



Published in final edited form as:

*Oral Oncol.* 2019 December ; 99: 104441. doi:10.1016/j.oraloncology.2019.104441.

## Single-cell sequencing and its applications in head and neck cancer

Zongtai Qi<sup>1,\*</sup>, Thomas Barrett<sup>1,\*</sup>, Anuraag S. Parikh<sup>2,3</sup>, Itay Tirosh<sup>4</sup>, Sidharth V. Puram<sup>1,5,†</sup>

<sup>1</sup>Department of Otolaryngology-Head and Neck Surgery, Washington University School of Medicine, St. Louis, USA

<sup>2</sup>Department of Otolaryngology, Massachusetts Eye and Ear, Boston, MA, USA

<sup>3</sup>Department of Otolaryngology, Harvard Medical School, Boston, MA, USA

<sup>4</sup>Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel

<sup>5</sup>Department of Genetics, Washington University School of Medicine, St. Louis, USA

### Abstract

Head and neck squamous cell carcinoma (HNSCC), like many tumors, is characterized by significant *intra*-tumoral heterogeneity, namely transcriptional, genetic, and epigenetic differences that define distinct cellular subpopulations. While it has been established that *intra*-tumoral heterogeneity may have prognostic significance in HNSCC, we are only beginning to describe and define such heterogeneity at a cellular resolution. Recent advances in single-cell sequencing technologies have been critical in this regard, opening new avenues in our understanding of more nuanced tumor biology by identifying distinct cellular subpopulations, dissecting signaling within the tumor microenvironment, and characterizing cellular genomic mutations and copy number aberrations. The combined effect of these insights is likely to be robust and meaningful changes in existing diagnostic and treatment algorithms through the application of novel biomarkers as well as targeted therapeutics. Here, we review single-cell technological and computational advances at the genomic, transcriptomic, and epigenomic levels, and discuss their applications in cancer research and clinical practice, with a specific focus on HNSCC.

### Introduction

Worldwide, squamous cell carcinoma of the head and neck (HNSCC) accounts for more than 650,000 cancer diagnoses and 330,000 deaths annually [1]. Despite its high incidence, management and control of this disease remains a challenge, particularly in patients with

† Corresponding author: sidpuram@wustl.edu, Washington University School of Medicine, 660 S. Euclid Ave., Campus Box 8115, St. Louis, MO 63110, Phone: 314-273-9004.

\*These authors contributed equally to this work

Conflicts of Interest

None declared

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

human papilloma virus (HPV)-negative disease [2]. Advances in treatment will critically depend on an improved understanding of the underlying biology, including characterization of HNSCC using high-resolution technologies and techniques.

To date, much of our understanding of the mutational landscape and mechanisms underlying HNSCC – its driver mutations, aberrant regulatory programs, and expression subtypes – has been derived from genomic and transcriptomic studies utilizing *bulk* sequencing technologies, such as whole exome sequencing (WES) and bulk RNA sequencing (e.g. The Cancer Genome Atlas; TCGA) [3,4] (Figure 1). These methodologies, however, fail to capture the diversity of individual cells and cellular subpopulations, known as *intra*-tumoral heterogeneity, that characterizes HNSCC. This diversity among malignant, stromal and immune cells leads to immune-cancer and stromal-cancer interactions that are further affected by spatial dynamics and clonal evolution [5]. Such changes in the HNSCC microenvironment can affect the development, progression, and metastasis of cancer [6]. Thus, a high-resolution strategy to effectively screen and systematically analyze this heterogeneity is critical to advance the field in new and unanticipated directions.

Single-cell sequencing (SCS) technologies have rapidly developed over the last decade, including methods to interrogate the single-cell genome, epigenome and transcriptome (Figure 1). These quickly evolving methods have and will continue to lead to new discoveries in tumor biology, metastasis, immuno-oncology, and therapeutic resistance and response. In this review, we will describe the current state of single-cell technology, summarize its applications in HNSCC and other human cancers, and propose major avenues of future investigation, with a focus on how this technology may impact clinical practice. Accordingly, the first portion of this review is directed towards those wishing to implement these technologies in their own research and offers a detailed review of current SCS methodologies, while the second portion focuses on clinicians who may be interested in how SCS could revolutionize head and neck cancer diagnosis and treatment.

## Single-cell isolation methods

The first step of all SCS is separation and isolation of viable, individual cells. Currently, several approaches are available for isolating single cells (Table 1), including limiting dilution [7], micromanipulation [8], laser capture microdissection (LCM) [9], fluorescence-activated cell sorting (FACS) [10], microfluidics [11], microwell [12,13] and *in situ* barcoding [14,15]. Of these methods, microfluidics has become popular due to its low sample consumption, precise fluid control, and low operating costs. Fluidigm C1 is one of the major microfluidic commercial platforms, providing automated single-cell lysis, RNA extraction, and cDNA synthesis for up to 800 cells in parallel on a single chip [16]. However, this approach is limited by its dependence on homogeneously sized cells. An alternative commercial option is the Chromium system from 10x Genomics, which offers high-throughput profiling of single-cell transcriptomes with high capture efficiency, arguably enabling discovery of rare cell types in a heterogeneous sample. More recently, a novel scalable isolation strategy, known as *in situ* barcoding, was introduced, which labels the cellular origin of RNA through *in situ* combinatorial barcoding. With this approach, more than 10,000 single-cell transcriptomes can be captured without requiring the physical

isolation of each cell [14,15]. The high-throughput capacity of these new technologies represents a marked departure from the early experience by our group and others with individual, well-based separation and amplification of single cells (e.g. Smart-seq2), and it is likely these approaches will increase the resolution of single-cell experiments and potentially improve the ability to detect rare cell subtypes or transitional states. However, we offer the caveat that these new methods may produce lower quality cells that may act as a trade-off to the large number of cells detected.

## Single-cell genome/DNA sequencing

Whole genome amplification (WGA) of the single-cell genome is always required for single-cell DNA sequencing (scDNA-seq). There are three main WGA methods with different performance with respect to genome coverage and uniformity (Table 2): (a) PCR amplification, (b) isothermal amplification by multiple displacement amplification, and (c) hybrid method. The PCR-based method uses degenerate oligonucleotide primers for amplification [17]. The isothermal WGA method employs polymerase Phi29 that has a lower error rate and better processivity and strand-displacement activity, providing higher genome coverage and lower false-positive rates[18]. The limitation of isothermal WGA is preferential amplification of the loci amplified first, leading to reduced uniformity and a high rate of allelic dropout. The third hybrid method combines advantages of the first two methods by using an initial limited isothermal amplification followed by PCR amplification and thus reduces associated biases. The most commonly used protocols are PicoPLEX [19] and MALBAC [20]. Notably, although WGA increases the sensitivity of the scDNA-seq, single-cell DNA studies mostly focus on copy number variations (CNVs) detection as accurate identification of single-nucleotide variations (SNVs) requires a reasonably high sensitivity and is still a challenge in the field. In a recent study [21], MALBAC was used for SNV detection in single cells derived from circulating tumor cell (CTC) isolation methods. All expected SNVs in control samples were identified as well as two different SNVs in several CTCs isolated from peripheral blood of breast cancer patients. Interestingly, the newly detected SNVs were associated with acquired resistance to treatment in patients not responding to endocrine therapy.

## Single-cell transcriptome/RNA sequencing

Single-cell transcriptome sequencing, often referred to as single-cell RNA sequencing (scRNA-seq), is used to measure gene expression at the single-cell level. ScRNA-seq has been more accessible than single-cell genome/DNA sequencing because each cell contains numerous RNA molecules, but just two copies of DNA molecules. It captures cellular differences underlying tumor biology at a higher resolution than regular bulk RNA-seq and has revolutionized studies of cancer mechanisms and personalized medicine. The workflow of scRNA-seq consists of single-cell capture, cDNA synthesis, cDNA amplification, library preparation and high-throughput sequencing. There are three major strategies to perform cDNA synthesis and amplification, which are essential steps and determine the sensitivity and specificity of scRNA-seq (Table 3). One strategy uses poly(A) tailing followed by PCR, as in the Tang-seq [22]. Another method uses second-strand synthesis followed by in vitro transcription, a linear amplification, as in CEL-seq/seq2 [23,24] and MARS-seq [25]. The

drawback of these two methods is the premature termination of reverse transcription which significantly reduces transcript coverage at the 5' end. A third approach uses a template-switching method, as in STRT-seq/seq2 [26,27] and Smart-seq/seq2 [28,29]. When possible, we favor this approach as it reduces 3' coverage biases from incomplete reverse transcription and yields full-length transcript coverage. The cDNA from this method can be amplified at a later point to specifically enrich and investigate mutational hot spots, although doing so is labor intensive. Regular gene expression analysis can be performed using either 3' or 5' fragments of genes, while for studies of T cell receptors (TCRs) and B cell receptors (BCRs), 5' end sequencing is critical for tracking immune clones. Investigators must be aware of these important differences before committing to a technological approach for their studies.

In order to control the technical variances among different cells, External RNA Control Consortium (ERCC) [30] and unique molecular identifiers (UMIs) [31] have been used. It is worth noting that ERCC and UMIs are not applicable to all scRNA-seq technologies due to inherent protocol differences. ERCC is used in approaches like Smart-seq2 but is not compatible with droplet-based methods, whereas UMIs are typically applied to 3'-end sequencing technologies (Table 3).

### Single-cell epigenome sequencing

In addition to single-cell genome and transcriptome sequencing, single-cell epigenome sequencing is an emerging area of interest that will enable studies of epigenetic or chromatin marks individual cells. Alterations of epigenetic marks and chromatin structure dictate gene expression, induce cellular heterogeneity, and are relevant to numerous disease states [32]. Thus, single-cell epigenetics will surely have an important role in cancer studies and clinical investigations.

