remains challenging technically. In *Arabidopsis*, a number of techniques have been used to isolate single cells, including protoplasting to remove plant cell walls followed by FACS,[170,171] and cell-wall digestion followed by glass micropipetting.[172] In maize kernels, RNA-Seq of LCM compartments has

166. Brady SM, Orlando DA, Lee JY, et al. A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science.* 2007;318:801-806

167. Birnbaum K, Shasha DE, Wang JY, et al. A gene expression map of the Arabidopsis root. *Science.* 2003;302:1956-1960

167. Efroni I and Birnbaum KD. The potential of single-cell profiling in plants. *Genome Biol.* 2016;17:65

169. Giacomello S, Salmen F, Terebieniec BK, et al. Spatially resolved transcriptome profiling in model plant species. *Nat Plants.* 2017;3:17061

170. Birnbaum K, Jung JW, Wang JY, et al. Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. *Nat Methods.* 2005;2:615-619

171. Adrian J, Chang J, Ballenger CE, et al. Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Dev Cell.* 2015;33:107-118

172. Efroni I, Ip PL, Nawy T, Mello A and Birnbaum KD. Quantification of cell identity from single-cell gene expression profiles. *Genome Biol.* 2015;16:9

173. Zhan J, Thakare D, Ma C, et al. RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *Plant Cell.* 2015;27:513-531

174. Li X, Li L and Yan J. Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nat Commun.* 2015;6:6648

175. Efroni I, Ip PL, Nawy T, Mello A and Birnbaum KD. Quantification of cell identity from single-cell gene expression profiles. *Genome Biol.* 2015;16:9

"The continued application of single-cell plant system biology is likely to transform our view of the relationships, development, and evolution of distinct cell-types in plants." - Libault et al. 2017

Single-cell analysis will allow a better assessment of the nature of plant stem cells, plant cell plasticity, and local cellular response to environmental changes.

revealed details of plant endosperm cell differentiation,[173] and WGS of individual maize microspores has advanced our understanding of plant meiotic recombination.[174] While technical challenges remain, scRNA-Seq methods are poised to revolutionize our knowledge of plant biology.[175]

Reviews

Libault M, Pingault L, Zogli P and Schiefelbein J. Plant Systems Biology at the Single-Cell Level. *Trends Plant Sci.* 2017;22:949-960

Efroni I and Birnbaum KD. The potential of single-cell profiling in plants. *Genome Biol.* 2016;17:65

References

Giacomello S, Salmen F, Terebieniec BK, et al. Spatially resolved transcriptome profiling in model plant species. *Nat Plants.* 2017;3:17061
Like mammalian systems, functional biology in plants is driven by spatially oriented gene expression programs that are differentially regulated in different tissues. Resolving this spatial organization in gene expression using single-cell RNA-Seq of individual plant cells is feasible, and it has been achieved by using laser-capture micro-dissection (LCM, see below) and FACS of plant tissues. In this study, the authors present a novel method for generating spatial transcriptome data in plants. They imaged *Arabidopsis thaliana* tissue cryosections on bar-coded arrays containing spot-specific barcodes coupled to oligo(dT) oligonucleotides. After cDNA synthesis, they used a NextSeq System to characterize gene expression within each array spot. By integrating the spatial information with gene expression data, the authors identified gene expression differences in 8 tissue domains of *Arabidopsis thaliana*.

Illumina Technology: NextSeq System

Adrian J, Chang J, Ballenger CE, Bargmann BO, Alassimone J, et al. Transcriptome dynamics of the stomatal lineage: birth, amplification, and termination of a self-renewing population. *Dev Cell.* 2015;33:107-118
Plant stomata facilitate plant gas exchange with the atmosphere. In *Arabidopsis*, the production and pattern of stomata proceeds from a discrete lineage that can be parsed into intermediate steps. Despite the biological significance of RNase L, the RNAs cleaved by this enzyme are poorly defined. In this study, the authors used Illumina sequencing to reveal the frequency and location of RNase L cleavage sites within host and viral RNAs. The method was optimized and validated using viral RNAs cleaved with RNase L and RNase A, and RNA from infected and noninfected HeLa cells. The authors identified discrete genomic regions susceptible to RNase L and other single-strand–specific endoribonucleases. Monitoring the frequency and location of these cleavage sites within host and viral RNAs may reveal how these enzymes contribute to health and disease.

Illumina Technology: TruSeq SBS Kit v3–HS, HiSeq 2000 System

Maize kernels.

Li X, Li L and Yan J. Dissecting meiotic recombination based on tetrad analysis by single-microspore sequencing in maize. *Nat Commun.* **2015;6:6648**

Meiotic recombination plays an important role in genetic diversity by contributing to allele assortment, creating a substrate for natural selection, and evolving eukaryotic genomes. Maize has been used successfully as a genetic model for the dissection of recombination variation, but understanding single-meiotic events at nucleotide-level resolution has been impossible previously, due to the difficulty in sequencing single plant cells and gametes. In this study, the authors developed a method for isolation and WGS of the 4 microspores of a single maize tetrad. They used the HiSeq 2000 system to sequence 96 individual microspores, identifying 600,000 high-quality SNPs, which allowed them to characterize recombination patterns at very high resolution. Their high-resolution recombination map revealed that crossovers were more likely to occur in genic rather than intergenic regions; further, they were especially common in the 5'- and 3'-end regions of annotated genes.

Illumina Technology: TruSeq DNA Sample Preparation Kit, HiSeq 2000 System

Zhan J, Thakare D, Ma C, Lloyd A, Nixon NM, et al. RNA sequencing of laser-capture microdissected compartments of the maize kernel identifies regulatory modules associated with endosperm cell differentiation. *Plant Cell.* **2015;27:513-531**

Cereal endosperm is a main source of food, feed, and raw material worldwide, yet genetic control of endosperm cell differentiation is not well defined. In this study, the authors coupled LCM and Illumina sequencing to profile mRNAs in 5 major cell types of differentiating endosperms and 4 compartments of maize (*Zea mays*) kernels. They identified mRNAs that specifically accumulate in each compartment, as well as genes predominantly expressed in 1 or multiple compartments. Their results demonstrate that the MRP-1 transcription factor can activate gene regulatory networks within the basal endosperm transfer layer. These data provide a high-resolution gene activity atlas of the compartments of the maize kernel. The study also uncovers the regulatory modules associated with differentiation of the major endosperm cell types.

Illumina Technology: TruSeq DNA Sample Preparation Kit, HiSeq 2000 System

Efroni I, Ip PL, Nawy T, Mello A and Birnbaum KD. Quantification of cell identity from single-cell gene expression profiles. *Genome Biol.* 2015;16:9

Ranjan A, Townsley BT, Ichihashi Y, Sinha NR and Chitwood DH. An intracellular transcriptomic atlas of the giant coenocyte Caulerpa taxifolia. *PLoS Genet.* 2015;11:e1004900

For Research Use Only. Not for use in diagnostic procedures.

46 Single-cell Research

Forensics

NGS has revolutionized many aspects of modern forensics, including short tandem repeat (STR) analysis, monozygotic twin discrimination, Y chromosome analysis, mitochondrial whole-genome studies, age estimation, cause-of-death determination, bodily fluid identification, forensic microbiological analysis, species identification, and ancestry inference.[176,177] Single-cell forensic analysis was first reported in 1997, when van Oorschot et al. performed STR analysis for 226 individual buccal cells isolated by micromanipulation.[178] Single-cell sequencing methods are suited to criminal investigations that are hampered by very small amounts of biological material. In sexual assault crimes, sperm cells can degrade or can be contaminated by the victim's epithelial cells, but mtDNA typing of individual sperm cells can resolve this issue.[179,180] Single-cell mtDNA analysis has also been applied to individual cells of other human tissues[181] and blood.[182,183]

Single-cell sequencing techniques can assist in criminal investigations where forensic analysis of evidence is hampered by extremely low amounts of sample.

Reviews

Naue J, Hoefsloot HCJ, Kloosterman AD and Verschure PJ. Forensic DNA methylation profiling from minimal traces: How low can we go? *Forensic Sci Int Genet.* 2017;33:17-23

Yao YG, Kajigaya S and Young NS. Mitochondrial DNA mutations in single human blood cells. *Mutat Res.* 2015;779:68-77

Yang Y, Xie B and Yan J. Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics.* 2014;12:190-197

References

Jayaprakash AD, Benson EK, Gone S, et al. Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. *Nucleic Acids Res.* 2015;43:2177-2187
In addition to the nuclear genome, eukaryotic cells also carry a mitochondrial genome, and mtDNA profiling is a useful tool in forensic analysis.[184] Heteroplasmy, the occurrence of multiple mtDNA haplotypes in a cell, can increase the strength of DNA evidence in cases of historical significance. However, since mtDNA makes up less than 1% of total cell DNA, characterizing mtDNA diversity has proved challenging. In this study, the authors developed Mseek, a method for purifying and sequencing mtDNA. They used the MiSeq system to sequence mtDNA from human PBMCs to show that heteroplasmy is maintained stably in individual daughter cells over multiple cell divisions.

Illumina Technology: MiSeq System

176. Yang Y, Xie B and Yan J. Application of next-generation sequencing technology in forensic science. *Genomics Proteomics Bioinformatics.* 2014;12:190-197

177. Vidaki A, Daniel B and Court DS. Forensic DNA methylation profiling--potential opportunities and challenges. *Forensic Sci Int Genet.* 2013;7:499-507

178. van Oorschot RA and Jones MK. DNA fingerprints from fingerprints. Nature. 1997;387:767

179. Pereira J, Neves R, Forat S, Huckenbeck W and Olek K. MtDNA typing of single-sperm cells isolated by micromanipulation. *Forensic Sci Int Genet.* 2012;6:228-235

180. Lu S, Zong C, Fan W, et al. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. *Science.* 2012;338:1627-1630

181. Nekhaeva E, Bodyak ND, Kraytsberg Y, et al. Clonally expanded mtDNA point mutations are abundant in individual cells of human tissues. *Proc Natl Acad Sci U S A.* 2002;99:5521-5526

182. Yao YG, Kajigaya S, Samsel L, McCoy JP, Jr., Torelli G and Young NS. Apparent mtDNA sequence heterogeneity in single human blood CD34+ cells is markedly affected by storage and transport. *Mutat Res.* 2013;751-752:36-41

183. Yao YG, Kajigaya S and Young NS. Mitochondrial DNA mutations in single human blood cells. *Mutat Res.* 2015;779:68-77

184. Just RS, Irwin JA and Parson W. Mitochondrial DNA heteroplasmy in the emerging field of massively parallel sequencing. *Forensic Sci Int Genet.* 2015;18:131-139

For Research Use Only. Not for use in diagnostic procedures.

Hanson E, Haas C, Jucker R and Ballantyne J. Specific and sensitive mRNA biomarkers for the identification of skin in 'touch DNA' evidence. *Forensic Sci Int Genet.* 2012;6:548-558

Forensic casework often focuses on microscopic or trace amounts of biological material left behind at a crime scene. Forensic profiles from these samples are demonstrated with "touch-DNA" evidence, which is understood to be DNA from skin cells transferred to an object through physical contact. In this study, the authors used the Genome Analyzer$_{IIx}$ to obtain transcriptome data from bulk human tissues, as well as from very low amounts of RNA (5-25 pg) from a few cells. By comparing the expression data across samples, they identified 5 mRNA markers highly specific to human skin that could be detected in almost all touch-DNA samples.

Illumina Technology: GAII$_x$ System

Geng T and Mathies RA. Minimizing inhibition of PCR-STR typing using digital agarose droplet microfluidics. *Forensic Sci Int Genet.* 2015;14:203-209

Pereira J, Neves R, Forat S, Huckenbeck W and Olek K. MtDNA typing of single-sperm cells isolated by micromanipulation. *Forensic Sci Int Genet.* 2012;6:228-235

185. Eckersley-Maslin MA and Spector DL. Random monoallelic expression: regulating gene expression one allele at a time. *Trends Genet.* 2014;30:237-244

186. Gendrel AV, Attia M, Chen CJ, et al. Developmental dynamics and disease potential of random monoallelic gene expression. *Dev Cell.* 2014;28:366-380

187. Saliba AE, Westermann AJ, Gorski SA and Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 2014;42:8845-8860

188. Eckersley-Maslin M. A., Thybert D., Bergmann J. H., Marioni J. C., Flicek P., et al. (2014) Random monoallelic gene expression increases upon embryonic stem cell differentiation. *Dev Cell.* 28: 351-365

Allele-Specific Gene Expression

Diploid organisms have 2 sets of chromosomes, 1 from each parent. Genes can be transcribed from 1 allele (monoallelic expression) or from 2 alleles (biallelic expression). Population sequencing provides a global representation of gene expression, but the expression levels of rare isoforms may be lost. Single-cell sequencing approaches can detect these rare isoforms, as well as changes between monoallelic and biallelic expression.

Compared to established methods such as RNA fluorescence *in situ* hybridization (FISH), RNA sequencing, PCR, and live-cell imaging, single-cell RNA sequencing provides the most accurate representation of monoallelic or biallelic expression in individual cells within a population.[185] Combining RNA expression data with SNP data can define specific SNPs that lead to preferential allele expression or silencing. Additionally, it can elucidate their subsequent role in cancer or tissue-specific differentiation.[186,187]
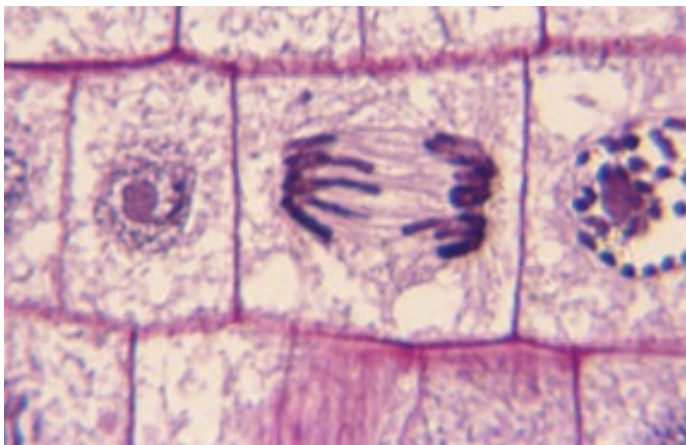
Mitotic cells accumulate SNPs that can play a role in determining random monoallelic gene expression. [188]

Reviews

Benitez JA, Cheng S and Deng Q. Revealing allele-specific gene expression by single-cell transcriptomics. Int J *Biochem Cell Biol.* 2017;90:155-160

Gawad C, Koh W and Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet.* 2016;17:175-188

Grun D and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell.* 2015;163:799-810

Huang L, Ma F, Chapman A, Lu S and Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genomics Hum Genet.* 2015;16:79-102

Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC and Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell.* 2015;58:610-620

References

Borel C, Ferreira PG, Santoni F, et al. Biased allelic expression in human primary fibroblast single cells. *Am J Hum Genet.* 2015;96:70-80
Mammalian cells have 2 alleles from which gene transcription can occur, but whether mRNAs are actively transcribed from 1 or both alleles is a subject of intense research. In this study, the authors used the HiSeq 2000 system to perform RNA-Seq of 203 single human primary fibroblasts, to determine allele-specific expression levels. Their data showed that, for the majority of genes in a cell, transcripts were derived from only 1 of 2 alleles. Moreover, genes expressing both alleles in a given cell were rare, and allele-specific expression correlated with cellular transcript levels.

Illumina Technology: Nextera XT DNA Kit, TruSeq RNA Kit, TruSeq DNA Kit, HiSeq 2000 System

Zhang CZ, Adalsteinsson VA, Francis J, et al. Calibrating genomic and allelic coverage bias in single-cell sequencing. *Nat Commun.* 2015;6:6822
In single-cell DNA-Seq, sequence artifacts are introduced by requisite DNA amplification methods, such as MDA and multiple annealing and looping–based amplification cycles (MALBAC). In this study, the authors developed a new statistical method for quantitative assessment of single-cell DNA amplification bias due to WGA. By comparing MDA and MALBAC DNA libraries, they provided a benchmark comparison of single-cell libraries generated by MDA and MALBAC and also identified universal features of genomic coverage bias at the amplicon level. Their statistical models allowed for calibration of allelic bias in single-cell WGA data.

Illumina Technology: MiSeq and HiSeq 2500 Systems

Soto M, Raaijmakers JA, Bakker B, et al. p53 Prohibits Propagation of Chromosome Segregation Errors that Produce Structural Aneuploidies. *Cell Rep.* 2017;19:2423-2431

Kim JK, Kolodziejczyk AA, Illicic T, Teichmann SA and Marioni JC. Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun.* 2015;6:8687

SAMPLE PREPARATION

Isolating individual cells is the first step in single-cell sequencing workflows, and many techniques are available.[189,190] In addition to well-established methods (including FACS, serial dilution, micropipetting, and LCM), microfluidics and drop-based techniques have increased the throughput of single-cell sequencing workflows, enabling greater accuracy and specificity in single-cell data analysis.[191,192,193] This section highlights some techniques used for isolating single cells from suspensions or tissues (Table 1).

Single cells are isolated from dissociated tissues and directly sorted into 96-well plates for analysis.

189. Saliba AE, Westermann AJ, Gorski SA and Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 2014;42:8845-8860

190. Wang Y and Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* 2015;58:598-609

191. Yu P and Lin W. Single-cell Transcriptome Study as Big Data. *Genomics Proteomics Bioinformatics.* 2016;14:21-30

192. Gawad C, Koh W and Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet.* 2016;17:175-188

193. Huang L, Ma F, Chapman A, Lu S and Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genomics Hum Genet.* 2015;16:79-102

Table 1. Methods for Single-Cell Isolation

Method	Description	Advantages[169, 170]	Disadvantages[169, 170]	Cost[170]
FACS[194]	Microdroplets with single cells isolated using electric charge	• Specific immuno-tagging of cell-surface markers improves accuracy • High throughput	• Requires specific antibodies/markers • Expensive equipment	$$
Serial dilution	Serial dilution to 1 cell per well melanogaster	• Simple approach	• Probability of isolating multiple cells	$
Mouth pipetting	Isolate single cells with glass pipettes	• Simple approach	• Technically difficult	$
Robotic micro-manipulation	Robotic micropipettes isolate single cells	• High accuracy	• Low throughput	$$$
Microfluidics platforms[195]	Microfluidic chips isolate cells in flow channels	• Isolate cells from small volumes • High throughput	• Requires uniform cell size • Expensive consumables	$$$
Optical tweezers[196]	Dissociated cell suspension	• Focused and controlled cell isolation • Fluorescence tagging of cells	• Technically challenging • Prolonged laser exposure can damage cells	$$$
Single nuclei[197]	Isolate nuclei from tissue homogenates and sort by FACS	• Gentle treatment avoids gene expression artifacts • High throughput	• Cytoplasmic transcripts and small RNAs are not detectable	$$
Nanofilters[198]	Size-exclusion filtration on filters	• Cells selected by size	• Cells can adhere to filters	$
Mag Sweeper[199]	Rotating magnet with EpCAM antibodies	• Enrichment of rare cells	• Requires markers for isolation	$$
Micro-manipulation[200]	Dissociated cell suspension	• Can isolate diverse cell types from mixed population	• Low throughput • Large starting volume needed	$
TIVA[201]	Photo-activatable mRNA capture molecule from live single cells	• Compatible with live tissues, retaining single-cell micro-environment • Noninvasive approach	• Low throughput	$$$
CellSearch[202]	Magnets with antibody-conjugated nanoparticles	• High throughput	• Bias toward isolation markers	$$$
CellCelector[203]	Robotic capillary micro-manipulator	• High throughput	• Expensive	$$$
DEP-Array[204]	Microchip with dielectric cages	• High sensitivity enables isolation of rare cells	• Low throughput • Time-consuming	$$$$
LCM[205]	Cells are cut from tissue section with laser under microscope	• Preserves spatial context	• Technically challenging • Potential UV damage to RNA/DNA	$$$

194. Vermeulen L, Todaro M, de Sousa Mello F, et al. Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proc Natl Acad Sci U S A.* 2008;105:13427-13432

195. Hsiao AP, Barbee KD and Huang X. Microfluidic Device for Capture and Isolation of Single Cells. *Proceedings of SPIE--the International Society for Optical Engineering.* 2010;7759:77590W_77591

196. Yoshimoto N, Kida A, Jie X, et al. An automated system for high-throughput single cell-based breeding. *Sci Rep.* 2013;3:1191

197. Krishnaswami SR, Grindberg RV, Novotny M, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524

198. Adams DL, Martin SS, Alpaugh RK, et al. Circulating giant macrophages as a potential biomarker of solid tumors. *Proc Natl Acad Sci U S A.* 2014;111:3514-3519

199. Powell AA, Talasaz AH, Zhang H, et al. Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. *PLoS One.* 2012;7:e33788

200. Kuppers R, Zhao M, Hansmann ML and Rajewsky K. Tracing B cell development in human germinal centres by molecular analysis of single cells picked from histological sections. *EMBO J.* 1993;12:4955-4967

201. Lovatt D, Ruble BK, Lee J, et al. Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods.* 2014;11:190-196

202. Yu M, Stott S, Toner M, Maheswaran S and Haber DA. Circulating tumor cells: approaches to isolation and characterization. *J Cell Biol.* 2011;192:373-382

203. Choi JH, Ogunniyi AO, Du M, et al. Development and optimization of a process for automated recovery of single cells identified by microengraving. *Biotechnol Prog.* 2010;26:888-895

204. Altomare L, Borgatti M, Medoro G, et al. Levitation and movement of human tumor cells using a printed circuit board device based on software-controlled dielectrophoresis. *Biotechnol Bioeng.* 2003;82:474-479

205. Suarez-Quian CA, Goldstein SR, Pohida T, et al. Laser capture microdissection of single cells from complex tissues. *Biotechniques.* 1999;26:328-335

Reviews

Gawad C, Koh W and Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet.* 2016;17:175-188

Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges [version 1; referees: 2 approved]. F1000Research. 2016;5(F1000 Faculty Review):182

Grun D and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell.* 2015;163:799-810

Huang L, Ma F, Chapman A, Lu S and Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genomics Hum Genet.* 2015;16:79-102

Kanter I and Kalisky T. Single cell transcriptomics: methods and applications. *Front Oncol.* 2015;5:53

Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC and Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell.* 2015;58:610-620

Stegle O, Teichmann SA and Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet.* 2015;16:133-145

Wang Y and Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* 2015;58:598-609

References

Binan L, Mazzaferri J, Choquet K, et al. Live single-cell laser tag. *Nat Commun.* 2016;7:11636
Since single-cell sequencing methods often involve dissociation of cells and loss of spatial information, methods that retain spatial information in single-cell genomic analysis are critically important. The authors developed a cell-labeling via photobleaching (CLaP) method that combines cellular labeling with single-cell genomics. Individual cells are labeled in culture by laser photobleaching, followed by isolation based on a wide variety of distinguishing characteristics. In this study, the authors used CLaP to tag a number of different cells from lines grown in monolayers. They isolated individual cells using drop-based microfluidics and performed RNA-Seq using the HiSeq 2500 system. The ability to combine spatial information with single-cell genomics makes this method well suited for studying tissue heterogeneity.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System

Cottinet D, Condamine F, Bremond N, Griffiths AD, Rainey PB, et al. Lineage Tracking for Probing Heritable Phenotypes at Single-Cell Resolution. *PLoS One.* 2016;11:e0152395
Determining the genotype and phenotype of individual microbial cells is fundamentally important in understanding microbial evolution. Single-cell sequencing techniques, including WGS, currently allow detection of mutants at high resolution. However, similar approaches for phenotypic analysis are lacking. In this study, the authors present a drop-based microfluidics system that allows the genetic detection of heritable phenotypes in evolving bacterial populations. At various time intervals, they sampled cells and isolated them in 100 nL drops, then monitored growth monitored using a fluorescent protein reporter. The authors used this approach to follow E. coli populations during 30 days of starvation. The data showed that the phenotypic diversity of the E. coli increased with starvation, and single-cell sequencing was able to identify mutations corresponding to each phenotypic class.