### Single-cell DNA methylation sequencing

DNA modification such as methylation in CG rich regions has been associated with transcriptional silencing. To study DNA methylation at the single-cell level, several sequencing technologies have been reported using either bisulfite conversion or methylation-sensitive restriction enzymes (Table 4). For bisulfite conversion, the first single-cell, multi-locus method utilized a modified version of reduced representation bisulfite sequencing (RRBS-seq), performing all reactions for a single cell in one tube in order to reduce DNA loss [33]. scRRBS-seq enabled the detection of around 1.5 million CpG sites (CpGs) within the genome of a single cell. Smallwood et al. [34] developed a single-cell genome wide bisulfite sequencing method (scBS), covering around 3.7 million CpGs. Several recent developments to the basic scRRBSseq and scBS-seq protocols have increased throughput, decreased amplification bias, and improved data analysis [35,36]. Alternatively, the endonuclease-based sequencing methods avoid random DNA loss from bisulfite treatment and thus improve coverage and efficiency. Genome-wide CpG island (CGI) methylation sequencing for single cells (scCGI-seq) is one such method that enriches for sequences with high CpG content, providing 72.7% CGI coverage per cell [37].

### Single-cell histone modification sequencing

Histone modifications regulate the affinity and accessibility of certain DNA regions. To decipher the role of histone modifications in different cellular functions, single-cell chromatin immunoprecipitation (ChIP)-seq (scChIP-seq) was firstly developed [38]. Then, three new single-cell methods for detecting histone modifications were reported recently (Table 4): single-cell chromatin integration labeling followed by sequencing (scChIL-seq) [39], single-cell cleavage under targets and tagmentation (scCUT&Tag) [40], and single-cell chromatin immune-cleavage followed by sequencing (scChIC-seq) [41]. These three methods replace traditional immunoprecipitation and sonication with *in situ* reactions inside nuclei. However, they suffer from issue with sensitivity and technical challenges which limit their use for unbiased analysis. It is expected that these advanced single-cell techniques will have more applications in cancer genomics and translational research as their feasibility matures.

### Single-cell chromatin structure sequencing

Chromatin structure and organization play a central role in embryogenesis, differentiation, lineage specification, and disease evolution. Several methods have been developed to characterize the chromatin features within an individual cell [42–44]. One method is single-cell assay for transposase accessible chromatin (scATAC-seq), which can be used to functionally identify relevant changes in chromatin structure among specific subpopulations of cancer cells (Table 4) [42]. Through these single-cell epigenomic studies, information about chromatin modifications and their potential regulatory effects can be analyzed at a single-cell resolution, thereby complementing data on DNA variation and RNA expression gained through single-cell DNA and RNA sequencing, respectively. Interestingly, a recent study that built an immune cell atlas has shown that analysis of chromatin accessibility by scATAC-seq at distal enhancers results in sharper cell classification than analysis based on RNA expression or accessibility of transcription start sites [45].

### Single-cell data analysis

Bulk technologies characterize samples via a feature-by-sample matrix whereas SCS methods add an orthogonal dimension of cellular layer between feature and sample, resulting in a feature-by-cell-by-sample data structure. Due to minimal amount of starting material, SCS data tends to be sparse and may have batch effects, amplification biases, and dropout events. Therefore, analysis of SCS data requires specific computational tools and expertise.

For single-cell genome/DNA sequencing, most of the analysis methods focus on quantifying single-cell CNAs and SNVs. SNV calling needs to manage mutations and the allelic imbalances that occur during genome amplification and sequencing. Mutations can be corrected by using either a bulk sample as a reference or two or three cells required to have the same variant at the same location [20]. The allelic imbalance can be removed by requiring that all variant calls be above the level of technical noise in control samples [46]. To decrease the sequencing error rate, molecular barcoding [47] and algorithm correction [48] can be used. CNA detection relies on algorithms that can normalize noisy coverage data after single-cell WGA to identify regions that are over- or under-represented compared with a

diploid genome [20,49]. CNA detection algorithms are currently developed to specifically address these technical artifacts introduced during specific types of single-cell WGA [50,51].

For single-cell RNA sequencing, many bioinformatic tools have been created to generate matrix counts and identify real signals from background noise (<https://www.scrna-tools.org/>). The basic analytical workflow includes several critical steps (Figure 2). Raw sequencing reads in FASTQ format are de-multiplexed by cell barcode and collapsed by UMI [52]. Then, a series of quality control (QC) analyses are required to eliminate the low-quality cells. First, cells that contain fewer reads, low genome mapping ratios, or high mitochondrial mapping ratios should be eliminated [53,54]. The number of detected genes per cell is then calculated. Although the number of detected genes is highly variable between cell types (high in transcriptionally active malignant cells compared to low in naïve T-cells), establishing rigorous thresholds is critical to avoid the incorporation of low-quality cells. We typically use a threshold of 3,000 genes among malignant and stromal cells and 2,000 genes for immune cells in our analyses. Lastly, the batch effects can be mitigated by several methods such as MNN (mutual nearest neighbor) [55] and kBET (k-nearest neighbor batch effect test) [56]. After low-quality cell elimination and batch effect correction, the gene count matrix is generated. To correctly interpret the results, normalization and imputation can be performed to obtain the meaningful signal by adjusting biases and missing values resulting from capture efficiency, sequencing depth, dropouts, and other technical effects. The common computational tools for normalization and imputation include SCnorm [57], SAVER [58], MAGIC [59], ScImpute [60], DrImpute [61], and AutoImpute [62]. The corrected count matrix can be further analyzed by R packages and python-based tools [63] at both cell and gene levels. Cell level analysis includes clustering to identify cell subpopulations by dimension reduction algorithms (PCA, t-SNE, UMAP) and cell lineage and pseudotime reconstruction by Monocle2 [64] or CellRouter [65]. On the other hand, the gene level analysis can be accomplished through differential gene expression algorithms such as DEsingle [66], Census[67], BCseq [68] and Expression Variation Analysis [69] as well as regulatory network reconstruction tools such as SCENIC [70] and PIDC [71]. Such network inference tools facilitate the identification of expression regulatory network from single-cell transcriptomic data and provide critical biological insights into the regulatory relationships between genes.

## Applications in HNSCC and cancer genomics

Single-cell technologies have been critical in beginning to dissect the cellular and biological complexities of human cancers, many of which are highly heterogeneous, owing to a variety of stromal and immune cell types, in addition to distinct subpopulations of malignant cells (Figure 3). Compared to melanoma [72–74], breast [75–80], lung [81–83], and colorectal carcinoma [84–86], relatively few studies have utilized single-cell methodologies to investigate HNSCC [87].

### Intra-tumoral heterogeneity

Over recent years, our appreciation for the diversity of cells – malignant, stromal, and immune– that can exist within a single tumor has grown immensely, but our understanding of how this heterogeneity influences tumor behavior and clinical outcomes remains nascent [88,89]. In patients with HNSCC, *intra*-tumoral genetic heterogeneity is associated with worse patient outcomes, as has been demonstrated by mutant-allele tumor heterogeneity (MATH) scores [90,91].

In the first scRNA-seq study in HNSCC, our group investigated *intra*-tumoral expression heterogeneity in oral cavity squamous cell carcinoma (SCC) tumors. Using this approach, we found that malignant cells varied within and between tumors in their expression of programs related to cell cycle, hypoxia, stress, epithelial differentiation, and partial epithelial-to-mesenchymal transition (p-EMT, see “Metastasis” below) [92]. In addition, we utilized our single-cell profiles to deconvolute bulk transcriptional profiles from TCGA, thereby refining previously described HNSCC expression subtypes (atypical, mesenchymal, basal, and classical) [93]. Strikingly, we discovered that malignant cells exclusively mapped to the basal, classical, and atypical subtypes, while only fibroblasts mapped to the mesenchymal subtype. Using a computational approach to normalize the expression of non-malignant cells and develop a malignant cell-specific signature, we found that the averages of malignant cells from TCGA mesenchymal tumors were, in fact, indistinguishable from those from TCGA basal tumors, suggesting that in oral cavity SCC, the majority of tumors (>80%) can be defined as a single ‘malignant-basal’ cohort, with mesenchymal subtype tumors defined only by a large fibroblast infiltrate. These observations raise the exciting possibility that targeted therapies directed against ‘malignant-basal’ subtype tumors may be effective in a large proportion of oral cavity SCC patients.

Interestingly, while HNSCC malignant cells of a given tumor map exclusively to one expression subtype, glioblastoma multiforme (GBM) tumors are composed of malignant cells mapping to multiple different subtypes. We recently used scRNA-seq to profile 28 primary GBM tumors and demonstrated that glioblastoma tumors are composed of malignant cells in four main cellular states, which are derived from distinct cells of origin [94]. Together, these findings suggest that unlike GBM where a tumor has varying proportions of multiple malignant subtypes, HNSCC tumors consist of a single malignant subtype, potentially enabling clinical trials stratified by tumor subtype. Moving forward, it will be critical to expand similar analyses to other tumors to determine the landscape of heterogeneity that is present and how well prior TCGA studies across oncology have captured tumor subtypes. Given the expanding interest in using TCGA subtypes to guide targeted clinical trials, clarifying this point has never been more important.

### Characterizing the tumor microenvironment

Single-cell technologies have been utilized to describe the immune and stromal cells of the tumor microenvironment (TME) of HNSCC [92], metastatic melanoma [74], GBM [95], breast cancer [79], colorectal cancer [85], and lung cancer [81] in a way that was not possible with bulk approaches. In our scRNA-seq analysis of oral cavity SCC [92], for example, we found that non-malignant cells clustered into eight main clusters by cell type: T

cells, B/plasma cells, macrophages, dendritic cells, mast cells, endothelial cells, fibroblasts, and myocytes. These non-malignant cells, however, did not cluster by tumor of origin, suggesting that the cell types and their expression states are consistent across tumors. In contrast, malignant cells clustered by patient, suggesting distinct genetic background of these cells drives distinct global transcriptional profiles. These findings suggest that therapies targeted toward stromal or immune cells (e.g. immune-checkpoint inhibition) may be applicable across tumors (and patients), while directed treatments against malignant cells may require a more personalized approach.

We defined four T-cell expression states specific to HNSCC, including regulatory T cells ( $T_{\text{regs}}$ ), CD4+ T-helper cells, and activated and exhausted subsets of cytotoxic CD8+ T cells [92]. The cytotoxic subsets expressed varying genes associated with T cell dysfunction and exhaustion (e.g. PD1, CTLA4) and could be clustered into CD8+ T and CD8+  $T_{\text{exhausted}}$  cells. These studies provided the first glimpse of *in vivo* signatures of exhaustion in HNSCC patients, defining a host of novel potentially-targetable markers and raising the possibility that immune checkpoint inhibitors (ICIs) that specifically target these exhaustion markers may be more efficacious. Outside of HNSCC, attempts to understand cancer-specific states of exhaustion have been used in melanoma [72,73,96] and non-small cell lung cancer (NSCLC). Guo *et al.* utilized deep scRNA-seq to identify two novel clusters of T cells exhibiting states appearing to precede exhaustion (“pre-exhausted” T cells) [81]. Furthermore, they determined that patients with a high ratio of “pre-exhausted” to exhausted T cells had an improved prognosis. These findings have important implications not only for non-small-cell lung cancer patients, but for other types of cancer being treated with immunomodulating agents. Future work must detail these T-cell expression states on the single-cell level at pre-treatment, on-treatment, and post-treatment timepoints to better identify and predict likely responders versus non-responders.

T cell receptor (TCR) analysis is also of great interest in understanding the dynamics of T cell clonality and identifying potential tumor antigens [97,98]. Understanding T cell clonality and diversity on a single-cell level will help better elaborate the tumor-immune response to ICI. Yost *et al.* utilized the combination of scRNA-seq and TCR sequencing on primarily cutaneous basal cell carcinoma samples before and after anti-PD-1 therapy [99]. They demonstrated clonal expansion of CD8+ T cells with treatment, with expanded clones being novel clonotypes not detected in the pre-treatment sample. Their findings suggest that CD8+ exhausted T cells have a limited capacity for reinvigoration and that ICI response in tumors rich in immune cells may be due to recruitment of “new” T cells. Ongoing ICI clinical studies should incorporate single-cell data to better understand the mechanisms of response and resistance. Furthermore, improved understanding of T cell clonality will improve neoantigen prediction, serving as a guide for tumor vaccine development and adoptive T cell interventions including chimeric antigen T cell (CAR-T cell) therapy.

In addition to the presence and influence of immune cells in the TME, stromal cells such as cancer-associated fibroblasts (CAFs) have a complex and not yet fully understood role in the TME. Normal fibroblasts exist in a quiescent state and can be activated – classically, in wound healing – to synthesize extracellular matrix (ECM), produce cytokines and chemokines, recruit immune cells, and exert physical forces to modify tissue architecture



[100]. Indeed, CAFs have been implicated in tumorigenesis, tumor survival, ECM remodeling, immune system suppression, and tumor invasion (see “Metastasis” below) [100–102]. Li *et al.* used scRNA-seq to define two populations of CAFs in colorectal carcinoma patients [85], termed CAF-A and CAF-B. Similar to CAF clusters we identified in HNSCC [92], CAF-A cells expressed ECM-related genes and CAF-B cells expressed markers of myofibroblasts (e.g. ACTA2, PDGFA). Future single-cell studies must continue to subtype CAFs, define their wide variety of expression states, their influence on therapeutic resistance, and their complex intercellular interactions. On a broader scale, studies of nearly all stromal cells of the TME including CAFs, endothelial cells, and myocytes are needed to comprehensively characterize cross-talk with malignant populations.

### Invasion and metastasis

Invasion and metastasis represents a highly complex process that is one of the least well understood aspects of cancer biology, yet accounts for the majority of cancer-associated deaths [103]. Both locoregional and distant failure are significant problems for patients with advanced HNSCC, with more than 50% having recurrence or developing metastases within 3 years of treatment [104].

In an attempt to understand metastasis in HNSCC, our scRNA-seq analysis characterized matched lymph node metastases from five patients [86]. We found that *intra*-tumoral heterogeneity and tumor nest architecture were largely recapitulated within lymph node metastases [86]. However, when we assessed the localization of specific malignant subpopulations, we found that p-EMT cells were exquisitely localized at the leading edge of tumors in close apposition to CAFs. In further studies, we found that TGF $\beta$  from surrounding CAFs likely induces a p-EMT state in malignant cells, promoting invasion into the surrounding tissue. We therefore assessed p-EMT in malignant cell-specific profiles based on bulk TCGA data and found that p-EMT was associated with lymph node metastases, lymphovascular invasion, and extracapsular extension in malignant-basal subtype tumors. Together, these findings indicate a role for p-EMT in locoregional metastasis and raise the intriguing possibility of a “collective migration” model of metastases, in which malignant and stromal cells move in clusters to spread lymphatogenously and form nodal metastases, rather than individual malignant cells extravasating and engrafting in the lymph node bed. These findings have implications for systemic therapy as metastases that heavily mirror the primary tumor may be amenable to therapy targeted at the primary tumor, while analyses of the primary tumor may provide insight into distant metastases that may not be readily biopsied or resected. In addition, our results raise the tantalizing possibility that analysis of p-EMT markers in primary tumors using IHC may improve prediction of occult nodal or distant metastasis [105].

Other proposed models of metastasis include late dissemination (“linear”) and early dissemination (“parallel”), and distinguishing between the two may be critical to predicting the success of surgical intervention. The linear model postulates that tumor cells acquire mutations in a stepwise manner until the tumor eventually gains invasive capacity, whereas the parallel model posits that metastases disseminate early in tumor development and gradually acquire their own distinct set of mutations divorced from the progression of the

primary tumor. The clinical implications are significant: The linear model would support early surgical excision of a malignant lesion to prevent metastases, while the parallel model suggests that such measures may prove futile in select patients and that both primary and metastatic sites may require separate treatment plans. Single-cell DNA sequencing may be the key to delineating between these two possibilities through phylogenetic analysis of mutations. For example, Leung *et al.* recently utilized single-cell DNA sequencing, exome sequencing, and targeted deep-sequencing to study clonal evolution during metastatic dissemination in two colorectal patients in whom primary tumor and liver metastases were analyzed [86]. While this study was limited by a very small sample size, one patient demonstrated a clear linear progression of acquired mutations prior to metastatic spread, while another patient demonstrated a linear acquisition of mutations prior to a first metastasis, then developed additional mutations in the primary before seeding a second metastasis. Future single-cell analyses that capture clonality will be critical for unraveling whether HNSCC consistently adheres to either of these models of metastasis, and HPV (+) oropharyngeal carcinoma may represent a particularly interesting cohort, given its propensity to present with neck metastases.

Understanding clonality at a single-cell level will be further enhanced through studies that can simultaneously capture and maintain spatial information. Casasent *et al.* developed a method called Topographic Single Cell Sequencing (TSCS), which utilizes LCM followed by single-cell DNA-sequencing, to measure genomic copy number profiles of single tumor cells while preserving their spatial context in breast cancer patients with both ductal carcinoma in situ (DCIS) and invasive ductal carcinoma (IDC) [76]. Prior studies had presented conflicting models for invasion in this disease, with one model suggesting that the *in situ* and invasive malignant cell subpopulations evolved in parallel without sharing genomic aberrations (“independent lineage model”) and the other suggesting that multiple clones evolve within the ducts until a single clone escapes the basement membrane to form the invasive tumor mass (“bottleneck model”). Using TSCS, they demonstrated that genome evolution occurred within the ducts, all subclones arose from a single initiating cell (as evidenced by truncal mutations and copy number aberrations), and one or more clones were able to invade through the basement membrane (“multiclonal invasion model”). Applying similar single-cell spatial transcriptomics and genomics to HNSCC is likely to provide further insights into invasion and metastasis (see “Outlook” below).

### Therapeutic response and resistance

Despite aggressive multi-modality therapy, a significant proportion of patients with HNSCC will develop disease recurrence, with up to 60% of patients experiencing locoregional failure and up to 30% of patients with distant failure [106]. Despite promising therapeutic responses to ICI in a subset of patients with persistent or recurrent disease, the prognosis for remainder of these patients remains poor [104,107]. Our understanding of the molecular and genetic underpinnings for treatment failure and the ability to predict patient response to therapy remains crude at best [87].

For treatment-resistant tumors, the question arises whether malignant cells acquired mutations that enabled survival, or whether there always existed a subpopulation of cells that

survived through the selective pressure of therapy. Outside of HNSCC, such efforts are just beginning. Kim *et al.* applied single-cell DNA and RNA sequencing with bulk exome sequencing to profile longitudinal triple-negative breast cancer (TNBC) samples from patients undergoing neoadjuvant therapy [75]. Among their 20 patients, ten had a complete response to therapy, while the remaining patients failed treatment. Eight samples – four from the clonal-extinction and four from the clonal-persistence groups – were analyzed at the single-cell level. Among the treatment resistant patients, they found that treatment resistant genotypes were present, though rare, in the pre-treatment samples and were selected for by the evolutionary pressure of neoadjuvant therapy. Interestingly, the transcriptional profiles of these samples changed throughout the course of treatment, suggesting that both adaptive and acquired mechanisms of therapeutic resistance occurred concurrently. Similar studies would be of great value in HNSCC but would require the optimization of single-nuclei RNA sequencing (sNuc-seq) to enable matched analyses of archived, frozen tissue.

Understanding the biology of ICI is another area of particular interest in head and neck oncology because responses to ICI are durable but highly variable and difficult to predict [106,107]. Recent data from melanoma suggests that there may be unique malignant cell signatures that define response to ICI [72], a provocative possibility if extended to HNSCC. In advanced melanoma, Sade-Feldman *et al.* used scRNA-seq to analyze sub-populations of T-cells [73]. Two main clusters of T-cells were seen in all 48 patient samples: One cluster linked to memory, activation, and cell survival (“CD8\_G”), and the other with genes linked to cell exhaustion (“CD8\_B”). Tumors with higher proportions of CD8\_G cells were linked with treatment response, while those enriched in CD8\_B cells often did not respond. The transcription factor TCF7 was among the chief markers predictive of a good clinical response. Using immunofluorescence, they stained an independent cohort of melanoma tissue samples treated with anti-PD-1 therapy and found significantly higher expression of TCF7 in treatment responders versus non-responders.