Illumina Technology: HiSeq 2500 System

Bigdeli S, Dettloff RO, Frank CW, Davis RW and Crosby LD. A simple method for encapsulating single cells in alginate microspheres allows for direct PCR and whole genome amplification. *PLoS One.* 2015;10:e0117738
FACS, followed by MDA, is the current standard for single-cell sample processing. Processing cells in individual wells can increase the cost of single-cell sequencing, due to increased costs for reagents, consumables, and equipment for high-throughput liquid handling. To reduce the cost of parallel single-cell sequencing, the authors developed an approach for isolating single cells and preparing DNA libraries in bulk, followed by sorting afterward. They embedded *Rhodobacter sphaeroides* cells in alginate microspheres and subjected them to MDA. They extracted DNA from individual microspheres and sequenced it using the MiSeq system. This approach has the potential to improve the process for generating sequencing-ready DNA from many individually isolated cells.

Illumina Technology: MiSeq System

Bose S, Wan Z, Carr A, et al. Scalable microfluidics for single-cell RNA printing and sequencing. *Genome Biol.* 2015;16:120

In this study, the authors present a new scalable high-density microfluidic platform for solid-phase capture of RNA on glass coverslips or on polymer beads. They trapped single-cell lysates in sealed picoliter microwells capable of printing RNA on glass or capturing RNA on beads. They combined this sample preparation approach with a scalable technology for scRNA-Seq based on CEL-Seq. The technology is relatively inexpensive, with consumable costs of $0.10–$0.20 per cell and is capable of processing hundreds of individual cells in parallel.

Illumina Technology: TruSeq RNA-Seq Library Preparation Kit, NextSeq 500 and HiSeq 2500 Systems

Lee JH, Daugharthy ER, Scheiman J, et al. Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat Protoc.* 2015;10:442-458

scRNA-Seq can profile gene expression over the entire cell transcriptome, but cell isolation typically results in loss of spatial context. *In situ* hybridization is an excellent technique for identifying the location of gene expression, but it is restricted to a fixed number of genes. In this study, the authors present a protocol for *in situ* profiling of gene expression in cells and tissues. In this approach, RNA is converted into crosslinked cDNA amplicons and sequenced manually on a confocal microscope. The approach has the added benefit of enriching for context-specific transcripts over housekeeping/structural genes, while preserving the tissue architecture for transcript localization.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq System

Nishikawa Y, Hosokawa M, Maruyama T, Yamagishi K, Mori T and Takeyama H. Monodisperse Picoliter Droplets for Low-Bias and Contamination-Free Reactions in Single-Cell Whole Genome Amplification. *PLoS One.* 2015;10:e0138733

WGA is a critical component of single-cell sequencing pipelines, and MDA is the most common WGA method in single-cell sequencing. Despite its widespread use, MDA typically produces uneven genome coverage due to amplification bias and the formation of DNA chimeras. To overcome this limitation, the authors developed droplet MDA that minimizes these technical artifacts. They used microfluidics to compartmentalize extracted DNA fragments into 67 pL droplets, where the individual fragments were then amplified using MDA. This approach was validated by sequencing the droplet MDA products of *E. coli* cells, with genome recovery improving to 89%, compared to 59% using traditional MDA.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq System

Gierahn TM, Wadsworth MH, 2nd, Hughes TK, et al. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods.* 2017;14:395-398

Zhang Q, Wang T, Zhou Q, et al. Development of a facile droplet-based single-cell isolation platform for cultivation and genomic analysis in microorganisms. *Sci Rep.* 2017;7:41192

Mora-Castilla S, To C, Vaezeslami S, et al. Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing. *J Lab Autom.* 2016;21:557-567

Ungai-Salanki R, Gerecsei T, Furjes P, et al. Automated single cell isolation from suspension with computer vision. *Sci Rep.* 2016;6:20375

Xin Y, Kim J, Ni M, et al. Use of the Fluidigm C1 platform for RNA sequencing of single mouse pancreatic islet cells. *Proc Natl Acad Sci U S A.* 2016;113:3293-3298

Campton DE, Ramirez AB, Nordberg JJ, et al. High-recovery visual identification and single-cell retrieval of circulating tumor cells for genomic analysis using a dual-technology platform integrated with automated immunofluorescence staining. *BMC Cancer.* 2015;15:360

Szulwach KE, Chen P, Wang X, et al. Single-Cell Genetic Analysis Using Automated Microfluidics to Resolve Somatic Mosaicism. *PLoS One.* 2015;10:e0135007

For Research Use Only. Not for use in diagnostic procedures.

An overview of recent publications featuring Illumina tecnology   53

DATA ANALYSIS

Single-cell sequencing poses unique challenges for data analysis. Individual mammalian cells contain 50,000–300,000 transcripts, and gene expression values among individual cells can vary significantly.[206] Although several hundred thousand transcripts may be expressed per individual cell, up to 85% of these are present at only 1–100 copies.[207] Therefore, it is critically important in scRNA-Seq to capture low-abundance mRNA transcripts and amplify the synthesized cDNA to ensure that all transcripts are ultimately represented uniformly in the library.[208,209] Spike-in quantification standards of known abundance can help distinguish technical variability/ noise from biologically meaningful gene expression changes.[210] Molecular indexing can also correct for sequencing biases,[211,212] and recent improvements in automated sample handling can reduce technical variability even more.[213]

> "Single-cell analysis provides a new venue for bioinformatics, as bulk-cell data analysis methods may not be directly applicable to single-cell data." - Yalcin et al. 2016

DNA amplification and single-cell DNA-Seq technical artifacts can be reduced by using computational algorithms specifically designed for this purpose.[214] This section highlights some analysis methods used for single-cell sequencing data (Table 2).

Table 2. Summary of Data Analysis Methods for Single-Cell Sequencing

Name	Algorithm	Data	Description
Daley & Smith[215]	Coverage	DNA-Seq	Estimates gain in coverage with increased sequencing depth from initial shallow sequencing using Bayes Poisson models.
Varbin[216]	CNV	DNA-Seq	Uses variable bin sizes to call CNVs.
SNS[217]	CNV	DNA-Seq	Uses variable bin sizes to call copy numbers.
Xu et al.[218]	CNV	DNA-Seq	Uses a simplified negative binomial distribution to call CNVs.
siCHILD[219]	Haplotype & CNV	DNA-Seq	Determines haplotypes, CNV, and segregational origin haplotypes across the genome of a single cell via haplarithmisis.
Velvet-SC[220]	Assembly	DNA-Seq	Addresses low-coverage regions by using de Bruijn graphs with a dynamic cut-off.
SPAdes[221]	Assembly	DNA-Seq	Single-cell assembler for both single-cell and multi-cell assembly.
SmashCell[222]	Assembly annotation	DNA-Seq	Uses a tree with branches representing different choice of algorithm or parameters, mostly used in metagenomics.
Kim & Simon[223]	Evolutionary tree	DNA-Seq	Likelihood function for allele dropouts, Bayesian approach for mutation ordering, temporal relationships among mutation sites.
PyClone[224]	Clonal population	DNA-Seq	A statistical model for inference of clonal population structures in cancers.
Subramanian & Schwartz[225]	Clonal population	DNA-Seq	Computational approach for learning tumor progression from single-cell sequencing data using *k*-mer counts.

206. Wu AR, Neff NF, Kalisky T, et al. Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods*. 2014;11:41-46

207. Macaulay IC and Voet T. Single cell genomics: advances and future perspectives. *PLoS Genet*. 2014;10:e1004126

208. Bhargava V, Head SR, Ordoukhanian P, Mercola M and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep*. 2014;4:3678

209. Grun D, Kester L and van Oudenaarden A. Validation of noise models for single-cell transcriptomics. *Nat Methods*. 2014;11:637-640

210. Marinov GK, Williams BA, McCue K, et al. From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. *Genome Res*. 2014;24:496-510

211. Fu GK, Xu W, Wilhelmy J, et al. Molecular indexing enables quantitative targeted RNA sequencing and reveals poor efficiencies in standard library preparations. *Proc Natl Acad Sci U S A*. 2014;111:1891-1896

212. Islam S, Zeisel A, Joost S, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods*. 2014;11:163-166

213. Streets AM, Zhang X, Cao C, et al. Microfluidic single-cell whole-transcriptome sequencing. *Proc Natl Acad Sci U S A*. 2014;111:7048-7053

214. Yalcin D, Hakguder ZM and Otu HH. Bioinformatics approaches to single-cell analysis in developmental biology. *Mol Hum Reprod*. 2016;22:182-192

215. Daley T and Smith AD. Modeling genome coverage in single-cell sequencing. *Bioinformatics*. 2014;30:3159-3165

216. Baslan T, Kendall J, Rodgers L, et al. Genome-wide copy number analysis of single cells. *Nat Protoc*. 2012;7:1024-1041

217. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472:90-94

218. Xu B, Cai H, Zhang C, Yang X and Han G. Copy number variants calling for single cell sequencing data by multi-constrained optimization. *Comput Biol Chem*. 2016;

219. Zamani Esteki M, Dimitriadou E, Mateiu L, et al. Concurrent whole-genome haplotyping and copy-number profiling of single cells. *Am J Hum Genet*. 2015;96:894-912

220. Chitsaz H, Yee-Greenbaum JL, Tesler G, et al. Efficient de novo assembly of single-cell bacterial genomes from short-read data sets. *Nat Biotech*. 2011;29:915-921

221. Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19:455-477

222. Harrington ED, Arumugam M, Raes J, Bork P and Relman DA. SmashCell: a software framework for the analysis of single-cell amplified genome sequences. *Bioinformatics*. 2010;26:2979-2980

223. Kim KI and Simon R. Using single cell sequencing data to model the evolutionary history of a tumor. *BMC Bioinformatics*. 2014;15:27

224. Roth A, Khattra J, Yap D, et al. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods*. 2014;11:396-398

225. Subramanian A and Schwartz R. Reference-free inference of tumor phylogenies from single-cell sequencing data. *BMC Genomics*. 2015;16 Suppl 11:S7

CheckM[226]	Genome quality	DNA-Seq	Estimates genome completeness and contamination using marker genes specific to a genome's inferred lineage within a reference genome tree.
Ji et al.[227]	Allele dropout	DNA-Seq	Method to control for false negatives from single-cell amplification data due to allele dropout in mutation calling.
GRM[228]	Normalization	RNA-Seq	Fits polynomial gamma regression model to fragments per kilobase of transcript per million mapped reads (FPKM) data from spike-ins.
SAMstrt[229]	Normalization	RNA-Seq	Uses spike-in controls to normalize and estimate transcript numbers per cell; tolerates variations in sequencing depth.
BASiCS[230]	Identifying variable genes	RNA-Seq	Fully Bayesian approach that jointly models extrinsic spike-in molecules with genes from cells of interest.
Brennecke et al.[231]	Identifying variable genes	RNA-Seq	Statistical method that allows the user to assess whether observed gene variation provides evidence of high biological variability.
Kim et al.[232]	Identifying variable genes	RNA-Seq	Uses spike-ins to estimate parameters related to technical noise, allowing for differences in variability across cells.
scLVM[233]	Noise reduction	RNA-Seq	Single-cell latent variable model estimates proportion of variation associated with hidden factors to identify subpopulations.
OEfinder[234]	Noise reduction	RNA-Seq	Uses orthogonal polynomial regression to identify genes with significantly increased expression artifacts in specific capture sites on the Fluidigm C1 platform.
PCA/t-SNE[235]	Subpopulation ID	RNA-Seq	Linear/nonlinear dimension-reduction approach for unsupervised clustering of cells.
ZIFA[236]	Subpopulation ID	RNA-Seq	Dimensionality reduction method that models dropout characteristics to improve simulated and biological data sets.
Destiny[237]	Subpopulation ID	RNA-Seq	Extends diffusion maps to handle zeros and sampling data heterogeneities in single-cell data.
SNN-Cliq[238]	Subpopulation ID	RNA-Seq	Uses shared nearest neighbor-based similarity graphs. Partitioning of the graphs automatically identifies subgroups of cells.
RaceID[239]	Subpopulation ID	RNA-Seq	Two technical noise sources: random sampling (Poissonian) noise and variability due to sequencing efficiency characterization.
SCUBA[240]	Subpopulation ID	RNA-Seq	Uses k-means to cluster data along a binary tree detailing bifurcation events for time-course data.
BackSPIN[241]	Subpopulation ID	RNA-Seq	A divisive biclustering method based on sorting points into neighborhoods.
PAGODA[242]	Subpopulation ID	RNA-Seq	Principal component analysis (PCA) for gene sets to identify those where first PCA exceeds significantly exceeds genome-wide background expectation.
MAST[243]	Differential detection	RNA-Seq	Two-part generalized linear model characterizing expression heterogeneity by parameterizing stochastic dropout and bimodal expression distributions.
SCDE[244]	Differential detection	RNA-Seq	Single-cell differential expression uses a separate model for dropouts and a Bayesian model for differential expression.
scDD[245]	Differential detection	RNA-Seq	Bayesian modeling framework characterizing expression within a biological condition and with differential distributions across conditions.
Monocle[246]	Pseudotem poral ordering	RNA-Seq	Uses independent component analysis for dimension reduction and minimum spanning tree for cell ordering.
Waterfall[247]	Pseudotem poral ordering	RNA-Seq	Clustering method for determining the temporal ordering of the expression profiles of individual cells assayed by RNA-Seq.

226. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P and Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25:1043-1055

227. Ji C, Miao Z and He X. A simple strategy for reducing false negatives in calling variants from single-cell sequencing data. *PLoS One.* 2015;10:e0123789

228. Ding B, Zheng L, Zhu Y, et al. Normalization and noise reduction for single cell RNA-seq experiments. *Bioinformatics.* 2015;31:2225-2227

229. Katayama S, Tohonen V, Linnarsson S and Kere J. SAMstrt: statistical test for differential expression in single-cell transcriptome with spike-in normalization. *Bioinformatics.* 2013;29:2943-2945

230. Vallejos CA, Marioni JC and Richardson S. BASiCS: Bayesian Analysis of Single-Cell Sequencing Data. *PLoS Comput Biol.* 2015;11:e1004333

231. Brennecke P, Anders S, Kim JK, et al. Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods.* 2013;10:1093-1095

232. Kim JK, Kolodziejczyk AA, Illicic T, Teichmann SA and Marioni JC. Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun.* 2015;6:8687

233. Buettner F, Natarajan KN, Casale FP, et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol.* 2015;33:155-160

234. Leng N, Choi J, Chu LF, Thomson JA, Kendziorski C and Stewart R. OEFinder: a user interface to identify and visualize ordering effects in single-cell RNA-seq data. *Bioinformatics.* 2016;32:1408-1410

235. Van der Maaten L and Hinton G. Visualizing data using t-SNE. *J Mach Learn Res.* 2008;9:2579-2605

236. Pierson E and Yau C. ZIFA: Dimensionality reduction for zero-inflated single-cell gene expression analysis. *Genome Biol.* 2015;16:241

237. Haghverdi L, Buettner F and Theis FJ. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics.* 2015;31:2989-2998

238. Xu C and Su Z. Identification of cell types from single-cell transcriptomes using a novel clustering method. *Bioinformatics.* 2015;31:1974-1980

239. Grun D, Lyubimova A, Kester L, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature.* 2015;525:251-255

240. Marco E, Karp RL, Guo G, et al. Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. *Proc Natl Acad Sci U S A.* 2014;111:E5643-5650

241. Zeisel A, Munoz-Manchado AB, Codeluppi S, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015;347:1138-1142

242. Fan J, Salathia N, Liu R, et al. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat Methods.* 2016;13:241-244

243. Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 2015;16:278

Sincell[248]	Pseudotem poral ordering	RNA-Seq	Method to assess cell-state hierarchies from single-cell data using a metric to assess cell-to-cell similarities and a graph-building algorithm.
Oscope[249]	Pseudotem poral ordering	RNA-Seq	Uses coregulation information among oscillators to identify groups of putative oscillating genes and cyclic order of samples for each group.
Wanderlust[250]	Pseudotem poral ordering	RNA-Seq	A graph-based trajectory detection algorithm that orders cells to a unified trajectory based on their developmental maturity.
Seurat[251]	Cellular localization	RNA-Seq	A computational strategy to infer cellular localization by integrating scRNA-Seq data with *in situ* RNA patterns.
Achim et al.[252]	Cellular localization	RNA-Seq	Compares complete specificity-weighted mRNA profiles of a cell with positional gene expression profiles derived from a gene expression atlas.
VDJPuzzle[253]	TCR reconstruction	RNA-Seq	Reconstructs the native TCRαβ from individual antigen-specific T cells and links these with the single-cell gene expression profiles.

Reviews

Hardwick SA, Deveson IW and Mercer TR. Reference standards for next-generation sequencing. *Nat Rev Genet.* 2017;18:473-484

Bacher R and Kendziorski C. Design and computational analysis of single-cell RNA-sequencing experiments. *Genome Biol.* 2016;17:63

Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges [version 1; referees: 2 approved]. F1000Research. 2016;5(F1000 Faculty Review):182

Yalcin D, Hakguder ZM and Otu HH. Bioinformatics approaches to single-cell analysis in developmental biology. *Mol Hum Reprod.* 2016;22:182-192

Yu P and Lin W. Single-cell Transcriptome Study as Big Data. *Genomics Proteomics Bioinformatics.* 2016;14:21-30

Grun D and van Oudenaarden A. Design and Analysis of Single-Cell Sequencing Experiments. *Cell.* 2015;163:799-810

Stegle O, Teichmann SA and Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet.* 2015;16:133-145

Woodhouse S, Moignard V, Gottgens B and Fisher J. Processing, visualising and reconstructing network models from single-cell data. *Immunol Cell Biol.* 2015;94:256-265

References

Dong X, Zhang L, Milholland B, et al. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat Methods.* 2017;14:491-493
NGS is routinely used to detect SNVs in genomic DNA from tissues, but analyzing SNVs in single cells is prone to artifacts that are associated with WGA. To overcome this technical challenge, the authors developed single-cell multiple displacement amplification (SCMDA) and an associated single-cell variant-calling algorithm, SCaller. In this study, the authors isolated nonamplified genomic DNA from fibroblast clones. They also isolated individual cells from these clones and used SCMDA to amplify the single-cell genomic DNA. They performed whole-genome sequencing using the HiSeq 2500 and HiSeq X Ten Systems on SCMDA-amplified and nonamplified samples and identified SNVs using SCaller. By comparing SNVs from single cells and parent clones, the authors validated their procedure for accurately analyzing SNVs in single-cell genomics.

Illumina Technology: HiSeq 2500 and HiSeq X Ten Systems

244. Kharchenko PV, Silberstein L and Scadden DT. Bayesian approach to single-cell differential expression analysis. *Nat Methods.* 2014;11:740-742

245. Korthauer KD, Chu L-F, Newton MA, et al. A statistical approach for identifying differential distributions in single-cell RNA-seq experiments. *Genome Biology.* 2016;17:222

246. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol.* 2014;32:381-386

247. Shin J, Berg DA, Zhu Y, et al. Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. *Cell Stem Cell.* 2015;17:360-372

248. Julia M, Telenti A and Rausell A. Sincell: an R/Bioconductor package for statistical assessment of cell-state hierarchies from single-cell RNA-seq. *Bioinformatics.* 2015;31:3380-3382

249. Leng N, Chu LF, Barry C, et al. Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. *Nat Methods.* 2015;12:947-950

250. Bendall SC, Davis KL, Amir el AD, et al. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell.* 2014;157:714-725

251. Satija R, Farrell JA, Gennert D, Schier AF and Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* 2015;33:495-502

252. Achim K, Pettit JB, Saraiva LR, et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol.* 2015;33:503-509

253. Eltahla AA, Rizzetto S, Rasoli M, et al. Linking the T cell receptor to the single cell transcriptome in antigen-specific human T cells. *Immunol Cell Biol.* 2016;94:604-611

Mende DR, Aylward FO, Eppley JM, Nielsen TN and DeLong EF. Improved Environmental Genomes via Integration of Metagenomic and Single-Cell Assemblies. *Front Microbiol.* **2016;7:143**

Single-cell genomics has led to a number of individual draft genomes for uncultivated microbes; however, MDA artifacts during the amplification step lead to incomplete and uneven coverage. Metagenomic data sets do not suffer the same sequence bias, but the genomic complexity of microbial communities precludes the recovery of draft genomes. In this study, the authors developed a new method for generating population genome assemblies from metagenomic-guided, single-cell amplified genome assembly data. They validated the approach by completing single-cell amplified genomes for Marine Group 1 Thaumarchaeota and SAR324 clade bacterioplankton. The improved method assembly of the SAR324 clade genome revealed the presence of many genes not present in the single-cell amplified genome.