To date, no similar multi-omics study has been performed for patients with HNSCC. Such studies could help optimize patient selection for systemic chemotherapy or immunotherapy. An important component of these analyses will be analyzing changes in *intra*-tumoral heterogeneity with exogenous agents such as various forms of chemotherapy (e.g. cisplatin), biologic therapies (e.g. cetuximab), radiation, and ICI, which will require the use of samples from ongoing clinical trials as well as a broad array of model systems. The improved ability to multiplex samples for single-cell analyses [104,108–111] has opened new, affordable avenues for high throughput single-cell analyses of a range of conditions, with some techniques now optimized for 96 conditions in a singular sequencing lane. Applying these treatment-related perturbations to models such as HNSCC cell lines, cell line orospheres, patient-derived xenografts (PDX), and PDX-derived organoids (PDX-O) will be essential to help define a baseline and investigate its changes with treatment.

## Outlook

### A single-cell head and neck tumor atlas

TCGA was an unprecedented endeavor, managing to collect thousands of tumor and normal tissue samples from patients with 33 types of malignancies at multiple collaborating

institutions as well as providing publicly available libraries of genomic, transcriptomic, and epigenomic data to the research community [112]. These efforts vastly expanded our collective understanding on a number of human malignancies and have helped to move research forward on basic, translational, and clinical fronts. To date, our study represents the only published single-cell dataset in head and neck cancer research [92]. Ongoing and future single-cell studies must explore HNSCC at subsites other than the oral cavity, as well as other primary tumor types, including salivary gland and endocrine malignancies. Preliminary observations from single-cell analyses of adenoid cystic and mucocystic carcinoma are likely to represent the forefront of these efforts and will contribute to a single-cell atlas of head and neck malignancies that should rapidly accelerate major lines of investigation.

### Spatial single-cell sequencing

Although SCS methods have begun to define genetic and transcriptional heterogeneity in cancers, commonly utilized methods notably suffer from the loss of spatial information. Recent single-cell sequencing studies have described malignant, stromal, and immune subpopulations with distinct spatial localization, such as p-EMT cells localized at the tumor-stromal interface in HNSCC [92] and subpopulations of perivascular, mammary fat pad, and epithelial CAFs with variable prognostic significance in breast cancer [113]. Studies have also hypothesized that intercellular interactions may drive invasion, metastasis, and immune exhaustion, such as interactions between CAFs and cancer cells in HNSCC and pancreatic ductal adenocarcinoma [114], and relationships between T cells and cancer cells in breast cancer [115]. These findings highlight the need to utilize SCS methods with preserved spatial information, as we develop a more nuanced understanding of the role of intercellular interactions in cancer biology.

In this vein, a few techniques have been developed but remain nascent for single-cell profiling of human tumor specimens. Fluorescent *in situ* sequencing (FISSEQ) utilizes 30-base reads from more than 8,000 genes to detect *in situ* expression patterns [116], though it may suffer from low read counts per cell and is predominantly useful for characterizing cell types. LCM coupled with Smart-seq2 has also been described [9] but is low throughput and labor intensive. Further optimization of these methods would be useful in utilizing spatial SCS in a hypothesis-generating manner.

Conversely, techniques for validating sequencing data using highly multiplexed *in situ* detection of markers have been much more widely utilized and, for the moment, may represent more practical approaches to translating single-cell sequencing data into clinically meaningful information. Highly multiplexed IHC-and immunofluorescence (IF)-based assays with image analysis that offers cellular resolution have been described for up to 61 markers in FFPE tissue [117]. Perkin Elmer's commercially available Vectra multispectral imaging system with the Opal staining kit offers a more accessible approach to seven-marker staining in FFPE sections, with the ability to computationally merge serial sections and quantify up to 14 markers [118]. The latter approach has been utilized to explore subpopulations of tumor infiltrating immune cells [119–122] and delineate tumor-immune cell interactions [123,124] in multiple tumor types, including ongoing work in HNSCC in

our group. An alternative to staining-based approaches is the use of mass cytometry at the cellular level (CyTOF), allowing the simultaneous quantification of more than 30 analytes at subcellular resolution [7]. In solid tumors, CyTOF has been used for immune cell profiling in colon cancer [125] and tumor cell profiling in pancreatic and prostate cancers [114,126]. Regardless of approach, these validation techniques have the advantage of being extremely high throughput, as compared with existing single-cell sequencing techniques, with the ability to profile hundreds of thousands of cells at a time.

### Temporal single-cell sequencing

Understanding the temporal progression of tumor heterogeneity, particularly with tumor recurrence or in response to systemic therapy, will provide another critical layer of insights into cancer biology and is increasingly accessible with decreasing sequencing costs and increasing throughput. True serial sequencing of individual patient samples has not been reported in any solid tumors beyond cutaneous malignancies. This approach is particularly powerful as biologic corollaries of clinical trials become more standardized. In particular, serial sequencing of individual patients pre-and post-therapy should be an expected correlate in all interventional clinical trials, providing important insights into predictors of treatment response and mechanisms of resistance. Current work is ongoing in HNSCC to investigate the immune landscape of tumors pre-and post-ICI therapy, but extending this approach to additional biologic and systemic agents will be important.

### Single-cell multi-omics

Despite the development of individual single-cell genomic, epigenomic, and transcriptomic profiling technologies, the ability to simultaneously assess single-cell DNA, RNA, and epigenetic changes remains limited due to challenges with separation and unbiased amplification of small quantities of individual genetic material in a cell. A number of single-cell multi-omic techniques have been described, including those that simultaneously sequence the genome and transcriptome [19,127], those that sequence the transcriptome and the epigenome [128–130], and those that attempt to do all three [131]. However, the large-scale application of these techniques to human tumor specimens has been limited and would provide an important layer of insight into the genomic and epigenomic contributions to expression heterogeneity. To date, most studies have been limited to analyses of inferred CNVs based on scRNA-seq data [74,92,132]. Moudgil *et al.* recently described the use of self-reporting transposons in single-cell “calling cards” as an assay for simultaneously capturing expression profiles and mapping transcription factor binding sites in single-cells [133], and this approach will certainly be of interest in HNSCC and oncology more broadly.

### Models of tumor heterogeneity

As *in vivo* tumor heterogeneity is better understood at a cellular level, it is increasingly important that tumor models accurately capture and reflect this heterogeneity. Accordingly, comprehensive SCS of cell lines may help to better delineate which cell lines recapitulate aspects of *in vivo* tumor heterogeneity. In our experience with oral cavity HNSCC cell lines, some reflect *in vivo* heterogeneity, while many are transcriptionally homogeneous, and yet others have heterogeneity but do not meaningfully mimic *in vivo* expression states [92]. Novel techniques such as cell hashing with barcoded antibodies [111] may enable

multiplexing for high throughput approaches to accomplishing these goals. Similarly, although PDX and PDX-O models exist in HNSCC, it remains to be seen whether these models faithfully recapitulate *in vivo* heterogeneity and particularly whether, after multiple passages, this heterogeneity is maintained. More comprehensively characterizing these models will be critical to determining which models are useful in future studies of *intra*-tumoral heterogeneity. If these models maintain tumor heterogeneity, then PDX or PDXO may provide a patient-specific “reactor” to test the efficacy of systemic therapies before they are utilized in a patient, eliminating the need for a trial-and-error approach in salvage cases.

## Conclusions

Since the role of tumor heterogeneity in prognosis and treatment resistance has become clear [90], there has been a flurry of work to better define heterogeneity at a cellular resolution. Of the variety of SCS technologies now available, single-cell transcriptomics is the best developed and most widely used to study human tumors, with novel insights to date into malignant, immune, and stromal cell subpopulations, as well as intercellular interactions that promote invasion and metastasis. Novel technologies in scRNA-seq have focused on higher throughput, lower cost methods that will enable more widespread use of these technologies, particularly as biological corollaries of clinical trials. Indeed, as the cost and accessibility of SCS further improves, it is not hard to imagine biopsy specimens being “scored” for a panel of genetic states or transcriptional programs (e.g. p-EMT), providing a detailed, molecularly-based predictor of adverse biologic features that drives clinical decision-making.

Conceptually, the widespread use of scRNA-seq has been hypothesis-generating, and a variety of approaches will be necessary to develop a better mechanistic understanding of the phenotypes that have been defined. This includes further development of single-cell genomic, epigenomic, and multi-omic sequencing technologies, as well as SCS technologies that retain spatial information. It also involves the better characterization of existing tumor models, including cell lines and PDX models, potentially by SCS, in order to determine which models truly capture the heterogeneity seen *in vivo*. While single-cell studies have uncovered many new layers and hypotheses in cancer biology, these findings also highlight the complexity and nuance of the tumor ecosystem and emphasize the significant work to be done.

## Acknowledgements

This work was supported by K08CA237732 (S.V.P) from the National Cancer Institute.