Illumina Technology: TruSeq LT Nano Kit, MiSeq System

Fan J, Salathia N, Liu R, et al. Characterizing transcriptional heterogeneity through pathway and gene set overdispersion analysis. *Nat Methods.* **2016;13:241-244**

scRNA-Seq methods present an unbiased approach for studying complex tissues and diseases. However, the data suffer from high levels of technical noise and strong dependence on expression magnitude. Cell-to-cell differences can prove challenging when clustering cells based on important biological differences. For example, partitioning methods including *k*-means clustering and a BackSPIN algorithm may classify cells based on cell cycle rather than tissue-specific signaling. The authors introduce pathway and gene set overdispersion analysis (PAGODA) that overcomes this challenge by detecting all significant and potentially overlapping pathways in which measured cells can be classified.

Illumina Technology: HiSeq 2000 System

Ilicic T, Kim JK, Kolodziejczyk AA, et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol.* **2016;17:29**

Modern single-cell sequencing techniques, particularly those involving massively parallel approaches, often result in the isolation of cells that are stressed, broken, or killed. These low-quality cells can lead to data artifacts, and they must be excluded from analysis. In this study, the authors present the first tool for scRNA-Seq that can process and remove low-quality cells in a simple and rigorous way. The analysis pipeline uses a highly-curated set of 20 biologic and technical features that are incorporated into a machine-learning algorithm. The authors validated the approach on CD4+ T cells, bone marrow dendritic cells, and mouse ESCs. The method also defined a new type of low-quality cell that was not detectable visually.

Illumina Technology: HiSeq 2000 System

Buettner F, Natarajan KN, Casale FP, et al. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol.* **2015;33:155-160**

scRNA-Seq data sets suffer from inherent technical noise that can challenge the identification of cell subpopulations. To overcome this challenge, as well as unknown hidden factors affecting gene expression heterogeneity, the authors developed a model (scLVM) to account for unobserved factors in RNA-Seq data sets, and validated their model using individual mouse ESCs. They also used the HiSeq 2000 system to perform RNA-Seq of individual T cells over the course of naïve T cells differentiating into T$_H$2 cells. They applied the scLVM model to differentiating T-cell RNA-Seq data sets and corrected for cell cycle gene expression. They were able to identify 2 subpopulations of differentiating T cells that were not revealed by using nonlinear PCA or k-means clustering alone.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2000 System

Grun D, Lyubimova A, Kester L, et al. Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature.* **2015;525:251-255**

Characterizing constituent cell types is critical for understanding the function of a given organ or tissue. Existing methods for characterizing cell types involve imaging and isolating cells based on specific markers, but this approach is challenging if the cell types are rare, such as CSCs or CTCs. In this study, the authors used the HiSeq 2500 system to perform RNA-Seq on hundreds of randomly selected cells from mouse intestinal organoids. To characterize cell subpopulations within the organoids, they developed RaceID, a computational method for identifying rare cell types in complex populations of cells. They validated this algorithm by identifying a single hormone-producing cell type in a population of sampled organoid cells, and they identified Reg4 as a novel marker for these rare enteroendocrine cells. Finally, they used Reg4 to capture these rare cells to investigate their genetic heterogeneity, identifying a number of enteroendocrine lineages.

Illumina Technology: HiSeq 2500 System

Leng N, Chu LF, Barry C, et al. Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. *Nat Methods.* **2015;12:947-950**

scRNA-Seq has the potential to capture oscillation dynamics in populations of individual cells and to discover oscillations missed in bulk sequencing experiments. However, continuous RNA-Seq time-series experiments are not feasible, and synchronization may not be possible for most oscillatory systems. The Monocle[254] computational algorithm was developed previously to address this challenge in scRNA-Seq data by pseudo-temporal ordering of the data from a few different time points. In this study, the authors developed Oscope, a computational method that identifies and characterizes the transcriptional dynamics of oscillating genes using scRNA-Seq data from unsynchronized cells. They validated Oscope by applying the model to various scRNA-Seq Illumina data sets, including human ESCs, and they discovered an oscillatory pattern related to capture-site and output-well positions on the Fluidigm C1 chip.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System
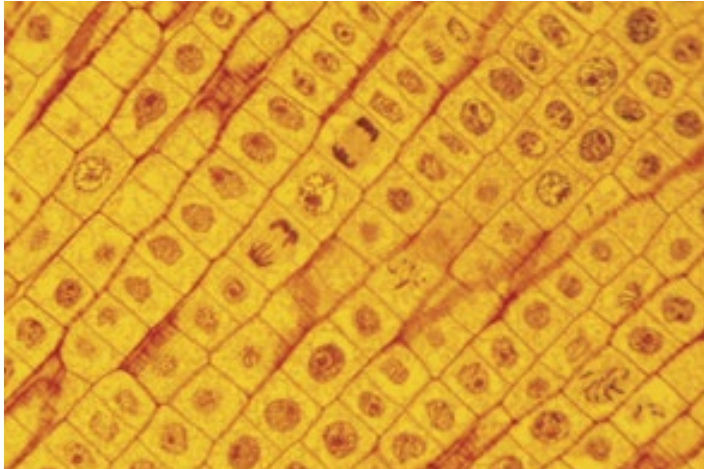
Statistical approaches can help identify and characterize oscillating genes, such as cell cycle genes, in scRNA-Seq data sets.

Satija R, Farrell JA, Gennert D, Schier AF and Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol.* **2015;33:495-502**

scRNA-Seq is an established method for discovering novel cell types, understanding regulatory networks, and reconstructing developmental processes. However, scRNA-Seq typically involves dissociating cells from tissues and thus disrupting their native spatial context. To capture spatial context in scRNA-Seq data, the authors developed Seurat, a computational strategy that combines scRNA-Seq with complementary *in situ* hybridization data for a smaller set of "landmark" genes that guides spatial assignment. They validated Seurat by spatially mapping 851 individual cells from dissociated zebrafish embryos and creating a transcriptome-wide map of spatial patterning. Seurat was able to localize rare subpopulations of cells correctly, and it could map spatially restricted cells as well as those with a more scattered pattern of expression.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System

Shin J, Berg DA, Zhu Y, et al. Single-Cell RNA-Seq with Waterfall Reveals Molecular Cascades underlying Adult Neurogenesis. *Cell Stem Cell.* **2015;17:360-372**

Somatic stem cells contribute to tissue development and regeneration, but a systematic molecular analysis of stem-cell behavior has proved elusive due to challenges in cellular heterogeneity. In this study, the authors used scRNA-Seq to characterize the developmental dynamics of adult hippocampal qNSCs. They also developed a bioinformatic pipeline, called Waterfall that quantified single-cell expression data along a reconstructed developmental trajectory. The combination of scRNA-Seq and Waterfall identified molecular signatures of adult qNSCs, and defined molecular cascades underlying qNSC activation and neurogenesis.

Illumina Technology: HiSeq 2500 System

254. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol.* 2014;32:381-386

Zamani Esteki M, Dimitriadou E, Mateiu L, Melotte C, Van der Aa N, et al. Concurrent whole-genome haplotyping and copy-number profiling of single cells. *Am J Hum Genet.* **2015;96:894-912**
Before analyzing single-cell DNA-Seq data, DNA copy-number aberrations must be differentiated from WGA artifacts. This requirement makes DNA copy-number profiling and haplotyping of single-cell sequencing data challenging. In this study, the authors developed a single-cell genome analysis method that determined haplotypes and copy number across the genome of a single cell—a process called, haplarithmisis. The method deciphered SNP allele fractions of single cells and integrated these data into a computational workflow for imputation of linked disease variants (siCHILD). The authors validated the method by determination of haplotypes carrying disease alleles in single-cell genomes from individual lymphocytes and human blastomeres derived from human IVF embryos.

Illumina Technology: TruSeq DNA LT Sample Preparation Kit, HumanCytoSNP-12v2.1 BeadChips, HiSeq 2000/2500 Systems

Zhang CZ, Adalsteinsson VA, Francis J, Cornils H, Jung J, et al. Calibrating genomic and allelic coverage bias in single-cell sequencing. *Nat Commun.* **2015;6:6822**
In single-cell DNA-Seq, sequence artifacts are introduced by requisite DNA amplification methods, such as MDA[255] and MALBAC.[256] In this study, the authors developed a new statistical method for quantitative assessment of single-cell DNA amplification bias due to WGA. By comparing MDA and MALBAC DNA libraries, they provided a benchmark comparison of single-cell libraries generated by MDA and MALBAC and also identified universal features of genomic coverage bias at the amplicon level. Their statistical models allowed for calibration of allelic bias in single-cell WGA data.

Illumina Technology: MiSeq and HiSeq 2500 Systems

Bacher R, Chu LF, Leng N, et al. SCnorm: robust normalization of single-cell RNA-seq data. *Nat Methods.* 2017;14:584-586

Gong W, Rasmussen TL, Singh BN, Koyano-Nakagawa N, Pan W and Garry DJ. Dpath software reveals hierarchical haemato-endothelial lineages of Etv2 progenitors based on single-cell transcriptome analysis. *Nat Commun.* 2017;8:14362

Wang B, Zhu J, Pierson E, Ramazzotti D and Batzoglou S. Visualization and analysis of single-cell RNA-seq data by kernel-based similarity learning. *Nat Methods.* 2017;14:414-416

Zhu X, Ching T, Pan X, Weissman SM and Garmire L. Detecting heterogeneity in single-cell RNA-Seq data by non-negative matrix factorization. *PeerJ.* 2017;5:e2888

Eltahla AA, Rizzetto S, Rasoli M, et al. Linking the T cell receptor to the single cell transcriptome in antigen-specific human T cells. *Immunol Cell Biol.* 2016;94:604-611

Knouse KA, Wu J and Amon A. Assessment of megabase-scale somatic copy number variation using single-cell sequencing. *Genome Res.* 2016;26:376-384

Mende DR, Aylward FO, Eppley JM, Nielsen TN and DeLong EF. Improved Environmental Genomes via Integration of Metagenomic and Single-Cell Assemblies. *Front Microbiol.* 2016;7:143

Achim K, Pettit JB, Saraiva LR, et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol.* 2015;33:503-509

Finak G, McDavid A, Yajima M, et al. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* 2015;16:278

Ji C, Miao Z and He X. A simple strategy for reducing false negatives in calling variants from single-cell sequencing data. *PLoS One.* 2015;10:e0123789

Scialdone A, Natarajan KN, Saraiva LR, et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods.* 2015;85:54-61

Vallejos CA, Marioni JC and Richardson S. BASiCS: Bayesian Analysis of Single-Cell Sequencing Data. *PLoS Comput Biol.* 2015;11:e1004333

255. Dean FB, Nelson JR, Giesler TL and Lasken RS. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* 2001;11:1095-1099
256. Zong C, Lu S, Chapman AR and Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science.* 2012;338:1622-1626

DNA METHODS

DNA replication during mitosis is not perfect, and progressive generations of cells accumulate somatic mutations. Consequently, each cell in our body has a unique genomic signature, which allows the reconstruction of cell-lineage trees with very high precision.[257] These cell-lineage trees can predict the existence of small subpopulations of stem cells. This information is instructive in cancer development,[258,259] as well as in preimplantation and genetic diagnoses.[260,261,262]

> "Single-cell sequencing has uncovered the breadth of genomic heterogeneity between cells in a variety of contexts, including somatic aneuploidy in the mammalian brain and intratumor heterogeneity."
>
> - Vitak et al. 2017

Single-cell DNA-Seq can identify acquired somatic mutations and CNVs, allowing researchers to trace back lineages of differentiated cells.[263] Single-cell genomics is also an effective approach for characterizing microorganisms that are difficult or impossible to culture *in vitro*. Advances in single-cell genomics have led to improvements in diagnosing infectious disease outbreaks, understanding antibiotic-resistant strains and food-borne pathogens, and classifying microbial diversity in the environment or in the gut.[264,265,266] Newer techniques are using multiplexing and microfluidics platforms to improve the throughput of single-cell DNA-Seq and to lower costs.[267,268,269]

This section highlights some single-cell DNA-Seq methods and recent publications demonstrating how Illumina technology is being used in single-cell DNA-Seq techniques. To learn more about Illumina sequencing methods, visit www.illumina.com/techniques/sequencing.html.
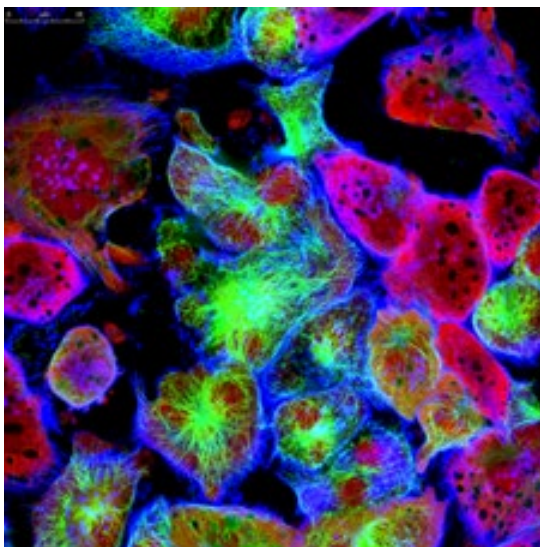
257. Frumkin D, Wasserstrom A, Kaplan S, Feige U and Shapiro E. Genomic variability within an organism exposes its cell lineage tree. *PLoS Comput Biol.* 2005;1:e50

258. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature.* 2011;472:90-94

259. Potter NE, Ermini L, Papaemmanuil E, et al. Single-cell mutational profiling and clonal phylogeny in cancer. *Genome Res.* 2013;23:2115-2125

260. Blainey PC. The future is now: single-cell genomics of bacteria and archaea. *FEMS Microbiol Rev.* 2013;37:407-427

261. Van der Aa N, Zamani Esteki M, Vermeesch JR and Voet T. Preimplantation genetic diagnosis guided by single-cell genomics. *Genome Med.* 2013;5:71

262. Hou Y, Fan W, Yan L, et al. Genome analyses of single human oocytes. *Cell.* 2013;155:1492-1506

263. Shapiro E, Biezuner T and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

264. Bergholz TM, Moreno Switt AI and Wiedmann M. Omics approaches in food safety: fulfilling the promise? *Trends Microbiol.* 2014;22:275-281

265. Stepanauskas R. Single cell genomics: an individual look at microbes. *Curr Opin Microbiol.* 2012;15:613-620

266. Hug LA, Baker BJ, Anantharaman K, et al. A new view of the tree of life. *Nature Microbiology.* 2016;1:16048

267. Baslan T, Kendall J, Ward B, et al. Optimizing sparse sequencing of single cells for highly multiplex copy number profiling. *Genome Res.* 2015;25:714-724

268. Fu Y, Li C, Lu S, et al. Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proc Natl Acad Sci U S A.* 2015;112:11923-11928

269. Leung ML, Wang Y, Kim C, et al. Highly multiplexed targeted DNA sequencing from single nuclei. *Nat Protoc.* 2016;11:214-235

Single-cell genomics can help characterize genetic and cellular heterogeneity within tumors.

For Research Use Only. Not for use in diagnostic procedures.

Reviews

Gawad C, Koh W and Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet.* 2016;17:175-188

Huang L, Ma F, Chapman A, Lu S and Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. *Annu Rev Genomics Hum Genet.* 2015;16:79-102

Voet T and Van Loo P. SNES makes sense? Single-cell exome sequencing evolves. *Genome Biol.* 2015;16:86

Wang Y and Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* 2015;58:598-609

270. Li H, Courtois ET, Sengupta D, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat Genet. 2017;49:708-718
270. Gerber T, Willscher E, Loeffler-Wirth H, et al. Mapping heterogeneity in patient-derived melanoma cultures by single-cell RNA-seq. Oncotarget. 2017;8:846-862

References

Chen C, Xing D, Tan L, et al. Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). *Science.* 2017;356:189-194
One of the challenges in single-cell genomics is accurate CNV analysis. Single-cell genomes must be amplified prior to copy-number analysis, and although there are several different WGA techniques, all suffer from a degree of sequence bias and low amplification fidelity. In this study, the authors developed Linear Amplification via Transposon Insertion (LIANTI), a superior WGA method. In this method, genomic DNA from single cells is fragmented and tagged using a LIANTI transposon, followed by addition of double-strand T7 promoters at both ends of the fragmented strands. These DNAs are then transcribed into genomic RNAs, followed by RNase digestion and second-strand synthesis to generate barcoded and double-stranded LIANTI amplicons. These amplicons are then ready for library prep and subsequent NGS. The authors demonstrated that LIANTI provided accurate CNV detection in individual human BJ cells, with the smallest false-positive rate of any single-cell WGA method.

Illumina Technology: HiSeq 2500/4000 Systems

Dong X, Zhang L, Milholland B, Lee M, Maslov AY, et al. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat Methods.* 2017;14:491-493
NGS is routinely used to detect SNVs in genomic DNA from tissues, but analyzing SNVs in single cells is prone to artifacts that are associated with WGA. To overcome this technical challenge, the authors developed single-cell multiple displacement amplification (SCMDA) and an associated single-cell variant-calling algorithm, SCaller. In this study, the authors isolated nonamplified genomic DNA from fibroblast clones. They also isolated individual cells from these clones and used SCMDA to amplify the single-cell genomic DNA. They performed whole-genome sequencing using the HiSeq 2500 and HiSeq X Ten Systems on SCMDA-amplified and nonamplified samples and identified SNVs using SCaller. By comparing SNVs from single cells and parent clones, the authors validated their procedure for accurately analyzing SNVs in single-cell genomics.

Illumina Technology: HiSeq 2500 and HiSeq X Ten Systems

Martelotto LG, Baslan T, Kendall J, et al. Whole-genome single-cell copy number profiling from formalin-fixed paraffin-embedded samples. *Nat Med.* 2017;23:376-385
Single-cell sequencing studies have revealed that tumors consist of diverse populations of genetically distinct cells.[270,271] Single-cell sequencing can also trace cell lineages, as well as identify genetic mutations that might be otherwise hidden in bulk tissue samples. Most human tumor samples are formalin-fixed and paraffin-embedded (FFPE) for clinical diagnostic purposes. While this technique does preserve the sample, formalin fixation chemically alters the genomic DNA and makes it unsuitable for single-cell sequencing. To address this challenge, the authors developed a method for single-cell, whole-genome copy-number profiling of DNA extracted from FFPE tissues. They observed matching copy-number profiles of nuclei isolated from matched FFPE, frozen cancer cell lines, and fresh neoplastic tissue. This approach enables single-cell genomics applications for research or clinical analysis of FFPE samples.

Illumina Technology: HiSeq 2000 System

Hou Y, Guo H, Cao C, Li X, Hu B, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26:304-319

To fully understand the mechanisms by which the genome, transcriptome, and DNA methylome interact at the single-cell level, these 3 separate methods ideally should be applied to the same individual cell. In this study, the authors report scTrio-Seq, a method that can analyze genomic CNVs, the DNA methylome, and the transcriptome of an individual mammalian cell simultaneously. They used scTrio-Seq in 25 individual hepatocellular carcinoma primary cells to identify 2 subpopulations of cells. They also found that large-scale CNVs can cause proportional changes in RNA expression in subsets of genes, but the CNVs did not affect DNA methylation in the relevant genomic regions.

Illumina Technology: HiSeq 2000/2500 Systems

Baslan T, Kendall J, Ward B, et al. Optimizing sparse sequencing of single cells for highly multiplex copy number profiling. *Genome Res.* 2015;25:714-724

Tumor cell heterogeneity is known to play a role in disease progression, therapeutic resistance, and metastasis. However, our understanding of tumor heterogeneity is limited, due to a lack of sensitive approaches for interrogating genetic heterogeneity at a genome-wide scale. In this study, the authors developed a DNA amplification method that combined bioinformatic and molecular approaches to enable highly multiplexed single-cell sequencing. They applied this technique to produce genome-wide CNV profiles of up to 100 individual human cancer cells as well as biopsied tissues on a single lane of a HiSeq system. The method enables rapid profiling of thousands of single-cell genomes.

Illumina Technology: HiSeq System

Fu Y, Li C, Lu S, et al. Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proc Natl Acad Sci U S A.* 2015;112:11923-11928

Current WGA methods can be limited by fluctuations in amplification yield, as well as false-positive and false–negative SNV errors. The authors developed an emulsion-based amplification method (eWGA) that can overcome amplification bias and detect SNVs with high accuracy. Single-cell DNA is divided into aqueous droplets in oil where DNA fragments can be amplified to saturation, minimizing the differences in amplification gain among the emulsified fragments. The method is compatible with MDA and can detect CNVs and SNVs in single cells with improved amplification evenness and accuracy.

Illumina Technology: MiSeq and HiSeq 2500 Systems

Borgstrom E, Paterlini M, Mold JE, Frisen J and Lundeberg J. Comparison of whole genome amplification techniques for human single cell exome sequencing. *PLoS One.* 2017;12:e0171566

Stepanauskas R, Fergusson EA, Brown J, et al. Improved genome recovery and integrated cell-size analyses of individual uncultured microbial cells and viral particles. *Nat Commun.* 2017;8:84

Vitak SA, Torkenczy KA, Rosenkrantz JL, et al. Sequencing thousands of single-cell genomes with combinatorial indexing. *Nat Methods.* 2017;14:302-308
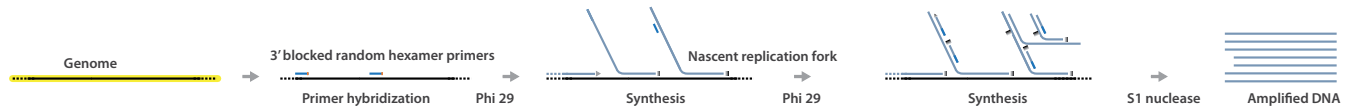
Zahn H, Steif A, Laks E, et al. Scalable whole-genome single-cell library preparation without preamplification. *Nature Methods.* 2017;14:167-173

Li N, Wang L, Wang H, Ma M, Wang X, et al. The Performance of Whole Genome Amplification Methods and Next-Generation Sequencing for Pre-Implantation Genetic Diagnosis of Chromosomal Abnormalities. *J Genet Genomics.* 2015;42:151-159

Ning L, Li Z, Wang G, et al. Quantitative assessment of single-cell whole genome amplification methods for detecting copy number variation using hippocampal neurons. *Sci Rep.* 2015;5:11415

Multiple-Strand Displacement Amplification

MDA is commonly used for sequencing microbial genomes due to its ability to amplify templates larger than 0.5 Mbp, but it can also be used to study genomes of other sizes. In this method, 3'-blocked random hexamer primers are hybridized to the template, followed by synthesis with Phi 29 polymerase. Phi 29 performs strand-displacement DNA synthesis, allowing for efficient and rapid DNA amplification. Deep sequencing of the amplified DNA allows for accurate representation of reads, while sequencing depth provides better alignment and consensus for sequences (Table 3).

272. Dean FB, Nelson JR, Giesler TL and Lasken RS. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* 2001;11:1095-1099

273. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature.* 2011;472:90-94

274. Woyke T, Sczyrba A, Lee J, et al. Decontamination of MDA reagents for single cell whole genome amplification. *PLoS One.* 2011;6:e26161

275. Leung ML, Wang Y, Waters J and Navin NE. SNES: single nucleus exome sequencing. *Genome Biol.* 2015;16:55

A schematic overview of MDA.

Table 3. Advantages and Disadvantages of MDA

Advantages	Disadvantages
• Template can be circular DNA (plasmids, bacterial DNA). • Can sequence large templates. • Can perform single-cell sequencing or sequencing for samples with very limited starting material.	• Strong amplification bias. Genome coverage as low as ~6%. [273] • PCR biases can underrepresent GC-rich templates. • Contaminated reagents can impact results.[274]

References

Dong X, Zhang L, Milholland B, Lee M, Maslov AY, et al. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat Methods.* **2017;14:491-493**
NGS is routinely used to detect SNVs in genomic DNA from tissues, but analyzing SNVs in single cells is prone to artifacts that are associated with WGA. To overcome this technical challenge, the authors developed single-cell multiple displacement amplification (SCMDA) and an associated single-cell variant-calling algorithm, SCaller. In this study, the authors isolated nonamplified genomic DNA from fibroblast clones. They also isolated individual cells from these clones and used SCMDA to amplify the single-cell genomic DNA. They performed whole-genome sequencing using the HiSeq 2500 and HiSeq X Ten Systems on SCMDA-amplified and nonamplified samples and identified SNVs using SCaller. By comparing SNVs from single cells and parent clones, the authors validated their procedure for accurately analyzing SNVs in single-cell genomics.

Illumina Technology: HiSeq 2500 and HiSeq X Ten Systems

Leung ML, Wang Y, Kim C, et al. Highly multiplexed targeted DNA sequencing from single nuclei. *Nat Protoc.* **2016;11:214-235**
This method is a refinement to SNES and includes the addition of DNA barcoding to allow multiplexing of 48–96 individual cells into single sequencing reactions. Compared to SNES, this new technique has higher throughput and reduced cost. The authors suggest that SNES is more suitable for detecting point mutations and indels at base-pair resolution.

Illumina Technology: HiSeq 2000 System

Leung ML, Wang Y, Waters J and Navin NE. SNES: single nucleus exome sequencing. *Genome Biol.* **2015;16:55**
Despite the great potential of single-cell sequencing methods to advance the understanding of tissue heterogeneity, current single-cell DNA-Seq methods are challenged by technical errors and poor physical coverage data. In this study, the authors developed single-nucleus exome sequencing (SNES), a single-cell DNA-Seq method that combines flow-sorting of G1/0- or G2/M nuclei, time-limited MDA, exome capture using the TruSeq Exome Enrichment Kit, and sequencing on the HiSeq 2000 system. They validated SNES by sorting and sequencing single nuclei from a fibroblast cell line. The method generated 96% coverage of individual cells and demonstrated 92% detection efficiency for SNVs and 85% for indels in single cells.

Illumina Technology: TruSeq Exome Enrichment Kit, HiSeq 2000 System

For Research Use Only. Not for use in diagnostic procedures.

Ning L, Li Z, Wang G, Hu W, Hou Q, et al. Quantitative assessment of single-cell whole genome amplification methods for detecting copy number variation using hippocampal neurons. *Sci Rep.* **2015;5:11415**

In this study, the authors compared MDA, MALBAC, and GenomePlex amplification methods in sequencing of individual hippocampal neurons. They amplified genomic DNA from individual hippocampal neurons using 3 different amplification methods, followed by sequencing at shallow depth on a HiSeq 2000 system. Their results showed that single-cell sequencing results from MALBAC and GenomePlex methods were highly reproducible and had high success rates. MALBAC did display significant GC bias, but it was overcome by using bioinformatics tools. Overall, they determined that MALBAC and GenomePlex performed better for detecting CNVs.

Illumina Technology: HiSeq 2000 System

Nishikawa Y, Hosokawa M, Maruyama T, Yamagishi K, Mori T, et al. Monodisperse Picoliter Droplets for Low-Bias and Contamination-Free Reactions in Single-Cell Whole Genome Amplification. *PLoS One.* **2015;10:e0138733**

WGA is a critical component of single-cell sequencing pipelines, and MDA is the most common WGA method in single-cell sequencing. Despite its widespread use, MDA typically produces uneven genome coverage due to amplification bias and the formation of DNA chimeras. To overcome this limitation, the authors developed droplet MDA that minimizes these technical artifacts. They used microfluidics to compartmentalize extracted DNA fragments into 67 pL droplets, where the individual fragments were then amplified using MDA. This approach was validated by sequencing the droplet MDA products of *E. coli* cells, with genome recovery improving to 89%, compared to 59% using traditional MDA.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq System

Li N, Wang L, Wang H, Ma M, Wang X, et al. The Performance of Whole Genome Amplification Methods and Next-Generation Sequencing for Pre-Implantation Genetic Diagnosis of Chromosomal Abnormalities. *J Genet Genomics.* 2015;42:151-159

276. Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* 2015;12:519-522

Genome & Transcriptome Sequencing

Genome & transcriptome sequencing (G&T-Seq) is a protocol that can separate and sequence genomic DNA and full-length mRNA from single cells.[276] In this method, single cells are isolated and lysed. RNA is captured using biotinylated oligo(dT) capture primers and separated from DNA using streptavidin-coated magnetic beads. Smart-Seq2 is used to amplify captured RNA on the bead, while MDA is used to amplify DNA. After sequencing, integrating DNA and RNA sequences provides insights into the gene-expression profile of single cells (Table 4).
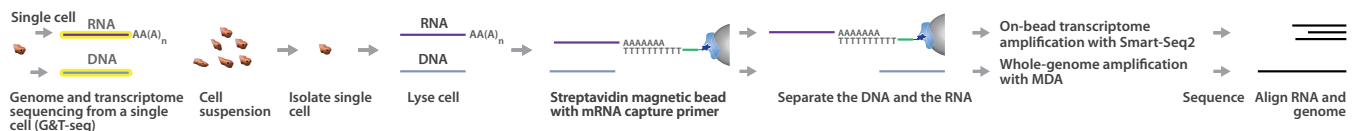
A schematic overview of G&T-Seq.

Table 4. Advantages and Disadvantages of G&T-Seq.

Advantages	Disadvantages
• Compatible with any WGA method. • No 3'-end bias in sequence reads because full-length transcripts are captured. • Because DNA and RNA are physically separated and amplified independently, there is no need to mask coding sequences during analysis.	• Physical separation of DNA and RNA can increase the risk of sample loss or contamination. • Physical separation of DNA and RNA increases handling time.

References

Angermueller C, Clark SJ, Lee HJ, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* **2016;13:229-232**

Multiparameter single-cell sequencing is a powerful tool that has uncovered relationships among genomic, transcriptional, and epigenetic heterogeneity. In this study, the authors developed single-cell methylome & transcriptome sequencing (scM&T-Seq), a multiparameter sequencing method that allows methylome and transcriptome profiling in the same cell. They used the G&T-Seq protocol to purify single-cell DNA that was then subjected to single-cell bisulfite conversion (scBS-Seq).[277] The authors performed scM&T-Seq on 61 mouse ESCs. They found that gene expression levels of many pluripotency factors were negatively associated with DNA methylation. These data demonstrate that epigenetic heterogeneity is an important mechanism of fluctuating pluripotency in ESCs. They also demonstrate that scM&T-Seq can illuminate the poorly understood relationship between transcriptional and DNA-methylation heterogeneity in single cells.

Illumina Technology: Nextera XT Kit, HiSeq 2000 System

Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* **2015;12:519-522**

Single-cell genomic sequencing has provided insights into cellular heterogeneity, as well as cellular lineage and development. Single-cell transcriptomic sequencing has refined our understanding of cell types and states. In this study, the authors developed G&T-Seq, a method that allows for the separation and subsequent sequencing of genomic DNA and full-length mRNA from single cells. It complements the genomic DNA and mRNA sequencing (DR-Seq) method,[278] but it can be used with any WGA method and also provides full-length transcripts from the same cell. The authors performed G&T-Seq-enabled transcriptome analysis by using a modified Smart-Seq2 protocol[279,280] and automated the method on a robotic liquid-handling platform. They used the HiSeq platform to sequence numerous single-cell types, including human cancer cells, reversine-treated mouse embryo blastomeres, and iPSC-derived neurons. Notably, G&T-Seq analysis of aneuploid blastomeres demonstrated that chromosomal gains/losses led to increases/losses in chromosome-wide relative gene expression during a single cell division.

Illumina Technology: Nextera XT Kit, MiSeq, HiSeq 2500 and HiSeq X Ten Systems

277. Smallwood SA, Lee HJ, Angermueller C, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods.* 2014;11:817-820

278. Dey SS, Kester L, Spanjaard B, Bienko M and van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol.* 2015;33:285-289

279. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G and Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods.* 2013;10:1096-1098

280. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S and Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc.* 2014;9:171-181

281. Zong C, Lu S, Chapman AR and Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science.* 2012;338:1622-1626

Multiple Annealing and Looping–Based Amplification Cycles

MALBAC is intended to address some of the shortcomings of MDA.[281] In this method, MALBAC primers randomly anneal to the DNA template. A polymerase with displacement activity at elevated levels amplifies the template, generating semiamplicons. As the amplification and annealing process is repeated, the semiamplicons are amplified into full amplicons that have a 3' end complementary to the 5' end. As a result, full-amplicon ends hybridize to form a looped structure that inhibits further amplification of the looped amplicon, while only the semiamplicons and genomic DNA undergo amplification. Deep sequencing full-amplicon sequences allows for accurate representation of reads, while sequencing depth provides improved alignment for consensus sequences (Table 5).
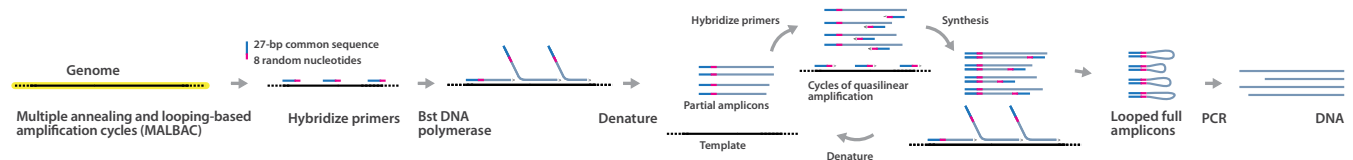
A schematic overview of MALBAC.

Table 5. Advantages and Disadvantages of MALBAC.

Advantages	Disadvantages
• Can sequence large templates. • Can perform single-cell sequencing or sequencing for samples with very limited starting material. • Full-amplicon looping inhibits overrepresentation of templates, reducing PCR bias. • Can amplify GC-rich regions. • Uniform genome coverage. • Lower allele drop-out rate compared to MDA.	• Polymerase is relatively error-prone compared to Phi 29 • Temperature-sensitive protocol. • Genome coverage up to ~90%,[282] but some regions of the genome are consistently underrepresented.[283]

282. Lovett M. The applications of single-cell genomics. *Hum Mol Genet.* 2013;22:R22-26
283. Lasken RS. Single-cell sequencing in its prime. *Nat Biotechnol.* 2013;31:211-212

References

Ning L, Li Z, Wang G, Hu W, Hou Q, et al. Quantitative assessment of single-cell whole genome amplification methods for detecting copy number variation using hippocampal neurons. *Sci Rep.* 2015;5:11415
In this study, the authors compared MDA, MALBAC, and GenomePlex amplification methods in sequencing of individual hippocampal neurons. They amplified genomic DNA from individual hippocampal neurons using 3 different amplification methods, followed by sequencing at shallow depth on a HiSeq 2000 system. Their results showed that single-cell sequencing results from MALBAC and GenomePlex methods were highly reproducible and had high success rates. MALBAC did display significant GC bias, but it was overcome by using bioinformatics tools. Overall, they determined that MALBAC and GenomePlex performed better for detecting CNVs.

Illumina Technology: HiSeq 2000

Yan L, Huang L, Xu L, Huang J, Ma F, et al. Live births after simultaneous avoidance of monogenic diseases and chromosome abnormality by next-generation sequencing with linkage analyses. *Proc Natl Acad Sci U S A.* 2015;112:15964-15969
NGS methods have improved the precision of PGS/PGD. Although the precision has been limited by false-positive and false-negative SNVs, linkage analysis can overcome this challenge. In this study, the authors developed MARSALA, a method that combines NGS using the HiSeq platform with single-cell WGA. The method allows for embryo diagnosis with a single-molecule precision and significantly reduces false-positive and false-negative errors. This is the first integrated NGS-based PGD procedure that simultaneously detects disease-causing mutations and chromosome abnormalities and performs linkage analyses.

Illumina Technology: HiSeq 2500 System

Li N, Wang L, Wang H, Ma M, Wang X, et al. The Performance of Whole Genome Amplification Methods and Next-Generation Sequencing for Pre-Implantation Genetic Diagnosis of Chromosomal Abnormalities. *J Genet Genomics.* 2015;42:151-159

Genomic DNA and mRNA Sequencing

DR-Seq studies the genomic and transcriptomic relationship of single cells via sequencing. Nucleic acid amplification prior to physical separation reduces sample loss and the risk of contamination. DR-Seq involves multiple amplification steps, including the quasilinear amplification technique similar to MALBAC.

First, mRNAs are reverse-transcribed from lysed single cells using poly(dT) primers with Ad-1x adapters, producing single-stranded cDNA (sscDNA). The Ad-1x adapter sequence contains cell-identifying barcodes, 5' Illumina adapters, and a T7 promoter. Next, both gDNA and sscDNA are amplified simultaneously via quasilinear WGA with Ad-2 primers. These primers are similar to MALBAC adapters, containing 8 random nucleotides for random priming followed by a constant 27-nucleotide

tag at the 5' end. Products of this amplification step are split in halves. One half is prepared for genome sequencing, in which gDNA are PCR-amplified and "liberated" of their Ad-2 adapters before DNA library prep and sequencing. The other half is prepared for transcriptome sequencing, whereby second strands are synthesized for the cDNAs and amplified by *in vitro* transcription. The resulting RNA products are produced only from cDNA fragments flanked with Ad-1x and Ad-2, omitting amplification of the gDNA fragments. The RNA library is prepared for sequencing following the Illumina small-RNA protocol. Sequencing gDNA and mRNA from the same cell preserves information between the genome and its expression levels (Table 6).

284. Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* 2015;12:519-522
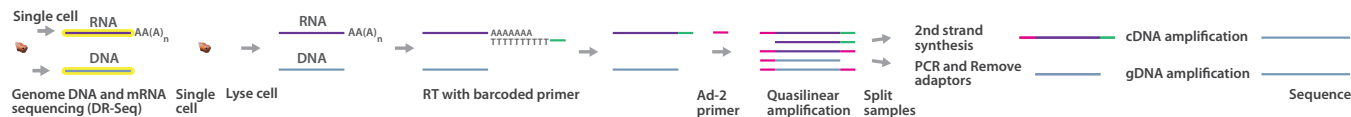
A schematic overview of DR-Seq.

Table 6. Advantages and Disadvantages of DR-Seq.

Advantages	Disadvantages
• Interrogates genomic and transcriptomic behavior from a single cell. • Amplification prior to separation reduces sample loss and contamination. • Length-based identifier used to remove duplicate reads. • Quasilinear amplification reduces PCR bias.	• Manual single-cell isolation prevents high-throughput adaptation. • Quasilinear amplification is temperature-sensitive. • RNA reads are 3'-end–biased.[284]

Reference

Dey SS, Kester L, Spanjaard B, Bienko M and van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol.* **2015;33:285-289**
Single-cell genomics and transcriptomics are promising tools for quantifying genetic and expression variability among individual cells. In this study, the authors describe DR-Seq, a method to quantify the genome and transcriptome of the same cell simultaneously. DR-Seq does not require physical separation of nucleic acids before amplification, which helps to minimize the chances for sample loss or contamination. The authors amplified gDNA and cDNA from mouse ESCs, subsequently divided the nucleic acids for further amplification and library construction, and sequenced both libraries using a HiSeq 2500 system. They demonstrated that genes with high cell-to-cell variability in transcript numbers have low CNVs, and *vice versa*.

Illumina Technology: HiSeq 2500 System

EPIGENOMICS METHODS

Epigenetics—the mechanisms of temporal and spatial control of gene activity, independent of DNA sequence—plays a crucial role in embryogenesis, differentiation, lineage specification, and cancer evolution.[285,286] During embryogenesis, differentiating cells acquire epigenetic alterations[287] that subsequently influence RNA expression and cellular phenotype.[298,289] In differentiated somatic cells, as well as in stem cells, epigenetic markers can be regulated by lifestyle, environmental factors, chemical exposure, stress, and other factors.[290] Epigenetic changes play a role in a number of diseases, including cancer, neurodegenerative diseases, cardiovascular disease, and respiratory disease.[291]

> "Today we can probe the majority of epigenetic dimensions with single-cell resolution."
>
> - Kelsey, Stegle, and Reik 2017

Bulk sequencing of tissues lacks the resolution required to understand how genotypically identical individual cells develop unique phenotypes as a result of unique spatial localization and temporal order. Single-cell epigenomics techniques, including DNA methylation and chromatin immunoprecipitation sequencing (ChIP-Seq), can be combined with RNA expression and SNP data to identify the mechanistic role of epigenetics in gene regulation precisely.[292] Recently, massively parallel sequencing techniques have been developed to analyze epigenomics in thousands of individual cells, allowing us to understand epigenomic heterogeneity at unprecedented resolution.[293,294] Further, the development of single-cell multiparameter methods has enabled simultaneous profiling of epigenomic and transcriptomic changes in individual cells,[295,296] with a recent study even demonstrating simultaneous single-cell profiling of genomic, transcriptomic, and epigenomic (triple-omics) changes within individual cells.[297]

This section highlights some single-cell epigenomics sequencing methods and recent publications demonstrating how Illumina technology is being used in single-cell epigenomics techniques. To learn more about Illumina sequencing methods, visit www.illumina.com/techniques/sequencing.html.

285. Clark SJ, Lee HJ, Smallwood SA, Kelsey G and Reik W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.* 2016;17:72

286. Guo H, Zhu P, Yan L, et al. The DNA methylation landscape of human early embryos. *Nature.* 2014;511:606-610

287. Guo H, Zhu P, Wu X, Li X, Wen L and Tang F. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* 2013;23:2126-2135

288. Xue Z, Huang K, Cai C, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature.* 2013;500:593-597

289. Weaver WM, Tseng P, Kunze A, et al. Advances in high-throughput single-cell microtechnologies. *Curr Opin Biotechnol.* 2014;25:114-123

290. Alegria-Torres JA, Baccarelli A and Bollati V. Epigenetics and lifestyle. *Epigenomics.* 2011;3:267-277

291. Hyun BR, McElwee JL and Soloway PD. Single molecule and single cell epigenomics. *Methods.* 2015;72:41-50

292. Greenleaf WJ. Assaying the epigenome in limited numbers of cells. *Methods.* 2015;72:51-56

293. Cusanovich DA, Daza R, Adey A, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science.* 2015;348:910-914

294. Rotem A, Ram O, Shoresh N, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol.* 2015;33:1165-1172

295. Angermueller C, Clark SJ, Lee HJ, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* 2016;13:229-232

296. Hu Y, Huang K, An Q, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* 2016;17:88

297. Hou Y, Guo H, Cao C, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26:304-319

For Research Use Only. Not for use in diagnostic procedures.

68   Single-cell Research
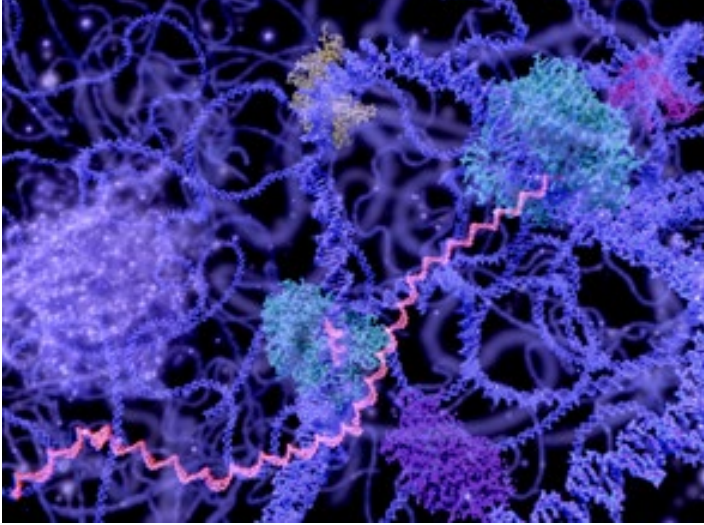
The dynamic composition of chromatin during different stages of the cell cycle, or from one cell type to another, is regulated through multiple epigenetic mechanisms.

Reviews

Kelsey G, Stegle O and Reik W. Single-cell epigenomics: Recording the past and predicting the future. *Science.* 2017;358:69-75

Clark SJ, Lee HJ, Smallwood SA, Kelsey G and Reik W. Single-cell epigenomics: powerful new methods for understanding gene regulation and cell identity. *Genome Biol.* 2016;17:72

Greenleaf WJ. Assaying the epigenome in limited numbers of cells. *Methods.* 2015;72:51-56

Hyun BR, McElwee JL and Soloway PD. Single molecule and single cell epigenomics. *Methods.* 2015;72:41-50

Schwartzman O and Tanay A. Single-cell epigenomics: techniques and emerging applications. *Nat Rev Genet.* 2015;16:716-726

References

Stevens TJ, Lando D, Basu S, et al. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature.* 2017;544:59-64

Understanding cell structure has typically been facilitated through cell biology studies involving microscopy imaging techniques. In the case of the cell nucleus, these cell biology data have been enhanced using chromosome conformation capture epigenomic methods, like 3C and Hi-C. In these epigenomic methods, DNA sequences that are in close physical proximity within the nucleus are digested, ligated, and analyzed by NGS. To more fully resolve the 3D structure of the mammalian genome, the authors combined imaging with improved single-cell Hi-C. Further, they show that integrating the epigenomic data with transcriptomic (RNA-Seq) and ChIP-Seq data provides finer detail regarding the organization of regulated genes. Their data show that structures of loops and topological-associated domains (TADs) are highly variable from cell to cell. Conversely, active enhancer and promoter structure is consistent genome-wide in every cell.