## References

- [1]. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries n.d. <https://www.ncbi.nlm.nih.gov/pubmed/30207593> (accessed July 26, 2019).
- [2]. Marur S, Forastiere AA. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clin Proc* 2016;91:386–96. doi:10.1016/j.mayocp.2015.12.017. [PubMed: 26944243]

- [3]. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015;517:576–82. doi:10.1038/nature14129. [PubMed: 25631445]
- [4]. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The mutational landscape of head and neck squamous cell carcinoma. *Science* 2011;333:1157–60. doi:10.1126/science.1208130. [PubMed: 21798893]
- [5]. Ren X, Kang B, Zhang Z. Understanding tumor ecosystems by single-cell sequencing: promises and limitations. *Genome Biol* 2018;19:211. doi:10.1186/s13059-018-1593-z. [PubMed: 30509292]
- [6]. Peltanova B, Raudenska M, Masarik M. Effect of tumor microenvironment on pathogenesis of the head and neck squamous cell carcinoma: a systematic review. *Mol Cancer* 2019;18:63. doi: 10.1186/s12943-019-0983-5. [PubMed: 30927923]
- [7]. Gross A, Schoendube J, Zimmermann S, Steeb M, Zengerle R, Koltay P. Technologies for Single-Cell Isolation. *Int J Mol Sci* 2015;16:16897–919. doi:10.3390/ijms160816897. [PubMed: 26213926]
- [8]. Guo F, Li L, Li J, Wu X, Hu B, Zhu P, et al. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res* 2017;27:967–88. doi:10.1038/cr.2017.82. [PubMed: 28621329]
- [9]. Nichterwitz S, Chen G, Aguila Benitez J, Yilmaz M, Storvall H, Cao M, et al. Laser capture microscopy coupled with Smart-seq2 for precise spatial transcriptomic profiling. *Nat Commun* 2016;7:12139. doi:10.1038/ncomms12139. [PubMed: 27387371]
- [10]. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* 2013;14:618–30. doi:10.1038/nrg3542. [PubMed: 23897237]
- [11]. Whitesides GM. The origins and the future of microfluidics. *Nature* 2006;442:368–73. doi: 10.1038/nature05058. [PubMed: 16871203]
- [12]. Fan HC, Fu GK, Fodor SPA. Combinatorial labeling of single cells for gene expression cytometry. *Science* 2015;347:1258367. doi:10.1126/science.1258367. [PubMed: 25657253]
- [13]. Gierahn TM, Wadsworth li MH, Hughes TK, Bryson BD, Butler A, Satija R, et al. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods* 2017;14:395–8. doi:10.1038/nmeth.4179. [PubMed: 28192419]
- [14]. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding n.d. <https://science.sciencemag.org/content/360/6385/176> (accessed July 26, 2019).
- [15]. Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 2017;357:661–7. doi:10.1126/science.aam8940. [PubMed: 28818938]
- [16]. Xin Y, Kim J, Ni M, Wei Y, Okamoto H, Lee J, 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;113:3293–8. doi:10.1073/pnas.1602306113. [PubMed: 26951663]
- [17]. Arneson N, Hughes S, Houlston R, Done S. Whole-Genome Amplification by Degenerate Oligonucleotide Primed PCR (DOP-PCR). *Cold Spring Harb Protoc* 2008;2008:pdb.prot4919. doi:10.1101/pdb.prot4919.
- [18]. Spits C, Le Caignec C, De Rycke M, Van Haute L, Van Steirteghem A, Liebaers I, et al. Whole-genome multiple displacement amplification from single cells. *Nat Protoc* 2006;1:1965–70. doi: 10.1038/nprot.2006.326. [PubMed: 17487184]
- [19]. Macaulay IC, Haerty W, Kumar P, Li YI, Hu TX, Teng MJ, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* 2015;12:519–22. doi:10.1038/nmeth.3370. [PubMed: 25915121]
- [20]. Zong C, Lu S, Chapman AR, Xie XS. Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. *Science* 2012;338:1622–6. doi:10.1126/science.1229164. [PubMed: 23258894]
- [21]. Paolillo C, Mu Z, Rossi G, Schiewer MJ, Nguyen T, Austin L, et al. Detection of Activating Estrogen Receptor Gene (ESR1) Mutations in Single Circulating Tumor Cells. *Clin Cancer Res Off J Am Assoc Cancer Res* 2017;23:6086–93. doi:10.1158/1078-0432.CCR-17-1173.

- [22]. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;6:377–82. doi:10.1038/nmeth.1315. [PubMed: 19349980]
- [23]. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol* 2016;17:77. doi:10.1186/s13059-016-0938-8. [PubMed: 27121950]
- [24]. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Rep* 2012;2:666–73. doi:10.1016/j.celrep.2012.08.003. [PubMed: 22939981]
- [25]. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types. *Science* 2014;343:776–9. doi:10.1126/science.1247651. [PubMed: 24531970]
- [26]. Islam S, Kjällquist U, Moliner A, Zajac P, Fan J-B, Lönnerberg P, et al. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc* 2012;7:813–28. doi:10.1038/nprot.2012.022. [PubMed: 22481528]
- [27]. Islam S, Kjällquist U, Moliner A, Zajac P, Fan J-B, Lönnerberg P, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 2011;21:1160–7. doi:10.1101/gr.110882.110. [PubMed: 21543516]
- [28]. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* 2014;9:171–81. doi:10.1038/nprot.2014.006. [PubMed: 24385147]
- [29]. Ramsköld D, Luo S, Wang Y-C, Li R, Deng Q, Faridani OR, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol* 2012;30:777–82. doi:10.1038/nbt.2282. [PubMed: 22820318]
- [30]. The External RNA Controls Consortium: a progress report. *Nat Methods* 2005;2:731–4. doi:10.1038/nmeth1005-731. [PubMed: 16179916]
- [31]. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* 2014;11:163–6. doi:10.1038/nmeth.2772. [PubMed: 24363023]
- [32]. Margueron R, Reinberg D. Chromatin Structure and the Inheritance of Epigenetic Information. *Nat Rev Genet* 2010;11:285–96. doi:10.1038/nrg2752. [PubMed: 20300089]
- [33]. Guo H, Zhu P, Guo F, Li X, Wu X, Fan X, et al. Profiling DNA methylome landscapes of mammalian cells with single-cell reduced-representation bisulfite sequencing. *Nat Protoc* 2015;10:645–59. doi:10.1038/nprot.2015.039. [PubMed: 25837417]
- [34]. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat Methods* 2014;11:817–20. doi:10.1038/nmeth.3035. [PubMed: 25042786]
- [35]. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schönegger A, Klughammer J, et al. Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. *Cell Rep* 2015;10:1386–97. doi:10.1016/j.celrep.2015.02.001. [PubMed: 25732828]
- [36]. Clark SJ, Smallwood SA, Lee HJ, Krueger F, Reik W, Kelsey G. Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat Protoc* 2017;12:534–47. doi:10.1038/nprot.2016.187. [PubMed: 28182018]
- [37]. Han L, Wu H-J, Zhu H, Kim K-Y, Marjani SL, Riester M, et al. Bisulfite-independent analysis of CpG island methylation enables genome-scale stratification of single cells. *Nucleic Acids Res* 2017;45:e77. doi:10.1093/nar/gkx026. [PubMed: 28126923]
- [38]. Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 2015;33:1165–72. doi:10.1038/nbt.3383. [PubMed: 26458175]
- [39]. Harada A, Maehara K, Handa T, Arimura Y, Nogami J, Hayashi-Takanaka Y, et al. A chromatin integration labelling method enables epigenomic profiling with lower input. *Nat Cell Biol* 2019;21:287. doi:10.1038/s41556-018-0248-3. [PubMed: 30532068]



- [40]. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat Commun* 2019;10:1930. doi: 10.1038/s41467-019-09982-5. [PubMed: 31036827]
- [41]. Ku WL, Nakamura K, Gao W, Cui K, Hu G, Tang Q, et al. Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. *Nat Methods* 2019;16:323. doi:10.1038/s41592-019-0361-7. [PubMed: 30923384]
- [42]. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, et al. Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 2015;523:486–90. doi: 10.1038/nature14590. [PubMed: 26083756]
- [43]. Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 2015;348:910–4. doi:10.1126/science.aab1601. [PubMed: 25953818]
- [44]. Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain | *Nature Biotechnology* n.d. <https://www.nature.com/articles/nbt.4038> (accessed August 24, 2019).
- [45]. Yoshida H, Lareau CA, Ramirez RN, Rose SA, Maier B, Wroblewska A, et al. The cis-Regulatory Atlas of the Mouse Immune System. *Cell* 2019;176:897–912.e20. doi:10.1016/j.cell.2018.12.036. [PubMed: 30686579]
- [46]. Gawad C, Koh W, Quake SR. Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. *Proc Natl Acad Sci* 2014;111:17947–52. doi:10.1073/pnas.1420822111. [PubMed: 25425670]
- [47]. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 2014;512:155–60. doi:10.1038/nature13600. [PubMed: 25079324]
- [48]. Nikolenko SI, Korobeynikov AI, Alekseyev MA. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 2013;14 Suppl 1:S7. doi: 10.1186/1471-2164-14-S1-S7.
- [49]. Baslan T, Kendall J, Rodgers L, Cox H, Riggs M, Stepansky A, et al. Genome-wide copy number analysis of single cells. *Nat Protoc* 2012;7:1024–41. doi:10.1038/nprot.2012.039. [PubMed: 22555242]
- [50]. Wang R, Lin D-Y, Jiang Y. SCOPE: a normalization and copy number estimation method for single-cell DNA sequencing. *BioRxiv* 2019:594267. doi:10.1101/594267.
- [51]. Single-cell genome sequencing: current state of the science | *Nature Reviews Genetics* n.d. <https://www.nature.com/articles/nrg.2015.16> (accessed August 4, 2019).
- [52]. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* 2017;27:491–9. doi:10.1101/gr.209601.116. [PubMed: 28100584]
- [53]. Ilicic T, Kim JK, Kolodziejczyk AA, Bagger FO, McCarthy DJ, Marioni JC, et al. Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* 2016;17:29. doi:10.1186/s13059-016-0888-1. [PubMed: 26887813]
- [54]. Bacher R, Kendzierski C. Design and computational analysis of single-cell RNA-sequencing experiments. *Genome Biol* 2016;17:63. doi:10.1186/s13059-016-0927-y. [PubMed: 27052890]
- [55]. Haghverdi L, Lun ATL, Morgan MD, Marioni JC. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* 2018;36:421–7. doi: 10.1038/nbt.4091. [PubMed: 29608177]
- [56]. Büttner M, Miao Z, Wolf FA, Teichmann SA, Theis FJ. A test metric for assessing single-cell RNA-seq batch correction. *Nat Methods* 2019;16:43–9. doi:10.1038/s41592-018-0254-1. [PubMed: 30573817]
- [57]. Bacher R, Chu L-F, Leng N, Gasch AP, Thomson JA, Stewart RM, et al. SCnorm: robust normalization of single-cell RNA-seq data. *Nat Methods* 2017;14:584–6. doi:10.1038/nmeth.4263. [PubMed: 28418000]
- [58]. Huang M, Wang J, Torre E, Dueck H, Shaffer S, Bonasio R, et al. SAVER: gene expression recovery for single-cell RNA sequencing. *Nat Methods* 2018;15:539–42. doi:10.1038/s41592-018-0033-z. [PubMed: 29941873]

- [59]. van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, et al. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* 2018;174:716–729.e27. doi:10.1016/j.cell.2018.05.061. [PubMed: 29961576]
- [60]. Li WV, Li JJ. An accurate and robust imputation method scImpute for single-cell RNA-seq data. *Nat Commun* 2018;9:997. doi:10.1038/s41467-018-03405-7. [PubMed: 29520097]
- [61]. Gong W, Kwak I-Y, Pota P, Koyano-Nakagawa N, Garry DJ. DrImpute: imputing dropout events in single cell RNA sequencing data. *BMC Bioinformatics* 2018;19:220. doi:10.1186/s12859-018-2226-y. [PubMed: 29884114]
- [62]. AutoImpute: Autoencoder based imputation of single-cell RNA-seq data | Scientific Reports n.d. <https://www.nature.com/articles/s41598-018-34688-x> (accessed August 4, 2019).
- [63]. Chen G, Ning B, Shi T. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Front Genet* 2019;10:317. doi:10.3389/fgene.2019.00317. [PubMed: 31024627]
- [64]. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 2017;14:979–82. doi:10.1038/nmeth.4402. [PubMed: 28825705]
- [65]. Lummertz da Rocha E, Rowe RG, Lundin V, Malleshaiah M, Jha DK, Rambo CR, et al. Reconstruction of complex single-cell trajectories using CellRouter. *Nat Commun* 2018;9:892. doi:10.1038/s41467-018-03214-y. [PubMed: 29497036]
- [66]. Miao Z, Deng K, Wang X, Zhang X. DEsingle for detecting three types of differential expression in single-cell RNA-seq data. *Bioinforma Oxf Engl* 2018;34:3223–4. doi:10.1093/bioinformatics/bty332.
- [67]. Qiu X, Hill A, Packer J, Lin D, Ma Y-A, Trapnell C. Single-cell mRNA quantification and differential analysis with Censur. *Nat Methods* 2017;14:309–15. doi:10.1038/nmeth.4150. [PubMed: 28114287]
- [68]. Chen L, Zheng S. BCseq: accurate single cell RNA-seq quantification with bias correction. *Nucleic Acids Res* 2018;46:e82. doi:10.1093/nar/gky308. [PubMed: 29718338]
- [69]. Davis-Marcisak EF, Orugunta P, Stein-O'Brien G, Puram SV, Torres ER, Hopkins A, et al. Expression variation analysis for tumor heterogeneity in single-cell RNA-sequencing data. *BioRxiv* 2018:479287. doi:10.1101/479287.
- [70]. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods* 2017;14:1083–6. doi:10.1038/nmeth.4463. [PubMed: 28991892]
- [71]. Chan TE, Stumpf MPH, Babbitt AC. Gene Regulatory Network Inference from Single-Cell Data Using Multivariate Information Measures. *Cell Syst* 2017;5:251–267.e3. doi:10.1016/j.cels.2017.08.014. [PubMed: 28957658]
- [72]. Jerby-Arnon L, Shah P, Cuoco MS, Rodman C, Su M-J, Melms JC, et al. A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. *Cell* 2018;175:984–997.e24. doi:10.1016/j.cell.2018.09.006. [PubMed: 30388455]
- [73]. Sade-Feldman M, Yizhak K, Bjorgaard SL, Ray JP, de Boer CG, Jenkins RW, et al. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell* 2018;175:998–1013.e20. doi:10.1016/j.cell.2018.10.038. [PubMed: 30388456]
- [74]. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016;352:189–96. doi:10.1126/science.aad0501. [PubMed: 27124452]
- [75]. Kim C, Gao R, Sei E, Brandt R, Hartman J, Hatschek T, et al. Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell* 2018;173:879–893.e13. doi:10.1016/j.cell.2018.03.041. [PubMed: 29681456]
- [76]. Casasent AK, Schalck A, Gao R, Sei E, Long A, Pangburn W, et al. Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* 2018;172:205–217.e12. doi:10.1016/j.cell.2017.12.007. [PubMed: 29307488]
- [77]. Gao R, Kim C, Sei E, Foukakis T, Crosetto N, Chan L-K, et al. Nanogrid single-nucleus RNA sequencing reveals phenotypic diversity in breast cancer. *Nat Commun* 2017;8:228. doi:10.1038/s41467-017-00244-w. [PubMed: 28794488]

- [78]. Gerdes MJ, Gökmen-Polar Y, Sui Y, Pang AS, LaPlante N, Harris AL, et al. Single-cell heterogeneity in ductal carcinoma in situ of breast. *Mod Pathol* 2018;31:406–17. doi:10.1038/modpathol.2017.143. [PubMed: 29148540]
- [79]. Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* 2018;174:1293–1308.e36. doi:10.1016/j.cell.2018.05.060. [PubMed: 29961579]
- [80]. Chung W, Eum HH, Lee H-O, Lee K-M, Lee H-B, Kim K-T, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun* 2017;8:15081. doi:10.1038/ncomms15081. [PubMed: 28474673]
- [81]. Guo X, Zhang Y, Zheng L, Zheng C, Song J, Zhang Q, et al. Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* 2018;24:978–85. doi:10.1038/s41591-018-0045-3. [PubMed: 29942094]
- [82]. Min J-W, Kim WJ, Han JA, Jung Y-J, Kim K-T, Park W-Y, et al. Identification of Distinct Tumor Subpopulations in Lung Adenocarcinoma via Single-Cell RNA-seq. *PLOS ONE* 2015;10:e0135817. doi:10.1371/journal.pone.0135817. [PubMed: 26305796]
- [83]. Kim K-T, Lee HW, Lee H-O, Kim SC, Seo YJ, Chung W, et al. Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. *Genome Biol* 2015;16:127. doi:10.1186/s13059-015-0692-3. [PubMed: 26084335]
- [84]. Roerink SF, Sasaki N, Lee-Six H, Young MD, Alexandrov LB, Behjati S, et al. Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature* 2018;556:457–62. doi:10.1038/s41586-018-0024-3. [PubMed: 29643510]
- [85]. Li H, Courtois ET, Sengupta D, Tan Y, Chen KH, Goh JLL, et al. Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* 2017;49:708–18. doi:10.1038/ng.3818. [PubMed: 28319088]
- [86]. Leung ML, Davis A, Gao R, Casasent A, Wang Y, Sei E, et al. Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res* 2017;27:1287–99. doi:10.1101/gr.209973.116. [PubMed: 28546418]
- [87]. Puram SV, Rocco JW. Molecular Aspects of Head and Neck Cancer Therapy. *Hematol Oncol Clin North Am* 2015;29:971–92. doi:10.1016/j.hoc.2015.07.003. [PubMed: 26568543]
- [88]. Hinohara K, Polyak K. Intratumoral Heterogeneity: More Than Just Mutations. *Trends Cell Biol* 2019;29:569–79. doi:10.1016/j.tcb.2019.03.003. [PubMed: 30987806]
- [89]. Levitin HM, Yuan J, Sims PA. Single-Cell Transcriptomic Analysis of Tumor Heterogeneity. *Trends Cancer* 2018;4:264–8. doi:10.1016/j.trecan.2018.02.003. [PubMed: 29606308]
- [90]. Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol* 2013;49:211–5. doi:10.1016/j.oraloncology.2012.09.007. [PubMed: 23079694]
- [91]. Mroz EA, Tward AD, Pickering CR, Myers JN, Ferris RL, Rocco JW. High intratumor genetic heterogeneity is related to worse outcome in patients with head and neck squamous cell carcinoma: Genetic Heterogeneity and HNSCC Outcome. *Cancer* 2013;119:3034–42. doi:10.1002/cncr.28150. [PubMed: 23696076]
- [92]. Puram SV, Tirosch I, Parikh AS, Patel AP, Yizhak K, Gillespie S, et al. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* 2017;171:1611–1624.e24. doi:10.1016/j.cell.2017.10.044. [PubMed: 29198524]
- [93]. The Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015;517:576–82. doi:10.1038/nature14129. [PubMed: 25631445]
- [94]. Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell* 2019;178:835–849.e21. doi:10.1016/j.cell.2019.06.024. [PubMed: 31327527]
- [95]. Darmanis S, Sloan SA, Croote D, Mignardi M, Chernikova S, Samghababi P, et al. Single-Cell RNA-Seq Analysis of Infiltrating Neoplastic Cells at the Migrating Front of Human Glioblastoma. *Cell Rep* 2017;21:1399–410. doi:10.1016/j.celrep.2017.10.030. [PubMed: 29091775]