Illumina Technology: Nextera XT DNA Library Preparation Kit, MiSeq and HiSeq Systems

Hou Y, Guo H, Cao C, Li X, Hu B, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* **2016;26:304-319**

To fully understand the mechanisms by which the genome, transcriptome, and DNA methylome interact at the single-cell level, these 3 separate methods ideally should be applied to the same individual cell. In this study, the authors report scTrio-Seq, a method that can analyze genomic CNVs, the DNA methylome, and the transcriptome of an individual mammalian cell simultaneously. They used scTrio-Seq in 25 individual hepatocellular carcinoma primary cells to identify 2 subpopulations of cells. They also found that large-scale CNVs can cause proportional changes in RNA expression in subsets of genes, but the CNVs did not affect DNA methylation in the relevant genomic regions.

Illumina Technology: HiSeq 2000/2500 Systems

Zhang CZ, Spektor A, Cornils H, et al. Chromothripsis from DNA damage in micronuclei. *Nature.* **2015;522:179-184**

Chromothripsis is a new mutational phenomenon in cancer and congenital disorders. In this process, extensive DNA rearrangements and oscillating patterns of DNA copy number are restricted to one or a few chromosomes. The mechanism underlying chromothripsis is not known, but it has been proposed to involve physical isolation of chromosomes in micronuclei. In this study, the authors combined single-cell genome sequencing with live cell imaging to demonstrate that micronucleus formation can lead to a spectrum of genomic rearrangements, including chromothripsis. Specifically, the mechanism for chromothripsis appears to involve the fragmentation and subsequent reassembly of single chromatids within single micronuclei.

Illumina Technology: HiSeq and MiSeq Systems

Ramani V, Deng X, Qiu R, et al. Massively multiplex single-cell Hi-C. *Nat Methods.* 2017;14:263-266

Sanders AD, Falconer E, Hills M, Spierings DCJ and Lansdorp PM. Single-cell template strand sequencing by Strand-seq enables the characterization of individual homologs. *Nat Protoc.* 2017;12:1151-1176

Wu X, Inoue A, Suzuki T and Zhang Y. Simultaneous mapping of active DNA demethylation and sister chromatid exchange in single cells. *Genes Dev.* 2017;31:511-523

Finegersh A and Homanics GE. Chromatin immunoprecipitation and gene expression analysis of neuronal subtypes after fluorescence activated cell sorting. *J Neurosci Methods.* 2016;263:81-88

Milani P, Escalante-Chong R, Shelley BC, Patel-Murray NL, Xin X, et al. Cell freezing protocol suitable for ATAC-Seq on motor neurons derived from human induced pluripotent stem cells. *Sci Rep.* 2016;6:25474

Qu W, Tsukahara T, Nakamura R, et al. Assessing Cell-to-Cell DNA Methylation Variability on Individual Long Reads. *Sci Rep.* 2016;6:21317
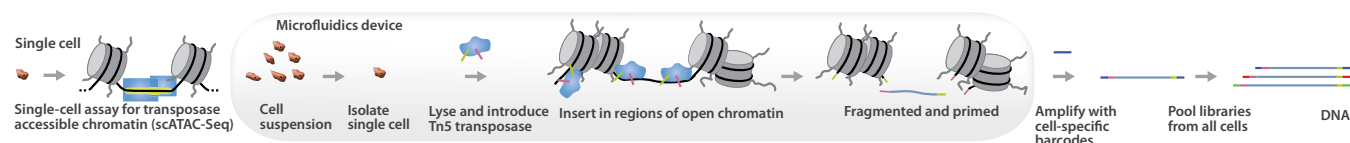
Fortin JP and Hansen KD. Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol.* 2015;16:180

Guo H, Zhu P, Guo F, et al. Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc.* 2015;10:645-659

Jin W, Tang Q, Wan M, et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. *Nature.* 2015;528:142-146

Single-Cell Assay for Transposase-Accessible Chromatin Using Sequencing

The single-cell assay for transposase-accessible chromatin using sequencing (scATAC-Seq) is a protocol for mapping accessible regions in the genome of single cells by combining microfluidics and Tn5 tagmentation. In scATAC-Seq, cell suspensions are loaded into a microfluidics system and sorted individually. Here, cells undergo lysis, and Tn5 transposase tags open chromatin regions with sequencing barcodes. Tagged DNA fragments are purified and amplified with cell-specific barcodes. Libraries from all single cells are then pooled, and deep sequencing provides base-pair resolution of nucleosome-free regions in the genome (Table 7).

298. Pott S and Lieb JD. Single-cell ATAC-seq: strength in numbers. *Genome Biol.* 2015;16:172

299. Buenrostro JD, Wu B, Litzenburger UM, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature.* 2015;523:486-490

A schematic overview of scATAC-Seq.

Table 7. Advantages and Disadvantages of scATAC-Seq.

Advantages	Disadvantages
• Deep sequencing of open chromatin regions in single cells. • High average reads per cell (70,000 reads) compared to combinatorial indexing scATAC-Seq. • Capture of each viable cell individually confirmed through microscopy in the microfluidics device.	• Lower throughput than combinatorial indexing scATAC-Seq (maximum of 96 cells in parallel).[298]

References

Buenrostro JD, Wu B, Litzenburger UM, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature.* 2015;523:486-490
Methods to investigate genome-wide DNA accessibility have revealed substantial variation in regulatory regions across a wide diversity of cells. In order to test whether this heterogeneity exists within individual cells, the authors developed scATAC-Seq. In this method, individual cells are captured and assayed using a microfluidics platform. After PCR amplification and barcoding, the authors sequenced these single-cell libraries using HiSeq and NextSeq systems. They generated DNA accessibility maps from 254 GM12878 lymphoblastoid cells, as well as other cell lines. Their data demonstrate single-cell epigenetic heterogeneity.

Illumina Technology: Nextera DNA Sample Prep Kit, HiSeq and NextSeq Systems

Cusanovich DA, Daza R, Adey A, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science.* 2015;348:910-914
This study used combinatorial indexing to measure chromatin accessibility in thousands of single cells. The authors isolated nuclei and tagged them in bulk with Tn5 transposases, in each of many wells. Next, they pooled these barcoded nuclei, diluted them, and redistributed them to a second set of wells where a second barcode was introduced using PCR. The authors integrated this combinatorial indexing with scATAC-Seq[299] to measure chromatin accessibility for more than 15,000 human and mouse single cells. They sequenced the scATAC-Seq libraries on the MiSeq system, and their data identified relevant differences in chromatin accessibility between cell types.

Illumina Technology: MiSeq and NextSeq Systems

Single-Cell Bisulfite Sequencing/Single-Cell Whole-Genome Bisulfite Sequencing

Single-cell bisulfite sequencing (scBS-Seq) or single-cell whole-genome bisulfite sequencing (scWGBS) are versions of the well-established bisulfite sequencing (BS-Seq) and whole-genome bisulfite sequencing (WGBS) post-bisulfite adapter-tagging (PBAT) protocols, modified to detect methylated cytosines in genomic DNA from single cells. In this method, after single cells are isolated, genomic DNA is treated with sodium bisulfite, which fragments the DNA. The converted DNA then undergoes random priming several times and is PCR-amplified for sequencing. Deep sequencing provides single-nucleotide resolution of methylated cytosines from single cells (Table 8).
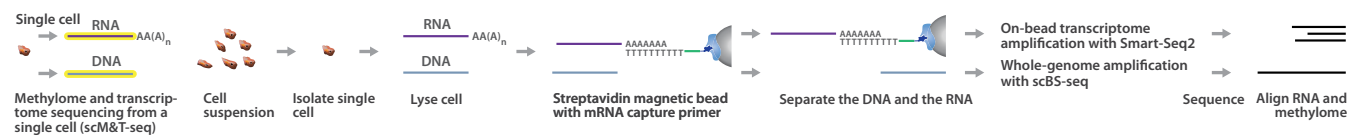
300. Tasic B, Menon V, Nguyen TN, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci.* 2016;19:335-346

301. Zeisel A, Munoz-Manchado AB, Codeluppi S, et al. Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science.* 2015;347:1138-1142

A schematic overview of scBS/scWGBS.

Table 8. Advantages and Disadvantages of scBS/scWGBS.

Pros	Cons
• Covers CpG and non-CpG methylation throughout the genome at single-base resolution. • Covers 5mC in dense, less dense, and repeat regions	• Bisulfite converts unmethylated cytosines to thymidines, reducing sequence complexity, which can make it difficult to create alignments. • SNPs where a cytosine is converted to thymidine will be missed upon bisulfite conversion. • Bisulfite conversion does not distinguish between 5mC and 5hmC.

References

Luo C, Keown CL, Kurihara L, Zhou J, He Y, et al. Single-cell methylomes identify neuronal subtypes and regulatory elements in mammalian cortex. *Science.* 2017;357:600-604
Single-cell RNA-Seq approaches have been useful in defining diverse neuronal subtypes within the brain.[300,301] In this study, the authors developed a technique to characterize the epigenomic diversity within individual neuronal cells. This technique, single-nucleotide methylcytosine sequencing (snmC-seq), takes advantage of the significant amount of 5-methylcytosine found in neurons. The authors isolated 3400 individual neurons from mouse cortex and collected nuclei by FACS. They treated the nuclear DNA treated with bisulfite to convert unmethylated cytosine into uracil, while 5-methyl- and 5-hydroxymethylcytosines remained unaffected. Finally, they pooled, amplified, and sequenced the samples using the HiSeq 4000 System. The authors used snmC-seq to generate 6000 methylomes from the isolated neuronal nuclei, and their data identified 16 mouse neuronal subtypes in mouse brain. They also performed snmC-seq on isolated nuclei from human cortex samples, and they identified 21 human neuronal subtypes. The single-cell methylome technique thus expanded the taxonomy of brain cell diversity. Moreover, it identified regulatory elements associated with neuronal diversity.

Illumina Technology: TruSeq Methylation kit, HiSeq 4000 System

Farlik M, Sheffield NC, Nuzzo A, et al. Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep.* 2015;10:1386-1397
WGBS is currently the most popular method for methylation mapping. In this study, the authors describe a WGBS method modified for single cells (scWGBS). They sorted cells using FACS and bisulfite-converted the DNA directly in lysed cells. Next, they prepared single-strand libraries and sequenced them using the HiSeq 2000/2500 system. They validated the method using more than 250 samples in 3 *in vitro* models of cellular differentiation, including the K562 erythroleukemia-derived cell line, the HL60 cell line, and induced mouse ESCs. In all 3 models, scWGBS detailed characteristic patterns of epigenome remodeling and cell-to-cell heterogeneity.

Illumina Technology: HiSeq 2000/2500 Systems

Single-Cell Methylome & Transcriptome Sequencing

scM&T-Seq allows parallel analysis of both epigenetic and gene expression patterns from single cells using Smart-Seq2 and scBS-Seq. scM&T-Seq is built upon G&T-Seq, but instead of using MDA for DNA sequencing, it uses scBS-Seq to interrogate DNA methylation patterns.

First, single cells are isolated and individually lysed. Then, mRNAs are isolated using streptavidin-coupled mRNA capture primers, physically separating them from DNA strands. Smart-Seq2 is used to generate cDNA libraries from the mRNA, which involves reverse transcription with template switching and tagmentation. DNA libraries are prepared via scBS-Seq, which involves bisulfite conversion of DNA strands to identify methylated cytosines. Both libraries are now ready for sequencing (Table 9).

302. Smallwood SA, Lee HJ, Angermueller C, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods.* 2014;11:817-820

A schematic overview of scM&T-Seq.

Table 9. Advantages and Disadvantages of scM&T-Seq.

Advantages	Disadvantages
• Investigates links between epigenetic and transcriptional heterogeneity in single cells. • Because DNA and RNA are physically separated and amplified independently, there is no need to mask coding sequences during analysis.	• Smart-Seq2 is not strand-specific and applicable to only poly(A)$^+$ RNA. • Does not distinguish between 5mC and 5hmC.

References

Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* 2016;13:229-232
Multiparameter single-cell sequencing is a powerful tool that has uncovered relationships among genomic, transcriptional, and epigenetic heterogeneity. In this study, the authors developed scM&T-Seq, a multiparameter sequencing method that allows methylome and transcriptome profiling in the same cell. They used the G&T-Seq protocol to purify single-cell DNA that was then subjected to single-cell bisulfite conversion (scBS-Seq).[302] Using the HiSeq 2000 system, the authors performed scM&T-Seq on 61 mouse ESCs. They found that gene expression levels of many pluripotency factors were negatively associated with DNA methylation. These data demonstrate that epigenetic heterogeneity is an important mechanism of fluctuating pluripotency in ESCs. They also demonstrate that scM&T-Seq can illuminate the poorly understood relationship between transcriptional and DNA-methylation heterogeneity in single cells.

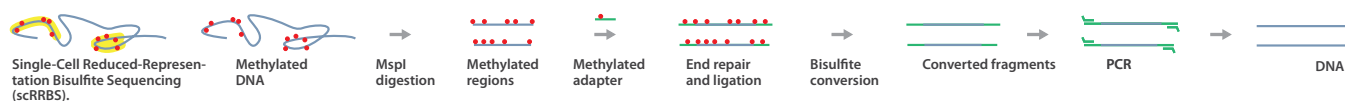Illumina Technology: Nextera XT Kit, HiSeq 2000 System

Hu Y, Huang K, An Q, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* 2016;17:88
In this study, the authors developed a method that simultaneously profiles the methylome and the transcriptome of the same individual cell (scMT-Seq). The method is very similar to scM&T-Seq, except that scMT-Seq uses single-cell reduced-representation bisulfite sequencing (scRRBS) for methylome analysis. The authors validated this method by simultaneously profiling the transcriptome and DNA methylome in individual sensory neurons of the dorsal root ganglion (DRG). Their data identified transcriptome and DNA methylome heterogeneity in DRG neurons. They also found that gene methylation and expression are positively correlated, but only for those genes that contain CpG island promoters.

Illumina Technology: MiSeq and HiSeq 2500 Systems

For Research Use Only. Not for use in diagnostic procedures.

Single-Cell Reduced-Representation Bisulfite Sequencing

scRRBS uses one or multiple restriction enzymes on genomic DNA to produce sequence-specific fragmentation.[303,304] The fragmented genomic DNA is then treated with bisulfite and sequenced. It is the method of choice to study specific regions of interest. It is particularly effective where methylation is high, such as in promoters and repeat regions (Table 10).

303. Guo H, Zhu P, Guo F, et al. Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc.* 2015;10:645-659

304. Guo H, Zhu P, Wu X, Li X, Wen L and Tang F. Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing. *Genome Res.* 2013;23:2126-2135

Single-Cell Reduced-Representation Bisulfite Sequencing (scRRBS). → **Methylated DNA** → **MspI digestion** → **Methylated regions** → **Methylated adapter** → **End repair and ligation** → **Bisulfite conversion** → **Converted fragments** → **PCR** → **DNA**

A schematic overview of scRRBS.

Table 10. Advantages and Disadvantages of scRRBS.

Advantages	Disadvantages
• Provides genome-wide coverage of CpGs in islands at single-base resolution. • Covers areas dense in CpG methylation	• Restriction enzymes cut at specific sites, providing biased sequence selection. • Measures 10%–15% of all CpGs in the genome. • Cannot distinguish between 5mC and 5hmC. • Does not cover non-CpG areas, genome-wide CpGs, and CpGs in areas without the enzyme restriction site.

Reference

Guo H, Zhu P, Guo F, Li X, Wu X, et al. Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc.* 2015;10:645-659

Single-Cell Chromatin Immunoprecipitation Sequencing

scChIP-Seq is a well-established method to map specific protein-binding sites.[305] In this method, DNA-protein complexes are crosslinked *in vivo*. Samples are then fragmented and treated with an exonuclease to trim unbound oligonucleotides. Protein-specific antibodies are used to immunoprecipitate the DNA-protein complex. The DNA is extracted and sequenced, giving high-resolution sequences of the protein-binding sites (Table 11).

305. Solomon MJ, Larsen PL and Varshavsky A. Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. *Cell.* 1988;53:937-947

306. Zentner GE and Henikoff S. Surveying the epigenomic landscape, one base at a time. *Genome Biol.* 2012;13:250

Chromatin immune precipitation (ChIP-Seq) — DNA-protein complex — Crosslink proteins and DNA — Sample fragmentation — Exonuclease digestion — Immunoprecipitation — DNA extraction — DNA
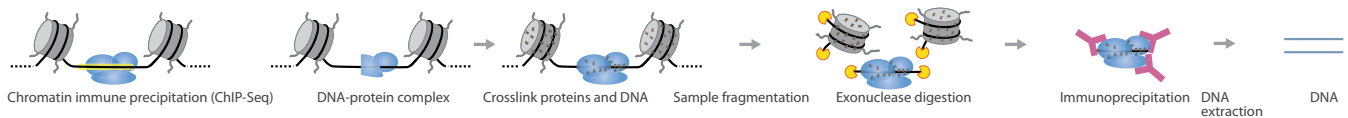
A schematic overview of scChIP-Seq.

Table 11. Advantages and Disadvantages of scChIP-Seq.

Advantages	Disadvantages
• Provides base-pair resolution of protein-binding sites. • Can map specific regulatory factors or proteins. • The use of exonuclease eliminates contamination by unbound DNA.[306]	• Nonspecific antibodies can dilute the pool of DNA-protein complexes of interest. • The target protein must be known and be able to raise an antibody.

Reference

Rotem A, Ram O, Shoresh N, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol.* 2015;33:1165-1172
In this study, the authors combined microfluidics, DNA barcoding, and sequencing to analyze chromatin data at single-cell resolution. They validated the technology by assaying thousands of individual cells, followed by deconvolution of a mixture of ESCs, fibroblasts, and hematopoietic progenitors into chromatin state maps for each cell type. Although the data from each single cell covered only 1000 reads, the ability to assay thousands of individual cells allowed them to identify a spectrum of subpopulations of ESCs, defined by differences in chromatin signatures of pluripotency and differentiation timing. The method revealed aspects of epigenetic heterogeneity not captured by scRNA-Seq alone.
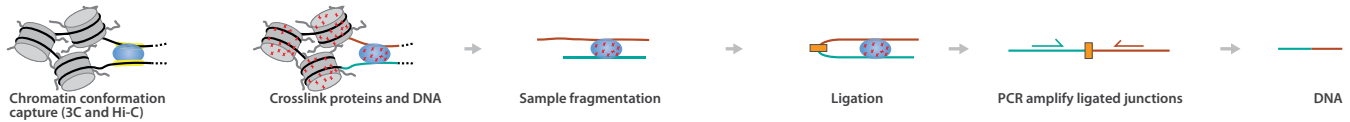
Illumina Technology: HiSeq 2500 System

Chromatin Conformation Capture Sequencing

Chromatin conformation capture sequencing (Hi-C[307] or 3C-Seq[308]) is used to analyze chromatin interactions. In this method, DNA-protein complexes are crosslinked using formaldehyde. The sample is fragmented, and the DNA is ligated and digested. The resulting DNA fragments are PCR-amplified and sequenced. Deep sequencing provides base-pair resolution of ligated fragments (Table 12).

307. Lieberman-Aiden E, van Berkum NL, Williams L, et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 2009;326:289-293

308. Duan Z, Andronescu M, Schutz K, et al. A genome-wide 3C-method for characterizing the three-dimensional architectures of genomes. *Methods.* 2012;58:277-288

Chromatin conformation capture (3C and Hi-C) → Crosslink proteins and DNA → Sample fragmentation → Ligation → PCR amplify ligated junctions → DNA

A schematic overview of Hi-C/3C-Seq.

Table 12. Advantages and Disadvantages of Hi-C/3C-Seq.

Advantages	Disadvantages
• Allows detection of long-range DNA interactions. • High-throughput method.	• Detection may result from random chromosomal collisions. • 3C PCR is difficult, and it requires careful controls and experimental design. • Needs further confirmation of interaction. • Requires large amounts of starting material due to multiple steps.

References

Kind J, Pagie L, de Vries SS, et al. Genome-wide maps of nuclear lamina interactions in single human cells. *Cell.* 2015;163:134-147
During interphase, mammalian chromosomes interact with the nuclear lamina (NL) through structures termed lamina-associated domains (LADs). In this study, the authors developed a modified DNA adenine methyltransferase identification (DamID) method to characterize genome-wide mapping of chromosome-NL interactions in 118 individual KBM7 cells. The data showed that 15% of the genome contacted the NL in most of the individual cells analyzed, and that this contact frequency was locus-specific. Chromosome-LN contact sites that were stable across cells were poor in genes, compared to those sites that were more variable across cells, suggesting that these sites may serve a structural rather than epigenetic role. Hi-C analysis also showed that loci with intrachromosomally coordinated NL contacts were in close proximity in the nuclear space.

Illumina Technology: HiSeq 2000/2500 Systems

Nagano T, Lubling Y, Yaffe E, et al. Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat Protoc.* 2015;10:1986-2003
Hi-C provides pairwise information on genomic regions that are within spatial proximity of each other in the nucleus. In this study, the authors modified single-cell Hi-C with in-nucleus ligation, in order to characterize the thousands of chromatin interactions that occur in individual cells. This modification allows for magnetic-bead capture of labeled, crosslinked ligation junctions and PCR amplification of single-cell Hi-C libraries. The authors validated this approach by performing single-cell Hi-C in individual mouse T helper 1 (T$_H$1) cells. The resulting T$_H$1 interactome maps provided information on nuclear genome organization and chromosome structure.

Illumina Technology: GA$_{IIx}$ System

Droplet-Based Chromatin Immunoprecipitation Sequencing

Single-cell droplet-based chromatin immunoprecipitation sequencing (Drop-ChIP-Seq) analyzes the chromatin states of single cells by utilizing microfluidics, unique molecular barcodes, and NGS.[309] First, single cells are isolated into droplets containing lysis buffer and MNase, and then fused with another droplet carrying distinct oligonucleotides. These oligonucleotides hold the sequences for cell-specific barcodes, sequencing adapter, and restriction sites. DNA ligase is also fused with the droplet to complete the tagging process. Next, carrier chromatins are introduced into the pooled droplets, followed by standard ChIP-Seq procedures (Table 13).