- [96]. Li H, van der Leun AM, Yofe I, Lubling Y, Gelbard-Solodkin D, van Akkooi ACJ, et al. Dysfunctional CD8 T Cells Form a Proliferative, Dynamically Regulated Compartment within Human Melanoma. *Cell* 2019;176:775–789.e18. doi:10.1016/j.cell.2018.11.043. [PubMed: 30595452]
- [97]. De Simone M, Rossetti G, Pagani M. Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges. *Front Immunol* 2018;9:1638. doi:10.3389/fimmu.2018.01638. [PubMed: 30072991]
- [98]. Bentzen AK, Marquard AM, Lyngaa R, Saini SK, Ramskov S, Donia M, et al. Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat Biotechnol* 2016;34:1037–45. doi:10.1038/nbt.3662. [PubMed: 27571370]
- [99]. Yost KE, Satpathy AT, Wells DK, Qi Y, Wang C, Kageyama R, et al. Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat Med* 2019;25:1251–9. doi:10.1038/s41591-019-0522-3. [PubMed: 31359002]
- [100]. Kalluri R The biology and function of fibroblasts in cancer. *Nat Rev Cancer* 2016;16:582–98. doi:10.1038/nrc.2016.73. [PubMed: 27550820]
- [101]. Gieniec KA, Butler LM, Worthley DL, Woods SL. Cancer-associated fibroblasts—heroes or villains? *Br J Cancer* 2019;121:293–302. doi:10.1038/s41416-019-0509-3. [PubMed: 31289350]
- [102]. Kraman M, Bambrough PJ, Arnold JN, Roberts EW, Magiera L, Jones JO, et al. Suppression of Antitumor Immunity by Stromal Cells Expressing Fibroblast Activation Protein— a 2010;330:5.
- [103]. Lambert AW, Pattabiraman DR, Weinberg RA. Emerging Biological Principles of Metastasis. *Cell* 2017;168:670–91. doi:10.1016/j.cell.2016.11.037. [PubMed: 28187288]
- [104]. Cohen EEW, Soulières D, Le Tourneau C, Dinis J, Licitra L, Ahn M-J, et al. Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *The Lancet* 2019;393:156–67. doi:10.1016/S0140-6736(18)31999–8.
- [105]. Puram SV, Parikh AS, Tirosh I. Single cell RNA-seq highlights a role for a partial EMT in head and neck cancer. *Mol Cell Oncol* 2018;5:e1448244. doi:10.1080/23723556.2018.1448244. [PubMed: 30250901]
- [106]. Sacco AG, Cohen EE. Current Treatment Options for Recurrent or Metastatic Head and Neck Squamous Cell Carcinoma. *J Clin Oncol* 2015;33:3305–13. doi:10.1200/JCO.2015.62.0963. [PubMed: 26351341]
- [107]. Ferris RL, Blumenschein G, Fayette J, Guigay J, Colevas AD, Licitra L, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *N Engl J Med* 2016;375:1856–67. doi:10.1056/NEJMoa1602252. [PubMed: 27718784]
- [108]. Gehring J, Park JH, Chen S, Thomson M, Pachter L. Highly Multiplexed Single-Cell RNA-seq for Defining Cell Population and Transcriptional Spaces. *Genomics*; 2018. doi:10.1101/315333.
- [109]. McGinnis CS, Patterson DM, Winkler J, Conrad DN, Hein MY, Srivastava V, et al. MULTI-seq: sample multiplexing for single-cell RNA sequencing using lipid-tagged indices. *Nat Methods* 2019;16:619–26. doi:10.1038/s41592-019-0433-8. [PubMed: 31209384]
- [110]. Gaublotte JT, Li B, McCabe C, Knecht A, Yang Y, Drokhlyansky E, et al. Nuclei multiplexing with barcoded antibodies for single-nucleus genomics. *Nat Commun* 2019;10:2907. doi:10.1038/s41467-019-10756-2. [PubMed: 31266958]
- [111]. Stoeckius M, Zheng S, Houck-Loomis B, Hao S, Yeung BZ, Mauck WM, et al. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol* 2018;19:224. doi:10.1186/s13059-018-1603-1. [PubMed: 30567574]
- [112]. Mathe EA, Davis S, editors. Statistical genomics: methods and protocols. New York: Humana Press; 2016.
- [113]. Bartoschek M, Oskolkov N, Bocci M, Lövrot J, Larsson C, Sommarin M, et al. Spatially and functionally distinct subclasses of breast cancer-associated fibroblasts revealed by single cell RNA sequencing. *Nat Commun* 2018;9:5150. doi:10.1038/s41467-018-07582-3. [PubMed: 30514914]
- [114]. Ligorio M, Sil S, Malagon-Lopez J, Nieman LT, Misale S, Di Pilato M, et al. Stromal Microenvironment Shapes the Intratumoral Architecture of Pancreatic Cancer. *Cell* 2019;178:160–175.e27. doi:10.1016/j.cell.2019.05.012. [PubMed: 31155233]

- [115]. Wagner J, Rapsomaniki MA, Chevrier S, Anzeneder T, Langwieder C, Dykgers A, et al. A Single-Cell Atlas of the Tumor and Immune Ecosystem of Human Breast Cancer. *Cell* 2019;177:1330–1345.e18. doi:10.1016/j.cell.2019.03.005. [PubMed: 30982598]
- [116]. Lee JH, Daugharthy ER, Scheiman J, Kalhor R, Yang JL, Ferrante TC, et al. Highly Multiplexed Subcellular RNA Sequencing in Situ. *Science* 2014;343:1360–3. doi:10.1126/science.1250212. [PubMed: 24578530]
- [117]. Gerdes MJ, Sevinsky CJ, Sood A, Adak S, Bello MO, Bordwell A, et al. Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc Natl Acad Sci* 2013;110:11982–7. doi:10.1073/pnas.1300136110. [PubMed: 23818604]
- [118]. Parra ER, Uraoka N, Jiang M, Cook P, Gibbons D, Forget M-A, et al. Validation of multiplex immunofluorescence panels using multispectral microscopy for immune-profiling of formalin-fixed and paraffin-embedded human tumor tissues. *Sci Rep* 2017;7:13380. doi:10.1038/s41598-017-13942-8. [PubMed: 29042640]
- [119]. Granier C, Dariane C, Combe P, Verkarre V, Urien S, Badoual C, et al. Tim-3 Expression on Tumor-Infiltrating PD-1 + CD8 + T Cells Correlates with Poor Clinical Outcome in Renal Cell Carcinoma. *Cancer Res* 2017;77:1075–82. doi:10.1158/0008-5472.CAN-16-0274. [PubMed: 27872087]
- [120]. Tsujikawa T, Kumar S, Borkar RN, Azimi V, Thibault G, Chang YH, et al. Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis. *Cell Rep* 2017;19:203–17. doi:10.1016/j.celrep.2017.03.037. [PubMed: 28380359]
- [121]. Vasaturo A, Di Blasio S, Verweij D, Blokx WAM, van Krieken JH, de Vries IJM, et al. Multispectral imaging for highly accurate analysis of tumour-infiltrating lymphocytes in primary melanoma. *Histopathology* 2017;70:643–9. doi:10.1111/his.13070. [PubMed: 27571246]
- [122]. Takahashi D, Kojima M, Suzuki T, Sugimoto M, Kobayashi S, Takahashi S, et al. Profiling the Tumour Immune Microenvironment in Pancreatic Neuroendocrine Neoplasms with Multispectral Imaging Indicates Distinct Subpopulation Characteristics Concordant with WHO 2017 Classification. *Sci Rep* 2018;8:13166. doi:10.1038/s41598-018-31383-9. [PubMed: 30177687]
- [123]. Barua S, Fang P, Sharma A, Fujimoto J, Wistuba I, Rao AUK, et al. Spatial interaction of tumor cells and regulatory T cells correlates with survival in non-small cell lung cancer. *Lung Cancer* 2018;117:73–9. doi:10.1016/j.lungcan.2018.01.022. [PubMed: 29409671]
- [124]. Granier C, Vinatier E, Colin E, Mandavit M, Dariane C, Verkarre V, et al. Multiplexed Immunofluorescence Analysis and Quantification of Intratumoral PD-1+ Tim-3+ CD8+ T Cells. *J Vis Exp* 2018:56606. doi:10.3791/56606.
- [125]. Zhang T, Lv J, Tan Z, Wang B, Warden AR, Li Y, et al. Immunocyte Profiling Using Single-Cell Mass Cytometry Reveals EpCAM+ CD4+ T Cells Abnormal in Colon Cancer. *Front Immunol* 2019;10:1571. doi:10.3389/fimmu.2019.01571. [PubMed: 31354723]
- [126]. Sehgal PD, Bauman TM, Nicholson TM, Vellky JE, Ricke EA, Tang W, et al. Tissue-specific quantification and localization of androgen and estrogen receptors in prostate cancer. *Hum Pathol* 2019;89:99–108. doi:10.1016/j.humpath.2019.04.009. [PubMed: 31054895]
- [127]. Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol* 2015;33:285–9. doi:10.1038/nbt.3129. [PubMed: 25599178]
- [128]. Angermueller C, Clark SJ, Lee HJ, Macaulay IC, Teng MJ, Hu TX, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods* 2016;13:229–32. doi:10.1038/nmeth.3728. [PubMed: 26752769]
- [129]. Hu Y, Huang K, An Q, Du G, Hu G, Xue J, et al. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol* 2016;17:88. doi:10.1186/s13059-016-0950-z. [PubMed: 27150361]
- [130]. Clark SJ, Argelaguet R, Kapourani C-A, Stubbs TM, Lee HJ, Alda-Catalinas C, et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* 2018;9:781. doi:10.1038/s41467-018-03149-4. [PubMed: 29472610]

- [131]. Hou Y, Guo H, Cao C, Li X, Hu B, Zhu P, et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. *Cell Res* 2016;26:304–19. doi:10.1038/cr.2016.23. [PubMed: 26902283]
- [132]. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 2014;344:1396–401. doi:10.1126/science.1254257. [PubMed: 24925914]
- [133]. Moudgil A, Wilkinson MN, Chen X, He J, Cammack AJ, Vasek MJ, et al. Self-reporting transposons enable simultaneous readout of gene expression and transcription factor binding in single cells. *Genomics*; 2019. doi:10.1101/538553.
- [134]. Macosko EZ, Basu A, Satija R, Nemes J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 2015;161:1202–14. doi:10.1016/j.cell.2015.05.002. [PubMed: 26000488]
- [135]. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell* 2015;161:1187–201. doi: 10.1016/j.cell.2015.04.044. [PubMed: 26000487]
- [136]. Sasagawa Y, Danno H, Takada H, Ebisawa M, Tanaka K, Hayashi T, et al. Quartz-Seq2: a high-throughput single-cell RNA-sequencing method that effectively uses limited sequence reads. *Genome Biol* 2018;19:29. doi:10.1186/s13059-018-1407-3. [PubMed: 29523163]