309. Rotem A, Ram O, Shoresh N, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol.* 2015;33:1165-1172
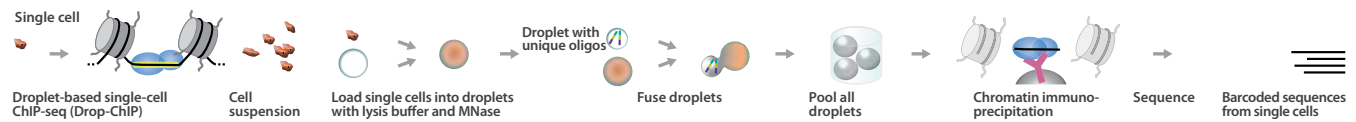
A schematic overview of Drop-ChIP-Seq.

Table 13. Advantages and Disadvantages of Drop-ChIP-Seq.

Advantages	Disadvantages
• Analyzes chromatin states from single cells in a highly parallel manner. • Unique molecular barcoding reduces the risk posed by nonspecific antibodies.	• Requires a large number of sample cells.

Reference

Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol.* 2015;33:1165-1172
In this study, the authors combined microfluidics, DNA barcoding, and sequencing to analyze chromatin data at single-cell resolution. They validated the technology by assaying thousands of individual cells, followed by deconvolution of a mixture of ESCs, fibroblasts, and hematopoietic progenitors into chromatin state maps for each cell type. Although the data from each single cell covered only 1000 reads, the ability to assay thousands of individual cells allowed them to identify a spectrum of subpopulations of ESCs, defined by differences in chromatin signatures of pluripotency and differentiation timing. The method revealed aspects of epigenetic heterogeneity not captured by scRNA-Seq alone.

**Illumina Technology: HiSeq 2500 System**

RNA METHODS

Low-level RNA detection refers to both detection of rare RNA molecules in a cell-free environment (such as circulating tumor RNA) and the expression patterns of single cells. Tissues consist of a multitude of different cell types, each with a distinctly different set of functions. Even within a single cell type, the transcriptomes are highly dynamic and reflect temporal, spatial, and cell cycle–dependent changes. Cell harvesting, handling, and technical issues with sensitivity and bias during amplification add additional levels of complexity. To resolve this multitiered complexity would require analyzing many thousands of cells. The use of unique barcodes has greatly increased the number of samples that can be multiplexed and pooled at little to no decrease in reads associated with each sample. Recent improvements in cell capture and sample preparation will provide more information, faster, and at lower cost.[310,311] This development promises to expand our understanding of cell function fundamentally, with significant implications for research and human health.[312]

> "Single-cell RNA-seq has become instrumental for interrogating cell types, dynamic states, and functional processes in complex tissues." - Habib et al. 2017

Recently, massively parallel sequencing techniques have been developed to analyze gene expression levels in thousands of individual cells, allowing us to understand transcriptional heterogeneity at unprecedented resolution.[313,314] Further, the development of single-cell multiparameter methods has enabled simultaneous profiling of transcriptomic and epigenomic changes in individual cells. A recent study even demonstrated simultaneous single-cell profiling of genomic, transcriptomic, and epigenomic (triple-omics) changes within individual cells.[315]

This section highlights some scRNA-Seq methods and recent publications demonstrating how Illumina technology is being used in scRNA-Seq. To learn more about Illumina sequencing methods, visit www.illumina.com/techniques/sequencing.html.
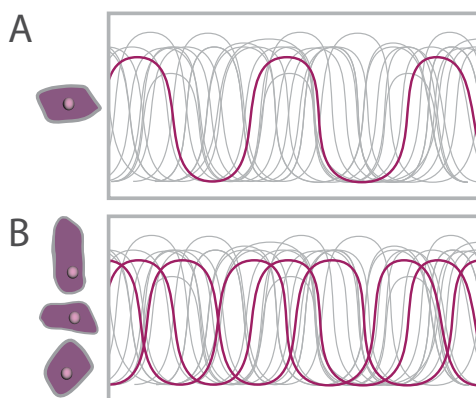
Single-cell transcriptomics approaches can characterize gene expression in individual cells of a tissue or organ.

310. Saliba AE, Westermann AJ, Gorski SA and Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.* 2014;42:8845-8860

311. Shapiro E, Biezuner T and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

312. Spaethling JM and Eberwine JH. Single-cell transcriptomics for drug target discovery. *Curr Opin Pharmacol.* 2013;13:786-790

313. Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* 2015;161:1202-1214

314. Rotem A, Ram O, Shoresh N, et al. High-Throughput Single-Cell Labeling (Hi-SCL) for RNA-Seq Using Drop-Based Microfluidics. *PLoS One.* 2015;10:e0116328

315. Hou Y, Guo H, Cao C, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* 2016;26:304-319

Reviews

Haque A, Engel J, Teichmann SA and Lonnberg T. A practical guide to single-cell RNA-sequencing for bio-medical research and clinical applications. *Genome Med.* 2017;9:75

Liu S and Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges [version 1; referees: 2 approved]. F1000Research. 2016;5(F1000 Faculty Review):182

Faridani OR and Sandberg R. Putting cells in their place. *Nat Biotechnol.* 2015;33:490-491

Kanter I and Kalisky T. Single cell transcriptomics: methods and applications. *Front Oncol.* 2015;5:53

Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC and Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell.* 2015;58:610-620

Trapnell C. Defining cell types and states with single-cell genomics. *Genome Res.* 2015;25:1491-1498

Wang Y and Navin NE. Advances and applications of single-cell sequencing technologies. *Mol Cell.* 2015;58:598-609

316. Kurata M, Yamamoto K, Moriarity BS, Kitagawa M and Largaespada DA. CRISPR/Cas9 library screening for drug target discovery. *J Hum Genet.* 2018;63:179-186

References

Datlinger P, Rendeiro AF, Schmidl C, et al. Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods.* **2017;14:297-301**
Pooled CRISPR screening has become a powerful tool in functional genomics, and it has been used to identify novel biological mechanisms, drug targets, and biomarkers.[316] In this approach, cells are transduced with a library of guide RNA (gRNA) vectors, and gRNA distribution is measured before and after selective challenge. These screening approaches are well suited for relatively simple biological outputs like cell survival, cell prolif-eration, or single gene expression. In this study, the authors developed CRISPR droplet screening (CROP-seq) to enable more complex CRISPR-based gene expression screens. They transduced a gRNA plasmid library into Jurkat cells to express 1 detectable gRNA per cell, then challenged the cells with anti-CD3 and anti-CD28 antibodies to activate T cells. They performed Drop-seq on the Jurkat cells to identify single-cell transcrip-tomes associated with individual gRNAs and identified genes known to be important for T cell activation. This approach links CRISPR screening with transcriptome responses in thousands of individual cells and should enable CRISPR screens with single-cell transcriptome resolution.

Illumina Technology: BaseSpace™, HiSeq 2500 System

Habib N, Avraham-Davidi I, Basu A, et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods.* **2017;14:955-958**
Single-cell RNA-Seq has become an indispensable technique for characterizing cell types, states, and function in complex tissues. Sequencing individual nuclei has similarly proved useful in cell-type discrimination, particu-larly for tissues where single cells are difficult to dissociate. Current approaches for single-nucleus sequenc-ing rely on the use of FACS or C1 microfluidics to sort individual nuclei into multiwell plates for downstream analysis. In this study, the authors developed DroNc-seq, a method that combines single-nucleus sequencing with the microfluidics-based Drop-seq method. The authors modified Drop-seq to incorporate single nuclei within microfluidic droplets, as opposed to single individual cells. They performed massively parallel RNA-Seq for approximate 39,000 individual nuclei from mouse and human brain, at low cost and with similar sensitivity to Drop-seq.

Illumina Technology: NextSeq 500 System

Hou Y, Guo H, Cao C, Li X, Hu B, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res.* **2016;26:304-319**
To fully understand the mechanisms by which the genome, transcriptome, and DNA methylome interact at the single-cell level, these 3 separate methods ideally should be applied to the same individual cell. In this study, the authors report scTrio-Seq, a method that can analyze genomic CNVs, the DNA methylome, and the tran-scriptome of an individual mammalian cell simultaneously. They used scTrio-Seq in 25 individual hepatocellular carcinoma primary cells to identify 2 subpopulations of cells. They also found that large-scale CNVs can cause proportional changes in RNA expression in subsets of genes, but the CNVs did not affect DNA methylation in the relevant genomic regions.

Illumina Technology: HiSeq 2000/2500 Systems

For Research Use Only. Not for use in diagnostic procedures.

Jaitin DA, Weiner A, Yofe I, et al. Dissecting Immune Circuits by Linking CRISPR-Pooled Screens with Single-Cell RNA-Seq. *Cell.* **2016;167:1883-1896 e1815**
In this study, the authors developed CRISP-seq, an NGS method that integrates massively parallel single-cell RNA-Seq with pooled CRISPR screens. They used this technique to characterize immune development and signaling pathways in thousands of individual mouse myeloid cells. By profiling the genomic perturbations and transcriptome in individual cells, they identified multiple interacting factors that act in concert to regulate immune signaling pathways. In particular, they identified specific sets of genes whose expression was regulated by Stat1 and Rela in myeloid cells.

Illumina Technology: MiSeq and NextSeq 500 Systems

Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Ferrante TC, et al. Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat Protoc.* **2015;10:442-458**
scRNA-Seq can profile gene expression over the entire cell transcriptome, but cell isolation typically results in loss of spatial context. *In situ* hybridization is an excellent technique for identifying the location of gene expression, but it is restricted to a fixed number of genes. In this study, the authors present a protocol for *in situ* profiling of gene expression in cells and tissues. In this approach, RNA is converted into crosslinked cDNA amplicons and sequenced manually on a confocal microscope. The approach has the added benefit of enriching for context-specific transcripts over housekeeping/structural genes, while preserving the tissue architecture for transcript localization.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, MiSeq System

Padovan-Merhar O, Nair GP, Biaesch AG, et al. Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. *Mol Cell.* **2015;58:339-352**
Individual mammalian cells display a wide diversity in cellular size and volume, yet they maintain the same absolute DNA content. Therefore, it is likely that compensatory mechanisms exist to maintain a constant concentration of gene expression products, despite differences in DNA concentration. In this study, the authors used single-molecule counting and single-cell image analysis to demonstrate that individual human primary foreskin fibroblasts globally control transcription to compensate for variability in the ratio of DNA to cellular content. They performed scRNA-Seq using the NextSeq 500 system and found that ubiquitously expressed "housekeeping" genes exhibited lower levels of expression noise than other genes.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, NextSeq 500 System

Lake BB, Codeluppi S, Yung YC, et al. A comparative strategy for single-nucleus and single-cell transcriptomes confirms accuracy in predicted cell-type expression from nuclear RNA. *Sci Rep.* 2017;7:6031

Sheng K, Cao W, Niu Y, Deng Q and Zong C. Effective detection of variation in single-cell transcriptomes using MATQ-seq. *Nat Methods.* 2017;14:267-270

Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods.* 2017;14:865-868

Xie S, Duan J, Li B, Zhou P and Hon GC. Multiplexed Engineering and Analysis of Combinatorial Enhancer Activity in Single Cells. *Mol Cell.* 2017;66:285-299 e285

Dixit A, Parnas O, Li B, Chen J, Fulco CP, et al. Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell.* 2016;167:1853-1866 e1817

Habib N, Li Y, Heidenreich M, et al. Div-Seq: A single nucleus RNA-Seq method reveals dynamics of rare adult newborn neurons in the CNS. *Science.* 2016;353:925-928

Krishnaswami SR, Grindberg RV, Novotny M, Venepally P, Lacar B, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524

Mora-Castilla S, To C, Vaezeslami S, Morey R, Srinivasan S, et al. Miniaturization Technologies for Efficient Single-Cell Library Preparation for Next-Generation Sequencing. *J Lab Autom.* 2016;21:557-567

Achim K, Pettit JB, Saraiva LR, Gavriouchkina D, Larsson T, et al. High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol.* 2015;33:503-509

Dueck H, Khaladkar M, Kim TK, et al. Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. *Genome Biol.* 2015;16:122

Kim JK, Kolodziejczyk AA, Illicic T, Teichmann SA and Marioni JC. Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun.* 2015;6:8687Fan

Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods.* 2015;85:54-61

Designed Primer–Based RNA Sequencing

Designed primer–based RNA sequencing (DP-Seq) is a method that amplifies mRNA from limited starting material, as low as 50 pg.[317] In this method, a specific set of heptamer primers is designed. Enriched poly(A)-selected mRNA undergoes first-strand cDNA synthesis. Designed primers are then hybridized to first-strand cDNA, followed by second-strand synthesis and PCR. Deep sequencing of amplified DNA allows for accurate detection of specific mRNA expression at the single-cell level (Table 14).

317. Bhargava V., Ko P., Willems E., Mercola M. and Subramaniam S. (2013) Quantitative transcriptomics using designed primer-based amplification. Sci Rep 3: 1740

318. Bhargava V, Head SR, Ordoukhanian P, Mercola M and Subramaniam S. Technical variations in low-input RNA-seq methodologies. Sci Rep. 2014;4:3678
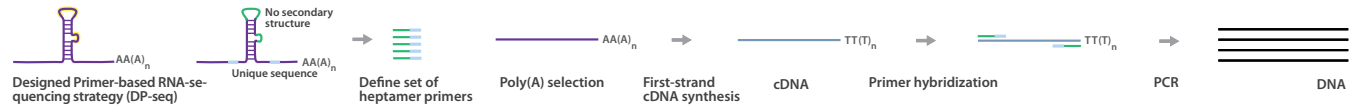
A schematic overview of DP-Seq.

Table 14. Advantages and Disadvantages of DP-Seq.

Advantages	Disadvantages
• Uses as little as 50 pg of starting material. • Low transcript-length bias.	• The sequences of the target areas must be known to design the heptamers. • Exponential amplification during PCR can lead to primer-dimers and spurious PCR products.[318] • Some read-length bias.

Reference

Bhargava V, Head SR, Ordoukhanian P, Mercola M and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep.* 2014;4:3678

Single-Cell Universal Poly(A)-Independent RNA Sequencing

Single-cell universal poly(A)-independent RNA sequencing (SUPeR-Seq) sequences non-poly(A) and poly(A)$^+$ RNAs from single cells. It is designed particularly for mapping circular RNA (circRNA) species. RNA samples from lysed single cells are annealed to random primers with universal anchor sequences (AnchorX-T15N6) and reverse-transcribed to generate the first strand of cDNA. Unreacted primers are digested to avoid primer dimers, prior to the addition of a poly(A) tract to the 3' end of the cDNA. This is done by introducing dATPs and ddATPs in a 100:1 ratio, respectively. A second set of random primers, also with a universal anchor sequence (AnchorY-T24) anneals to the newly synthesized poly(A) tract. A second cDNA strand is generated by reverse transcription, and the cDNA molecules are purified by gel electrophoresis. The purified cDNA molecules are PCR-amplified using 5'-amine-terminated primers and prepared for sequencing by the TruSeq DNA library preparation protocol. After sequencing the cDNA library, circRNAs are identified from the dataset by finding 2 exonic reads that are distal in the reference genome but adjacent to each other, with 1 inverted over the other in the dataset. The inversion of 1 adjacent exon signifies the circularization of the RNA (Table 15).

A schematic overview of SUPeR-Seq.

Table 15. Advantages and Disadvantages of SUPeR-Seq.

Advantages	Disadvantages
• Identifies circular RNA from single cells. • Avoids 3' bias by using random primers with anchor sequences. • Able to identify novel circRNAs due to random primers.	• Relies on dataset analysis to identify circRNAs.

Reference

Fan X, Zhang X, Wu X, et al. Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. *Genome Biol.* 2015;16:148
Although numerous scRNA-Seq methods have been developed, all of them specifically detect polyadenylated RNAs. A substantial amount of RNA expressed in mammalian cells lacks a poly(A) tail. In this study, the authors describe SUPeR-Seq, a poly(A)-independent method for scRNA-Seq. By performing SUPeR-Seq on mouse preimplantation embryos, they discovered 2891 circRNAs and 913 novel linear transcripts. This discovery allowed them to analyze the abundance of circRNAs in mammalian embryonic development and to identify sequence features of circRNAs.

Illumina Technology: TruSeq DNA Sample Preparation Kit, TruSeq RNA Sample Preparation Kit, HiSeq 2000/2500 Systems

For Research Use Only. Not for use in diagnostic procedures.

82 Single-cell Research

Quartz-Seq

The Quartz-Seq method optimizes whole-transcript amplification (WTA) of single cells.[319] In this method, a reverse-transcription (RT) primer with a T7 promoter and PCR target is first added to extracted mRNA. Reverse transcription synthesizes first-strand cDNA, after which the RT primer is digested by exonuclease I. A poly(A) tail is then added to the 3' ends of first-strand cDNA, along with a dT primer containing a PCR target. After second-strand generation, a blocking primer is added to ensure PCR enrichment in sufficient quantity for sequencing. Deep sequencing allows for accurate, high-resolution representation of the whole transcriptome of a single cell (Table 16).

319. Sasagawa Y, Nikaido I, Hayashi T, et al. Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. *Genome Biol.* 2013;14:R31
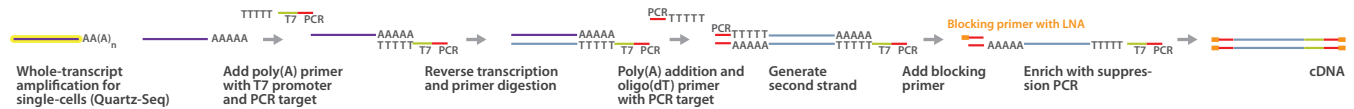
A schematic overview of Quartz-Seq.

Table 16. Advantages and Disadvantages of Quartz-Seq.

Advantages	Disadvantages
• Single-tube reaction suitable for automation. • Digestion of RT primers by exonuclease I eliminates amplification of byproducts. • Short fragments and byproducts are suppressed during enrichment.	• PCR biases can underrepresent GC-rich templates. • Amplification errors caused by polymerases will be represented and sequenced incorrectly. • Targets smaller than 500 bp are preferentially amplified by polymerases during PCR.

Reference

Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, et al. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods.* 2015;85:54-61
RNA-Seq and transcriptional profiling of single cells have expanded our understanding of cellular heterogeneity at levels not achievable using bulk sequencing. Importantly, the cell cycle can be a major driver of transcriptional heterogeneity in scRNA-Seq. In this study, the authors analyzed 6 supervised computational methods to predict G1, S, or G2M phase using ESC transcriptome data. These data were generated, using the Quartz-Seq method, from libraries constructed with the TruSeq Stranded RNA Sample Preparation Kit. By comparing the performance of each algorithm on various scRNA-Seq datasets from various organisms, the authors conclude that a principal component analysis–based approach provides the best results.

Illumina Technology: TruSeq Stranded RNA Sample Preparation Kit

Smart-Seq

Smart-Seq was developed as a single-cell sequencing protocol with improved read coverage across transcripts.[320] Complete coverage across the genome allows the detection of alternative transcript isoforms and SNPs. In this protocol, cells are lysed, and the RNA is hybridized to an oligo(dT)-containing primer. The first strand is then created with the addition of a few untemplated C nucleotides. This poly(C) overhang is added exclusively to full-length transcripts. An oligonucleotide primer is then hybridized to the poly(C) overhang and used to synthesize the second strand. Full-length cDNAs are PCR-amplified to obtain nanogram amounts of DNA. The PCR products are purified for sequencing (Table 17).

320. Ramskold D, Luo S, Wang YC, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol.* 2012;30:777-782

321. Shapiro E, Biezuner T and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

322. Bhargava V, Head SR, Ordoukhanian P, Mercola M and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep.* 2014;4:3678

323. Tang DT, Plessy C, Salimullah M, et al. Suppression of artifacts and barcode bias in high-throughput transcriptome analyses utilizing template switching. *Nucleic Acids Res.* 2013;41:e44

A schematic overview of Smart-Seq.

Table 17. Advantages and Disadvantages of Smart-Seq.

Advantages	Disadvantages
• Uses as little as 50 pg of starting material. • Can be used with unknown mRNA sequences. • Provides improved coverage across transcripts. • Results in high levels of mappable reads.	• Not strand-specific. • No early multiplexing.[321] • Transcript length bias with inefficient transcription of reads over 4 kb.[322] • Preferential amplification of high-abundance transcripts. • The purification step may lead to loss of material. • Could be subject to strand-invasion bias.[323]

Reference

Bhargava V, Head SR, Ordoukhanian P, Mercola M and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep.* 2014;4:3678

Smart-Seq2

Smart-Seq2 incorporates several improvements over the original Smart-Seq protocol.[324,325] The new protocol includes a locked nucleic acid (LNA), an increased MgCl2 concentration, betaine, and elimination of the purification step to improve the yield significantly. In this protocol, single cells are lysed in a buffer that contains free dNTPs and tailed oligo(dT) oligonucleotides with a universal 5′ anchor sequence. Reverse transcription is performed, which also adds 2–5 untemplated nucleotides to the cDNA 3′ end. A template-switching oligo (TSO) is added, which carries 2 riboguanosines and a modified guanosine to produce an LNA as the last base at the 3′ end. After the first-strand reaction, the cDNA is amplified using a limited number of cycles. Tagmentation is then used to construct sequencing libraries quickly and efficiently from the amplified cDNA.
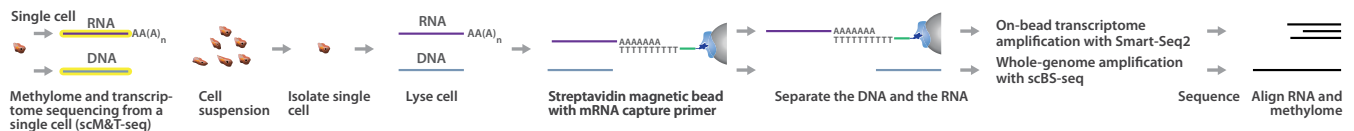
324. Picelli S., Bjorklund A. K., Faridani O. R., Sagasser S., Winberg G., et al. (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10: 1096-1098

325. Picelli S., Faridani O. R., Björklund Å. K., Winberg G., Sagasser S., et al. (2014) Full-length RNA-seq from single cells using Smart-seq2. Nat. Protocols 9: 171-181
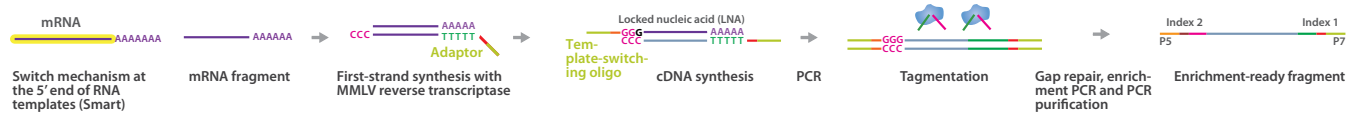
A schematic overview of Smart-Seq2.

Table 18. Advantages and Disadvantages of Smart-Seq2.

Advantages	Disadvantages
• Uses as little as 50 pg of starting material. • Can be used with unknown mRNA sequences. • Provides improved coverage across transcripts. • Results in high levels of mappable reads.	• Not strand-specific. • No early multiplexing. • Only suitable for poly(A)$^+$ RNA.

References

Krishnaswami SR, Grindberg RV, Novotny M, Venepally P, Lacar B, et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc.* 2016;11:499-524
To perform RNA-Seq of single cells, protease treatment has been used to dissociate individual cells from tissues. In this study, the authors showed that this protease-digestion approach altered the transcriptome of individual neurons. To overcome this challenge, they isolated nuclei from postmortem human brain homogenates and sorted them by FACS. They also used Smart-Seq2 to perform cDNA synthesis from nuclear mRNAs and the MiSeq system for sequencing of Nextera XT barcoded libraries. This approach is amenable to any tissue in which single-cell dissociation requires harsh treatment.

Illumina Technology: Nextera XT DNA Library Preparation Kit, MiSeq System

Ziegenhain C, Vieth B, Parekh S, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell.* 2017;65:631-643.e634

Single-Cell Methylome & Transcriptome Sequencing

scM&T-Seq allows parallel analysis of both epigenetic and gene expression patterns from single cells using Smart-Seq2 and scBS-Seq. scM&T-Seq is built upon G&T-Seq, but instead of using MDA for DNA sequencing, it uses scBS-Seq to interrogate DNA methylation patterns.

First, single cells are isolated and individually lysed. Then, mRNAs are isolated using streptavidin-coupled mRNA capture primers, physically separating them from DNA strands. Smart-Seq2 is used to generate cDNA libraries from the mRNA, which involves reverse transcription with template switching and tagmentation. DNA libraries are prepared via scBS-Seq, which involves bisulfite conversion of DNA strands to identify methylated cytosines. Both libraries are now ready for sequencing (Table 19).

326. Smallwood SA, Lee HJ, Angermueller C, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods.* 2014;11:817-820

A schematic overview of scM&T-Seq.

Table 19. Advantages and Disadvantages of scM&T-Seq.

Advantages	Disadvantages
• Investigates links between epigenetic and transcriptional heterogeneity in single cells. • Because DNA and RNA are physically separated and amplified independently, there is no need to mask coding sequences during analysis.	• Smart-Seq2 is not strand-specific and applicable to only poly(A)$^+$ RNA. • Does not distinguish between 5mC and 5hmC.

References

Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* 2016;13:229-232
Multiparameter single-cell sequencing is a powerful tool that has uncovered relationships among genomic, transcriptional, and epigenetic heterogeneity. In this study, the authors developed scM&T-Seq, a multiparameter sequencing method that allows methylome and transcriptome profiling in the same cell. They used the G&T-Seq protocol to purify single-cell DNA that was then subjected to scBS-Seq.[326] Using the HiSeq 2000 system, the authors performed scM&T-Seq on 61 mouse ESCs. They found that gene expression levels of many pluripotency factors were negatively associated with DNA methylation. These data demonstrate that epigenetic heterogeneity is an important mechanism of fluctuating pluripotency in ESCs. They also demonstrate that scM&T-Seq can illuminate the poorly understood relationship between transcriptional and DNA-methylation heterogeneity in single cells.

Illumina Technology: Nextera XT Kit, HiSeq 2000 System

Hu Y, Huang K, An Q, Du G, Hu G, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* 2016;17:88
In this study, the authors developed a method that simultaneously profiles the methylome and the transcriptome of the same individual cell (scMT-Seq). The method is very similar to scM&T-Seq, except that scMT-Seq uses scRRBS for methylome analysis. The authors validated this method by simultaneously profiling the transcriptome and DNA methylome in individual sensory neurons of the DRG. Their data identified transcriptome and DNA methylome heterogeneity in DRG neurons. They also found that gene methylation and expression are positively correlated, but only for those genes that contain CpG island promoters.

Illumina Technology: MiSeq and HiSeq 2500 Systems

Genome & Transcriptome Sequencing

G&T-Seq is a protocol that can separate and sequence genomic DNA and full-length mRNA from single cells.[327] In this method, single cells are isolated and lysed. RNA is captured using biotinylated oligo(dT) capture primers and separated from DNA using streptavidin-coated magnetic beads. Smart-Seq2 is used to amplify captured RNA on the bead, while MDA is used to amplify DNA. After sequencing, integrating DNA and RNA sequences provides insights into the gene-expression profile of single cells.
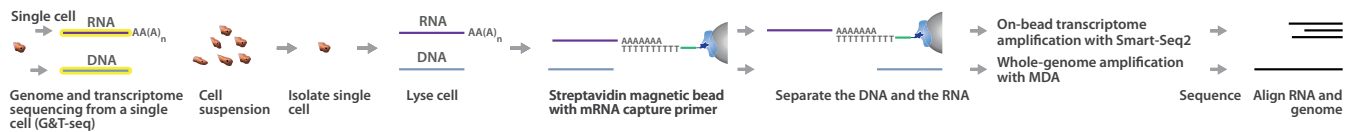
A schematic overview of G&T-Seq.

Table 20. Advantages and Disadvantages of G&T-Seq.

Advantages	Disadvantages
• Compatible with any WGA method. • No 3'-end bias in sequence reads because full-length transcripts are captured. • Because DNA and RNA are physically separated and amplified independently, there is no need to mask coding sequences during analysis.	• Physical separation of DNA and RNA can increase the risk of sample loss or contamination. • Physical separation of DNA and RNA increases handling time.

References

Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods.* 2016;13:229-232

Multiparameter single-cell sequencing is a powerful tool that has uncovered relationships among genomic, transcriptional, and epigenetic heterogeneity. In this study, the authors developed scM&T-Seq, a multiparameter sequencing method that allows methylome and transcriptome profiling in the same cell. They used the G&T-Seq protocol to purify single-cell DNA that was then subjected to scBS-Seq.[328] Using the HiSeq 2000 system, the authors performed scM&T-Seq on 61 mouse ESCs. They found that gene expression levels of many pluripotency factors were negatively associated with DNA methylation. These data demonstrate that epigenetic heterogeneity is an important mechanism of fluctuating pluripotency in ESCs. They also demonstrate that scM&T-Seq can illuminate the poorly understood relationship between transcriptional and DNA-methylation heterogeneity in single cells.

Illumina Technology: Nextera XT Kit, HiSeq 2000 System

Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* 2015;12:519-522

Single-cell genomic sequencing has provided insights into cellular heterogeneity, as well as cellular lineage and development. Single-cell transcriptomic sequencing has refined our understanding of cell types and states. In this study, the authors developed G&T-Seq, a method that allows for the separation and subsequent sequencing of genomic DNA and full-length mRNA from single cells. It complements the DR-Seq method,[329] but it can be used with any WGA method and also provides full-length transcripts from the same cell. The authors performed G&T-Seq-enabled transcriptome analysis by using a modified Smart-Seq2 protocol[330,331] and automated the method on a robotic liquid-handling platform. They used the HiSeq platform to sequence numerous single-cell types, including human cancer cells, reversine-treated mouse embryo blastomeres, and iPSC-derived neurons. Notably, G&T-Seq analysis of aneuploid blastomeres demonstrated that chromosomal gains/losses led to increases/losses in chromosome-wide relative gene expression during a single cell division.
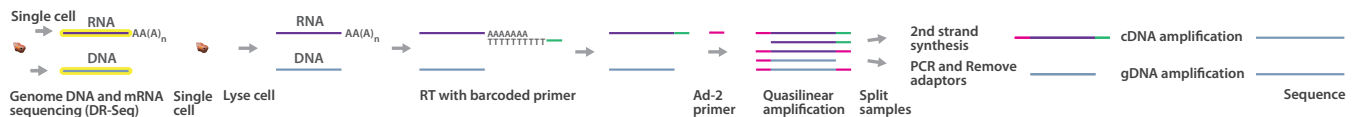
Illumina Technology: Nextera XT Kit, MiSeq, HiSeq 2500 and HiSeq X Ten Systems

327. Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* 2015;12:519-522

328. Smallwood SA, Lee HJ, Angermueller C, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods.* 2014;11:817-820

329. Dey SS, Kester L, Spanjaard B, Bienko M and van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol.* 2015;33:285-289

330. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G and Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods.* 2013;10:1096-1098

331. Picelli S, Faridani OR, Bjorklund AK, Winberg G, Sagasser S and Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc.* 2014;9:171-181

Genomic DNA and mRNA Sequencing

DR-Seq studies the genomic and transcriptomic relationship of single cells via sequencing. Nucleic acid amplification prior to physical separation reduces sample loss and the risk of contamination. DR-Seq involves multiple amplification steps, including the quasilinear amplification technique similar to MALBAC.

First, mRNAs are reverse-transcribed from lysed single cells using poly(dT) primers with Ad-1x adapters, producing sscDNA. The Ad-1x adapter sequence contains cell-identifying barcodes, 5' Illumina adapters, and a T7 promoter. Next, both gDNA and sscDNA are amplified simultaneously via quasilinear WGA with Ad-2 primers. These primers are similar to MALBAC adapters, containing 8 random nucleotides for random priming followed by a constant 27-nucleotide tag at the 5' end. Products of this amplification step are split in halves. One half is prepared for genome sequencing, in which gDNA are PCR-amplified and "liberated" of their Ad-2 adapters before DNA library prep and sequencing. The other half is prepared for transcriptome sequencing, whereby second strands are synthesized for the cDNAs and amplified by *in vitro* transcription. The resulting RNA products are produced only from cDNA fragments flanked with Ad-1x and Ad-2, omitting amplification of the gDNA fragments. The RNA library is prepared for sequencing following the Illumina small-RNA protocol. Sequencing gDNA and mRNA from the same cell preserves information between the genome and its expression levels (Table 21).

332. Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods.* 2015;12:519-522

A schematic overview of DR-Seq.

Table 21. Advantages and Disadvantages of DR-Seq.

Advantages	Disadvantages
• Interrogates genomic and transcriptomic behavior from a single cell. • Amplification prior to separation reduces sample loss and contamination. • Length-based identifier used to remove duplicate reads. • Quasilinear amplification reduces PCR bias.	• Manual single-cell isolation prevents high-throughput adaptation. • Quasilinear amplification is temperature-sensitive. • RNA reads are 3'-end–biased.[332]

Reference

Dey SS, Kester L, Spanjaard B, Bienko M and van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol.* 2015;33:285-289
Single-cell genomics and transcriptomics are promising tools for quantifying genetic and expression variability among individual cells. In this study, the authors describe DR-Seq, a method to quantify the genome and transcriptome of the same cell simultaneously. DR-Seq does not require physical separation of nucleic acids before amplification, which helps to minimize the chances for sample loss or contamination. The authors amplified gDNA and cDNA from mouse ESCs, subsequently divided the nucleic acids for further amplification and library construction, and sequenced both libraries using a HiSeq 2500 system. They demonstrated that genes with high cell-to-cell variability in transcript numbers have low CNVs, and vice versa.

Illumina Technology: HiSeq 2500 System

T Cell–Receptor Chain Pairing

Functional TCRs are heterodimeric proteins composed of unique combinations of α and β chains. For an accurate functional analysis, both subunits must be sequenced together to avoid disrupting the α- and β-chain pairing during the cell lysis step.[333]
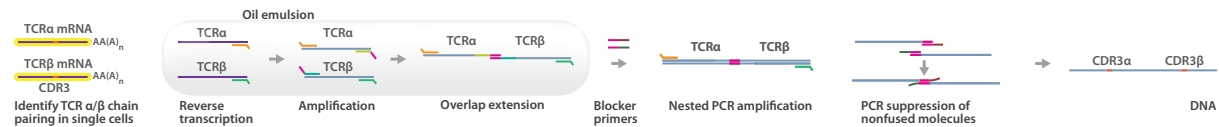
Cell-based emulsion RT-PCR technique for identifying TCR α-β chain pairing. Released TCR-α and TCR-β mRNAs are reverse-transcribed, amplified, and overlap-extended within each droplet. Products are extracted from the emulsion and fused molecules of interest are selectively amplified. Nonfused molecules are suppressed with blocking primers.[334]

References

Ma Y, Mattarollo SR, Adjemian S, et al. CCL2/CCR2-dependent recruitment of functional antigen-presenting cells into tumors upon chemotherapy. *Cancer Res.* 2014;74:436-445

Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet.* 2014;46:116-125

333. Woodsworth DJ, Castellarin M and Holt RA. Sequence analysis of T-cell repertoires in health and disease. Genome Med. 2013;5:98
334. Turchaninova MA, Britanova OV, Bolotin DA, et al. Pairing of T-cell receptor chains via emulsion PCR. Eur J Immunol. 2013;43:2507-2515
335. Kivioja T, Vaharautio A, Karlsson K, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods.* 2012;9:72-74

Unique Molecular Identifiers

UMIs are molecular tags that can be used to detect and quantify unique mRNA transcripts.[335] In this method, mRNA libraries are generated by fragmentation and then reverse-transcribed to cDNA. Oligo(dT) primers with specific sequencing linkers are added to the cDNA. Another sequencing linker with a 10 bp random label and an index sequence is also added to the 5' end of the template, which is amplified and sequenced. Sequencing allows for high-resolution reads, enabling accurate detection of true variants (Table 22).
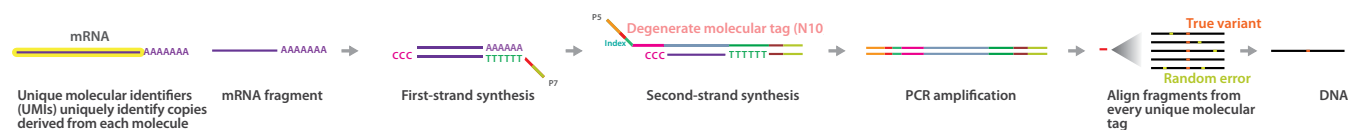
A schematic overview of UMIs.

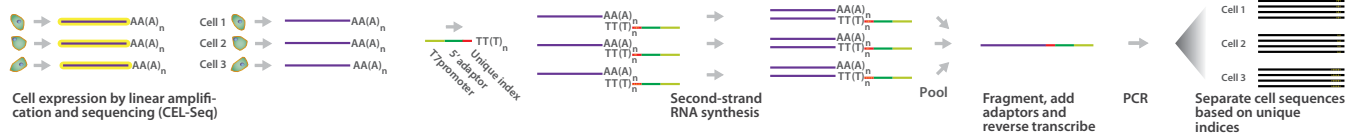Table 22. Advantages and Disadvantages of UMIs.

Advantages	Disadvantages
• Can sequence unique mRNA transcripts. • Can be used to detect transcripts occurring at low frequencies. • Transcripts can be quantified based on sequencing reads specific to each barcode. • Can be applied to multiple platforms to karyotype chromosomes as well.	• Targets smaller than 500 bp are preferentially amplified by polymerases during PCR.

For Research Use Only. Not for use in diagnostic procedures.

Reference

Cooper DA, Jha BK, Silverman RH, Hesselberth JR and Barton DJ. Ribonuclease L and metal-ion-independent endoribonuclease cleavage sites in host and viral RNAs. *Nucleic Acids Res.* 2014;42:5202-5216

Islam S, Zeisel A, Joost S, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods.* 2014;11:163-166

336. Hashimshony T., Wagner F., Sher N. and Yanai I. (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2: 666-673

337. Bhargava V, Head SR, Ordoukhanian P, Mercola M and Subramaniam S. Technical variations in low-input RNA-seq methodologies. *Sci Rep.* 2014;4:3678

338. Shapiro E, Biezuner T and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

Cell Expression by Linear Amplification Sequencing

CEL-Seq utilizes barcoding and pooling of RNA to overcome challenges from low input.[336] In this method, each cell undergoes reverse transcription with a unique barcoded primer in its individual tube. After second-strand synthesis, cDNA from all reaction tubes are pooled and PCR-amplified. Paired-end deep sequencing of the PCR products allows for accurate detection of sequence derived from sequencing both strands (Table 23).

A schematic overview of CEL-Seq.

Table 23. Advantages and Disadvantages of CEL-Seq.

Pros	Cons
• Barcoding and pooling allows for multiplexing and studying many different single cells at a time. • Contamination between samples is greatly reduced due to processing a single tube per cell. • Uses fewer steps than single-cell tagged reverse-transcription sequencing (STRT-Seq). • Shows very little read-length bias.[337] • Strand-specific.	• Strongly 3' biased. [338] • Abundant transcripts are preferentially amplified. • Requires at least 400 pg of total RNA.

Reference

Bose S, Wan Z, Carr A, Rizvi AH, Vieira G, et al. Scalable microfluidics for single-cell RNA printing and sequencing. *Genome Biol.* 2015;16:120

In this study, the authors present a new scalable high-density microfluidic platform for solid-phase capture of RNA on glass coverslips or on polymer beads. They trapped single-cell lysates in sealed picoliter microwells capable of printing RNA on glass or capturing RNA on beads. They combined this sample preparation approach with a scalable technology for scRNA-Seq based on CEL-Seq. The technology is relatively inexpensive, with consumable costs of $0.10–$0.20 per cell and is capable of processing hundreds of individual cells in parallel.

Illumina Technology: TruSeq RNA-Seq Library Preparation Kit, NextSeq 500 and HiSeq 2500 Systems

Flow Cell–Surface Reverse-Transcription Sequencing

Flow cell–surface reverse-transcription sequencing (FRT-Seq) is a transcriptome-sequencing technique developed in 2010.[339] It is strand-specific, free of amplification, and is compatible with paired-end sequencing. To begin with, poly(A)$^+$ RNA samples are fragmented by metal-ion hydrolysis and dephosphorylated. Next, P7 primers are ligated to the 3' end of the fragments. The adapter sequence starts at the 5' terminus with 20 nucleotides of RNA, followed by DNA nucleotides. The primers are also 5' phosphorylated and blocked with dideoxycytosine (ddC) at the 3' end. Following 3'-adapter ligation, fragments are size-selected for nucleotide fragments longer than the adapter. The 5' ends of the fragments are phosphorylated and ligated to P5 adapters. These adapters are blocked with an amino-C6 linker at the 5' end. Now that the fragments are flanked with adapters, they are hybridized to the flow cell and reverse-transcribed before cluster generation and sequencing (Table 24). FRT-Seq has the potential to overcome problems associated with RNA amplification in single-cell sequencing, but it has not yet been used in scRNA-Seq applications.[340]

339. Mamanova L, Andrews RM, James KD, et al. FRT-seq: amplification-free, strand-specific transcriptome sequencing. *Nat Methods.* 2010;7:130-132
340. Wang X. Single Cell Sequencing and Systems Immunology. *Translational Bioinformatics.* 2015;5:177
341. Islam S., Kjallquist U., Moliner A., Zajac P., Fan J. B., et al. (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome Res 21: 1160-1167

A schematic overview of FRT-Seq.

Table 24. Advantages and Disadvantages of FRT-Seq.

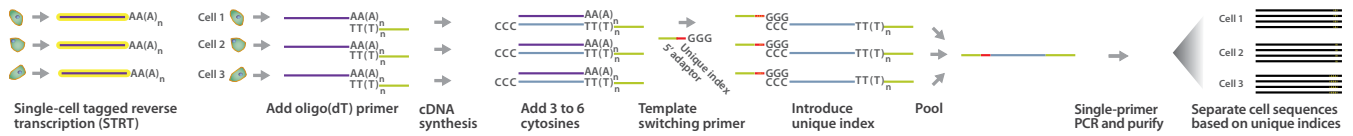| Pros	Cons
• Strand-specific poly(A)+ mRNA sequencing for transcriptome analysis. • No amplification step—gives more accurate representation of the total mRNA population, preventing amplification bias.	• Requires a large amount of input RNA material (250 ng). • Selects only poly(A)$^+$ mRNA samples.

Reference

Vergara-Irigaray M, Fookes MC, Thomson NR and Tang CM. RNA-seq analysis of the influence of anaerobiosis and FNR on Shigella flexneri. *BMC Genomics.* 2014;15:438

Single-Cell Tagged Reverse-Transcription Sequencing

STRT-Seq is a method similar to CEL-Seq that involves unique barcoding and sample pooling to overcome the challenges of samples with limited material.[341] In this method, single cells are first picked in individual tubes, where first-strand cDNA synthesis occurs using an oligo(dT) primer with the addition of 3–6 cytosines. A helper oligo promotes template switching, which introduces the barcode in the cDNA. Barcoded cDNA is then amplified by single-primer PCR. Deep sequencing allows for accurate transcriptome sequencing of individual cells (Table 25).

For Research Use Only. Not for use in diagnostic procedures.

An overview of recent publications featuring Illumina tecnology   **91**

A schematic overview of STRT-Seq.

Table 25. Advantages and Disadvantages of STRT-Seq.

Advantages	Disadvantages
• Barcoding and pooling allows for multiplexing and studying many different single cells at a time. • Contamination between samples is greatly reduced due to processing a single tube per cell.	• PCR biases can underrepresent GC-rich templates. • Nonlinear PCR amplification can lead to biases affecting reproducibility and accuracy. • Amplification errors caused by polymerases will be represented and sequenced incorrectly. • Targets smaller than 500 bp are preferentially amplified by polymerases during PCR.

Reference

Islam S, Zeisel A, Joost S, La Manno G, Zajac P, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods.* 2014;11:163-166

Fixed and Recovered Intact Single-Cell RNA Sequencing

FRISCR sequencing characterizes transcriptome profiles from fixed and stained single cells. First, the cell suspension is fixed with paraformaldehyde, permeabilized, and immunostained. Individual cells are then sorted into tubes using FACS. These cells are lysed and crosslinking is reversed crosslinkby incubation at 56°C for 1 hour. mRNA from the cells is isolated by dT25 magnetic bead pull-down. The mRNA sequencing library is prepared by following Smart-Seq2 procedures: 1) template-switching reverse transcription using Moloney murine leukemia virus reverse transcriptase; 2) PCR-amplifying the resulting cDNAs; and 3) preparing a cDNA library using the Nextera XT Library Preparation Kit. The fragments are now flanked with adapters and are ready for sequencing (Table 26).
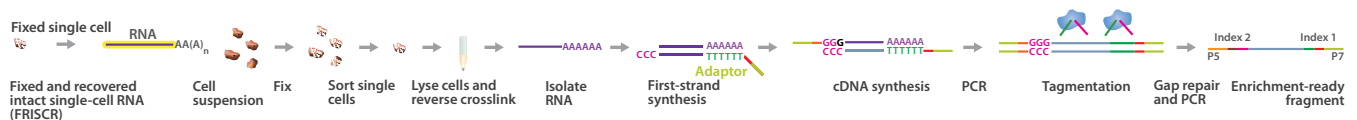
A schematic overview of FRISCR sequencing.

For Research Use Only. Not for use in diagnostic procedures.

92   Single-cell Research

Table 26. Advantages and Disadvantages of FRISCR Sequencing.

Advantages	Disadvantages
• Full-length mRNA transcriptome profiling from fixed and stained single cells. • Immunostaining enables targeting of rare cell populations. • Generates full-length mRNA reads. • Significantly more mRNA recovered compared to fixed cells from Triton-X100 lysis.	• 3' to 5' bias.

342. Picelli S, Bjorklund AK, Faridani OR, Sagasser S, Winberg G and Sandberg R. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods.* 2013;10:1096-1098

Reference

Thomsen ER, Mich JK, Yao Z, Hodge RD, Doyle AM, et al. Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat Methods.* 2016;13:87-93

The human neocortex develops from rare progenitor cells, especially RG. These cells have been difficult to characterize, since they are rare and are defined by a combination of position, morphology, and intracellular markers. The authors developed a method that allows RNA-Seq of individual fixed, stained, and sorted cells, known as FRISCR sequencing. They sorted individual RG cells by FACS and prepared single-cell mRNA libraries using Smart-Seq2[342] followed by sequencing using the MiSeq system. They demonstrated that expression data from fixed and purified single cells were similar to that obtained from live cells. Their data also identified subpopulations of ventricular zone–enriched RG and subventricular zone–localized RG, as well as new molecular markers for each subtype.

Illumina Technology: Nextera XT Library Preparation Kit, MiSeq System

Cell Labeling via Photobleaching

CLaP is a noninvasive, laser-based labeling technique for single cells. CLaP uses lasers to crosslink specific cells with fluorescent tags before isolating the single cells for sequencing.

In CLaP, cells of interest are tagged by crosslinking biotin-4-fluorescein (B4F) with the cell membrane using laser irradiation. Streptavidin-conjugate fluorescent labels are then bound to biotinylated cells. These steps can be repeated to tag multiple cell types with a variety of fluorescent tags. Tagged cells are subsequently isolated and processed to generate cDNA libraries before sequencing (Table 27).

Single cell

Cell labeling via photobleaching (CLaP) → Confluent cells in culture → Biotin-4-fluorescein (B4F) → Photobleach and crosslink with 473 nm laser → Rinse → Cy5-streptavidin labeling → Tagged cells isolated, reverse transcribed and → Barcoded mRNA from single cells
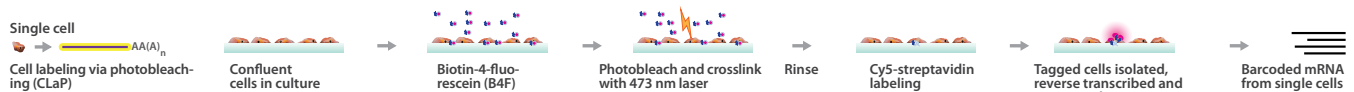
A schematic overview of CLaP.

Table 27. Advantages and Disadvantages of CLaP.

Advantages	Disadvantages
• Noninvasive, targeted laser-based single-cell labeling. • Automated image-based cell selection is possible. • Fluorescence-based tags can be substituted with other labels, such as electron-dense molecules. • Multicolored fluorescent stains can be used.	• Image-based selection limits the potential for high-throughput applications. • Diffusion of reagents through the extracellular matrix and continuous laser illumination limit the procedure for 3-dimensional environments/tissues. • Cellular specificity may be decreased slightly in primary cell cultures.

Reference

Binan L, Mazzaferri J, Choquet K, Lorenzo LE, Wang YC, et al. Live single-cell laser tag. *Nat Commun.* 2016;7:11636
Since single-cell sequencing methods often involve dissociation of cells and loss of spatial information, methods that retain spatial information in single-cell genomic analysis are critically important. The authors developed CLaP, a method that combines cellular labeling with single-cell genomics. Individual cells are labeled in culture by laser photobleaching, followed by isolation based on a wide variety of distinguishing characteristics. In this study, the authors used CLaP to tag a number of different cells from lines grown in monolayers. They isolated individual cells using drop-based microfluidics and performed RNA-Seq using the HiSeq 2500 system. The ability to combine spatial information with single-cell genomics makes this method well suited for studying tissue heterogeneity.

Illumina Technology: Nextera XT DNA Sample Preparation Kit, HiSeq 2500 System

343. Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* 2015;161:1202-1214

344. Hashimshony T, Wagner F, Sher N and Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* 2012;2:666-673

Indexing Droplets

Indexing droplets (inDrop) are used for high-throughput single-cell labeling. This approach is similar to Drop-Seq,[343] but it uses hydrogel microspheres to introduce the oligos.

Single cells from a cell suspension are isolated into droplets containing lysis buffer. After cell lysis, cell droplets are fused with a hydrogel microsphere containing cell-specific barcodes and another droplet with enzymes for reverse transcription. Droplets from all the wells are pooled and subjected to isothermal reactions for reverse transcription. The barcode-oligos anneal to poly(A)$^+$ mRNAs and act as primers for reverse transcriptase. Each mRNA strand now has cell-specific barcodes. The droplets are pooled, broken, and the mRNAs are purified. The 3' ends of the cDNA strands are ligated to adapters, amplified, annealed to indexed primers, and amplified further before sequencing (Table 28). The sequencing method is similar to CEL-Seq.[344]
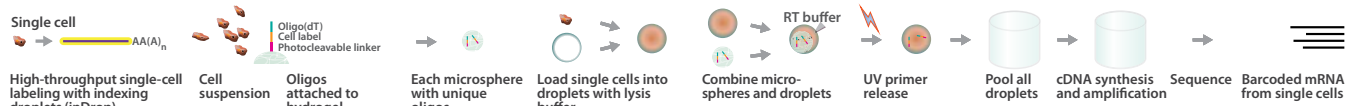
A schematic overview of inDrop.

Table 28. Advantages and Disadvantages of inDrop.

Advantages	Disadvantages
• High throughput single-cell transcriptome profiling using microfluidics. • Low cost: $0.1 per cell (experiments require 100 cells). • Highly scalable to larger cell quantities. • No fragmentation step.	• Droplets may contain 2 cells or 2 different types of barcodes.

References

Klein AM, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell.* 2015;161:1187-1201

A significant barrier to sequencing large numbers of single cells has been the challenge in physically isolating individual cells into separate chambers. Microfluidics platforms can yield highly scalable single-cell sequencing methods, but these systems remain costly with limits in sample throughput. The authors developed inDrop, a method based on high-throughput single-cell labeling (Hi-SCL) that physically separates individual cells in hydrogel drops, followed by RNA-Seq using cell-specific barcodes. In this study, they used the MiSeq and HiSeq systems to perform RNA-Seq of hundreds of droplet-isolated mouse ESCs and mouse embryonic fibroblasts (MEFs). Their data confirmed that inDrop data recapitulated bulk data, and that inDrop data could distinguish mouse ESCs from MEFs.

Illumina Technology: MiSeq, HiSeq 2500, and NextSeq Systems

Rotem A, Ram O, Shoresh N, et al. High-Throughput Single-Cell Labeling (Hi-SCL) for RNA-Seq Using Drop-Based Microfluidics. *PLoS One.* 2015;10:e0116328

Given the importance of single-cell data, there is a great need to increase the throughput of sequencing pipelines. Methods that physically separate large numbers of individual cells into wells or chambers of microfluidics chips are vital to this effort. Hi-SCL uses drop-based libraries of oligonucleotide barcodes to index individual cells. The drops are used as containers on a microfluidics platform, and the tagged molecules from different cells can be mixed without losing cell-of-origin information. In this study, the authors used MiSeq and HiSeq systems to validate Hi-SCL by performing RNA-Seq on hundreds of mouse ESCs and MEFs. They demonstrated that single-cell data could recapitulate bulk expression data and that single-cell data could distinguish ESCs from fibroblasts. Compared to Fluidigm C1 and CEL-Seq, Hi-SCL proved to be a faster and cheaper method for massively parallel sequencing.

Illumina Technology: MiSeq and HiSeq Systems

Cacchiarelli D, Trapnell C, Ziller MJ, et al. Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. *Cell.* 2015;162:412-424

Drop-Seq

Drop-Seq analyzes mRNA transcripts from droplets of individual cells in a highly parallel fashion. This single-cell sequencing method utilizes a microfluidic device to compartmentalize droplets containing a single cell, lysis buffer, and a microbead covered with barcoded primers. Each primer contains: 1) a 30 bp oligo(dT) sequence to bind mRNAs; 2) an 8 bp molecular index to uniquely identify each mRNA strand; 3) a 12 bp barcode unique to each cell; and 4) a universal sequence identical across all beads. Following compartmentalization, cells in the droplets are lysed, and the released mRNA hybridizes to the oligo(dT) tract of the primer beads. All droplets are then pooled and broken to release the beads within. After the beads are isolated, they are reverse-transcribed with a template-switching polymerase. This generates the first cDNA strand with a PCR primer sequence in place of the universal sequence. cDNAs are PCR-amplified and sequencing adapters added using the Nextera XT Library Preparation Kit. The barcoded samples are ready for sequencing (Table 29).
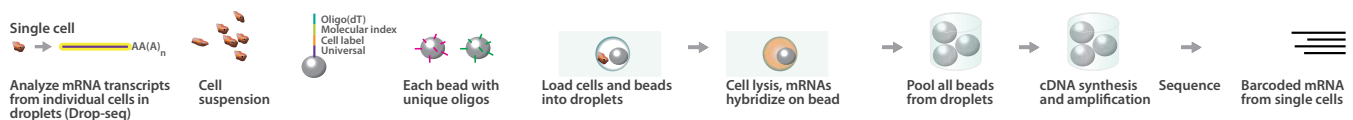
A schematic overview of Drop-Seq.

Table 29. Advantages and Disadvantages of Drop-Seq.

Advantages	Disadvantages
• Analyzes sequences of single cells in a highly parallel manner. • Unique molecular and cell barcodes enable cell- and gene-specific identification of mRNA strands. • Reverse transcription with template-switching PCR produces high-yield reads from single cells. • Low cost—$0.07 per cell ($653 per 10,000 cells)—and fast library prep (10,000 cells per day).	• Requires custom microfluidics device to perform droplet separation. • Low gene-per-cell sensitivity compared to other scRNA-Seq methods.[345] • Limited to mRNA transcripts.

345. Ziegenhain C, Vieth B, Parekh S, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. Mol Cell. 2017;65:631-643.e634

346. Fan HC, Fu GK and Fodor SP. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science.* 2015;347:1258367

References

Habib N, Avraham-Davidi I, Basu A, Burks T, Shekhar K, et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. *Nat Methods.* **2017;14:955-958**
Single-cell RNA-Seq has become an indispensable technique for characterizing cell types, states, and function in complex tissues. Sequencing individual nuclei has similarly proved useful in cell-type discrimination, particularly for tissues where single cells are difficult to dissociate. Current approaches for single-nucleus sequencing rely on the use of FACS or C1 microfluidics to sort individual nuclei into multiwell plates for downstream analysis. In this study, the authors developed DroNc-seq, a method that combines single-nucleus sequencing with the microfluidics-based Drop-seq method. The authors modified Drop-seq to incorporate single nuclei within microfluidic droplets, as opposed to single individual cells. They performed massively parallel RNA-Seq for approximate 39,000 individual nuclei from mouse and human brain, at low cost and with similar sensitivity to Drop-seq.

Illumina Technology: NextSeq 500 System

Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* **2015;161:1202-1214**
One of the bottlenecks in scRNA-Seq is the limitation in the number of individual cells that can be separated and analyzed. In this study, the authors developed Drop-Seq, a massively parallel scRNA-Seq method that uses uniquely barcoded primer beads together with captured single cells in droplets. This encapsulation method allows for processing of thousands of individual cells by RNA-Seq on a microfluidics platform. The authors validated this technique by applying it to the mouse retina. After Drop-Seq, they used the NextSeq system to perform RNA-Seq on approximately 45,000 cells, and they identified 39 distinct cell populations within mouse retina. Their results demonstrate that Drop-Seq can be used to understand the biology of complex tissues with diverse cell types.

Illumina Technology: Nextera XT DNA Sample Prep Kit, NextSeq 500 System

CytoSeq

The gene expression cytometry protocol known as CytoSeq enables gene expression profiling of thousands of single cells.[346] In this method, single cells are first randomly deposited into wells. A combinatorial library of beads with specific capture probes is added to each well. After cell lysis, mRNAs hybridize to the beads, which are then pooled for reverse transcription, amplification, and sequencing. Deep sequencing provides accurate, high-coverage gene expression profiles of several single cells (Table 30).
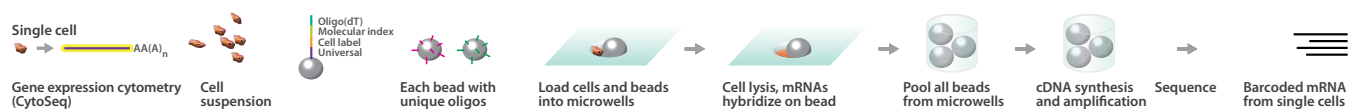
A schematic overview of CytoSeq.

Table 30. Advantages and Disadvantages of CytoSeq.

Advantages	Disadvantages
- Can readily scale to tens/hundreds of thousands of cells. - Complements and expands the capabilities of fluorescence or mass spectrometry–based cytometry. - Detects any transcribed mRNA without the limitations of antibody availability. - Enables rare cell characterization on small samples with insufficient cells for traditional flow cytometry. - Allows direct analysis of complex samples of heterogeneous cell size and shape.	- Sequencing depth requires large number of reads (eg, 200,000 transcripts per cell require 2 million reads for 10_x coverage: 2 billion reads for 1000 cells). - A single run can be relatively expensive and time consuming. - Involves a trade-off between depth of sequencing and differential gene expression.

347. Buenrostro JD, Wu B, Litzenburger UM, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature.* 2015;523:486-490

References

Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science.* **2015;348:910-914**

This study used combinatorial indexing to measure chromatin accessibility in thousands of single cells. The authors isolated nuclei and tagged them in bulk with Tn5 transposases, in each of many wells. Next, they pooled these barcoded nuclei, diluted them, and redistributed them to a second set of wells where a second barcode was introduced using PCR. The authors integrated this combinatorial indexing with scATAC-Seq[347] to measure chromatin accessibility for more than 15,000 human and mouse single cells. They sequenced the scATAC-Seq libraries on the MiSeq system, and their data identified relevant differences in chromatin accessibility between cell types.

Illumina Technology: MiSeq and NextSeq Systems

Fan HC, Fu GK and Fodor SP. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science.* **2015;347:1258367**

Combinatorial labeling of single cells is rapid and relatively inexpensive, and it can boost the throughput of massively parallel single-cell sequencing approaches dramatically. In this study, the authors developed CytoSeq, a method to label large numbers of individual cells combinatorially. Individual cells are placed in single wells, along with combinatorial libraries of beads containing cell- and transcript-barcoding probes. The authors performed CytoSeq on human PBMCs and used the MiSeq system to sequence amplified cDNAs. They analyzed several genes and were able to identify major subsets of PBMCs. In addition, by comparing cellular heterogeneity in naïve and CMV-activated CD8+ T cells, they identified rare cells specific to the CMV antigen. CytoSeq can be applied to complex mixtures of cells of varying size and shape, as well as to other biomolecules.

Illumina Technology: MiSeq System

Single-Cell RNA Barcoding and Sequencing

Single-cell RNA barcoding and sequencing (SCRB-Seq) is a cost-efficient, multiplexed scRNA-Seq technique. SCRB-Seq isolates single cells into wells using FACS. After cell lysis, poly(A)+ mRNAs are annealed to a custom primer containing a poly(T) tract, UMI, well barcode, and biotin. Template-switching reverse transcription and PCR amplification are carried out on the mRNA, generating barcoded full-length cDNA. cDNA strands from all wells are pooled together to be purified. They are amplified by PCR and purified further. cDNA libraries are prepared using the Nextera XT kit with modified i5 primers. The resultant cDNA fragments are size-selected for 300–800 bp and sequenced (Table 31).
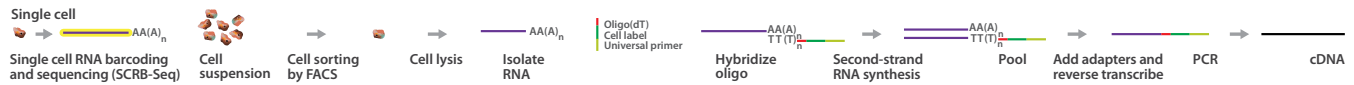
A schematic overview of SCRB-Seq.

Table 31. Advantages and Disadvantages of SCRB-Seq.

Advantages	Disadvantages
• Cost-efficient, high-throughput single-cell transcriptome profiling. • Highly sensitive gene-detection results compared to popular scRNA-Seq techniques.[348]	• Template-switching reverse transcription is heavily biased to full-length mRNA.[349]

Reference

Soumillon M, Cacchiarelli D, Semrau S, van Oudenaarden A and Mikkelsen TS. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. *bioRxiv.* 2014;10.1101/003236

348. Ziegenhain C, Vieth B, Parekh S, et al. Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell.* 2017;65:631-643. e634

349. Shapiro E, Biezuner T and Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet.* 2013;14:618-630

350. Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell.* 2015;161:1202-1214

351. Klein AM, Mazutis L, Akartuna I, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell.* 2015;161:1187-1201

High-Throughput Single-Cell Labeling

Hi-SCL generates transcriptome profiles for thousands of single cells using a custom microfluidics system, similar to Drop-Seq[350] and inDrop.[351] Single cells from a cell suspension are isolated into droplets containing lysis buffer. After cell lysis, cell droplets are fused with a droplet containing cell-specific barcodes and another droplet with enzymes for reverse transcription. Droplets from all the wells are pooled and subjected to isothermal reactions for reverse transcription. The barcode-oligos anneal to poly(A)+ mRNAs and act as primers for reverse transcriptase. Now that each mRNA strand has cell-specific barcodes, droplets are broken and the mRNAs are purified. The 3' ends of the cDNA strands are ligated to adapters, amplified, annealed to indexed primers, and amplified further before sequencing (Table 32).
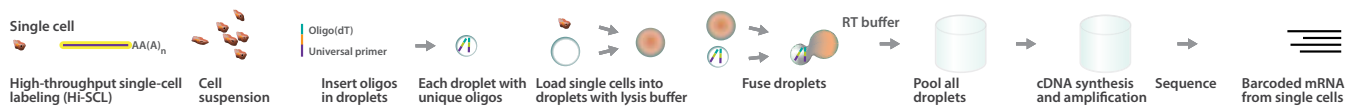
A schematic overview of Hi-SCL.

Table 32. Advantages and Disadvantages of Hi-SCL.

Advantages	Disadvantages
• High-throughput, single-cell transcriptome profiling using a microfluidics system. • Low cost—$0.1 per cell (experiment requires 100 cells). • Highly scalable to larger cell quantities. • No fragmentation step.	• Lack of UMIs in oligos may create amplification noise. • Droplets may contain 2 cells or 2 different types of barcodes.

References

Rotem A, Ram O, Shoresh N, Sperling RA, Schnall-Levin M, et al. High-Throughput Single-Cell Labeling (Hi-SCL) for RNA-Seq Using Drop-Based Microfluidics. *PLoS One.* **2015;10:e0116328**

Given the importance of single-cell data, there is a great need to increase the throughput of sequencing pipelines. Methods that physically separate large numbers of individual cells into wells or chambers of microfluidics chips are vital to this effort. Hi-SCL uses drop-based libraries of oligonucleotide barcodes to index individual cells. The drops are used as containers on a microfluidics platform, and the tagged molecules from different cells can be mixed without losing cell-of-origin information. In this study, the authors used MiSeq and HiSeq systems to validate Hi-SCL by performing RNA-Seq on hundreds of mouse ESCs and MEFs. They demonstrated that single-cell data could recapitulate bulk expression data and that single-cell data could distinguish ESCs from fibroblasts. Compared to Fluidigm C1 and CEL-Seq,[352] Hi-SCL proved to be a faster and cheaper method for massively parallel sequencing.

Illumina Technology: MiSeq and HiSeq Systems

Cacchiarelli D, Trapnell C, Ziller MJ, Soumillon M, Cesana M, et al. Integrative Analyses of Human Reprogramming Reveal Dynamic Nature of Induced Pluripotency. *Cell.* 2015;162:412-424

352. Hashimshony T, Wagner F, Sher N and Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep.* 2012;2:666-673

Scientific Publication Reviews can be accessed at
www.illumina.com/pubreviews

Genomic Solutions for Cell Biology and Complex Disease Research
Complex diseases are the result of multiple genetic and environmental factors. They are distinguished from Mendelian traits (or simple traits) as they do not follow a specific model of inheritance and are usually more frequent in the population. Although some of these diseases are highly heritable, currently known genetic variants can explain only some of the estimated heritability. This review gives a general overview on how genomic technologies and NGS can help in the study of complex diseases.

Cancer and Immune System Research Review
Advances in high-throughput sequencing have dramatically improved our knowledge of the cancer genome and the intracellular mechanisms involved in tumor progression and response to treatment. While the primary focus to date has been on the cancer cell, this technology can also be used to understand the interaction of the tumor cells and the cells in the surrounding tumor microenvironment.
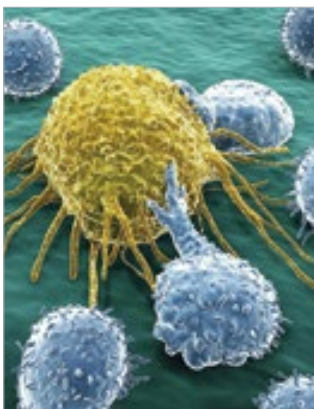
Expression analysis of the RNA levels can be used to determine the activation of pathways in the tumor microenvironment. Since common signaling pathways are involved in manifestation of several hallmarks of cancer, including cancer cell proliferation, survival, invasion, metastasis, and immunosuppression, targeting these shared signaling pathways in combination with immunotherapy may be a promising strategy for cancer treatment.

This Scientific Publication Review is brought to you by Illumina, Inc.

illumına®