### Highlights

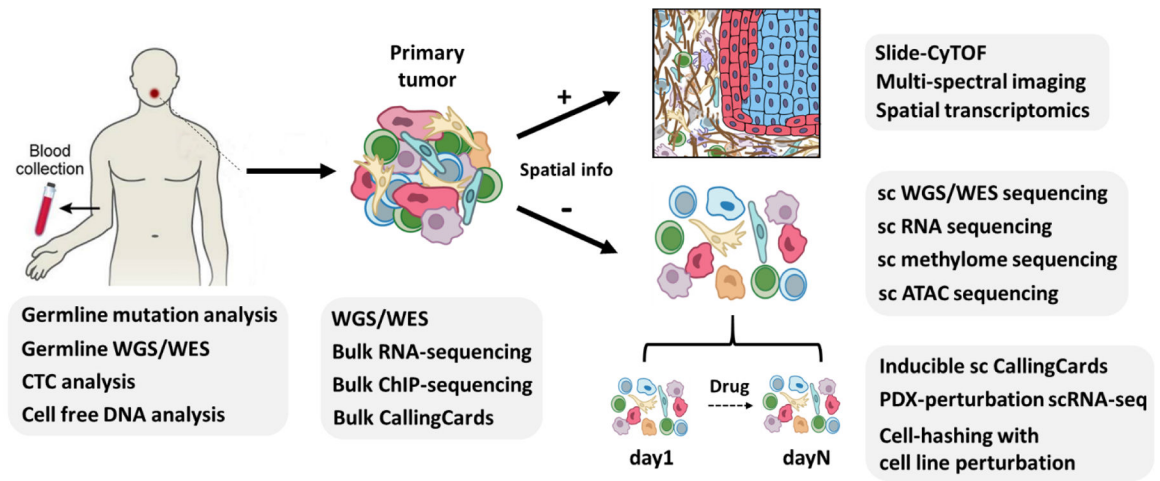
Head and neck squamous cell carcinoma (HNSCC) is characterized by significant intratumoral heterogeneity.

Single-cell sequencing (SCC) technologies provide valuable insights into tumor heterogeneity.

SCS reveals genetic, transcriptional, and epigenetic diversity among malignant, stromal, and immune cells

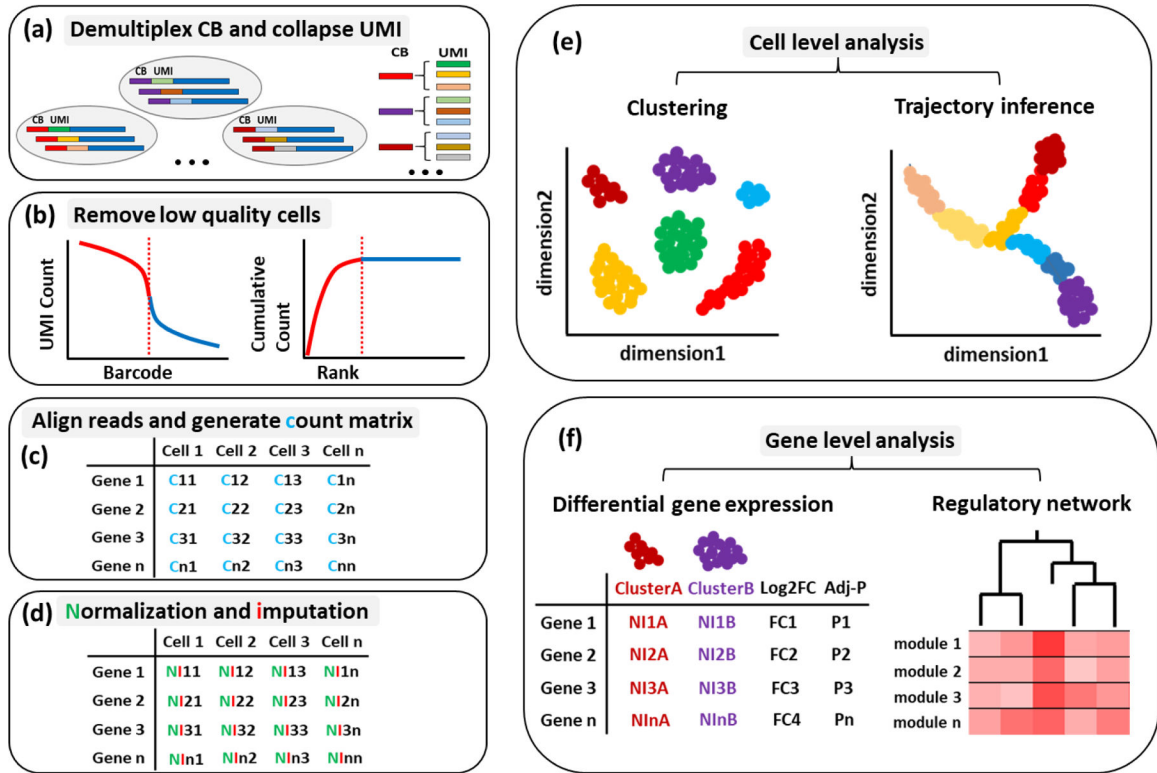
SCS is likely to alter the management of HNSCC as it improves diagnosis and facilitates novel therapeutic design

Advances in SCS should allow this technology to be integrated into clinical decision-making

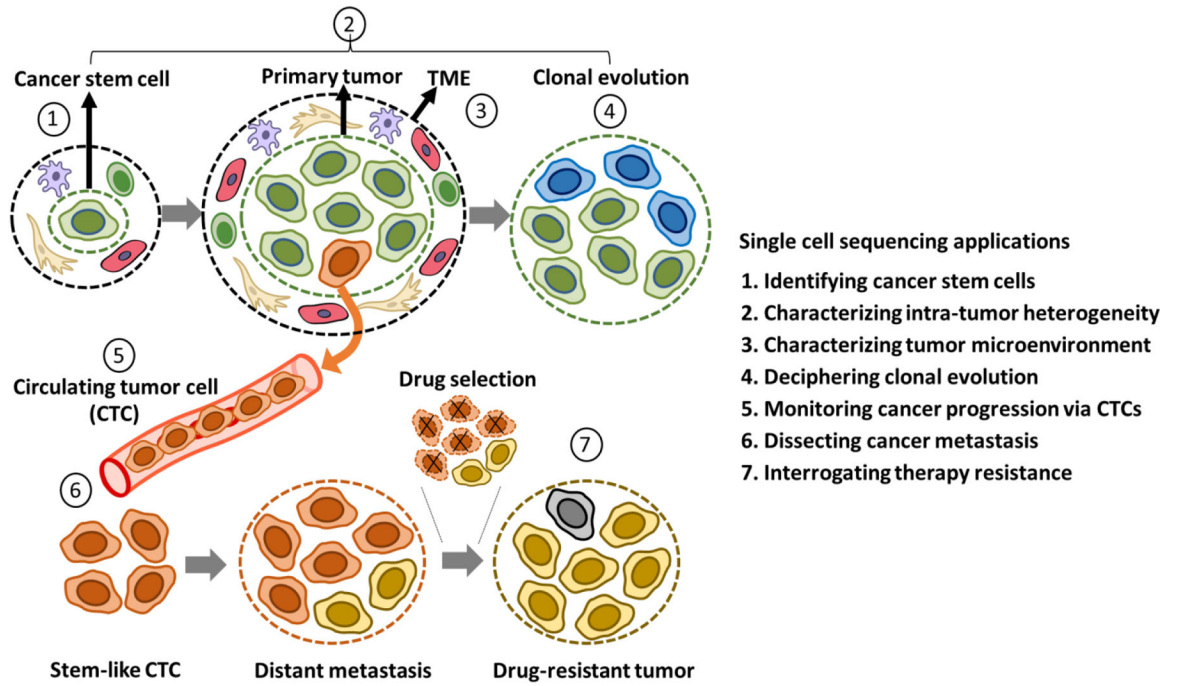


**Figure 1. Overview of the technologies applied at different clinical sample collections.**  
WGS: whole genome sequencing; WES: whole exome sequencing; CTC: circulating tumor cell; sc: single cell; PDX: patient-derived xenograft.





**Figure 2. A schematic overview of scRNA-seq analysis workflow.** scRNA-seq data is noisy due to technical and biological factors. The analysis workflow involves (a) demultiplexing the sequencing reads by cell barcode and collapsing on UMI, (b) removing low quality cells (doublets, multiplets or dead cells), (c) aligning trimmed reads and generating count matrix, (d) normalization by scaling factors and imputation for sparse counts, (e) cell level analysis including clustering and trajectory inference and (f) gene level analysis including differential gene expression and regulatory network.



**Figure 3. Schematic illustration of applications of single-cell sequencing to malignant cells.** The applications of single-cell sequencing are numbered as: 1) Identifying cancer stem cell; 2) Characterizing Intra-tumor heterogeneity; 3) Characterizing tumor microenvironment; 4) Deciphering clonal evolution; 5) Monitoring cancer progress via circulating tumor cells (CTCs); 6) Dissecting cancer metastasis; 7) Interrogating therapy resistance.

**Table 1.**

Summary of single-cell isolation methods

Isolation method	Throughput (cells/run)	Applied sequencing technology	Commercial Platform	Ref
Limiting dilution	Low (10–200)	None	None	[7]
Micromanipulation	Low (10–200)	Smart-seq2	None	[8]
Laser capture microdissection (LCM)	Low (10–200)	Smart-seq2	None	[9]
Flow-activated cell sorting (FACS)	Medium (100–1000)	Smart-seq2 CEL-seq2 STRT-seq MARS-seq	None	[10]
Circuit microfluidics	Medium (100–1000)	Smart-seq2 CEL-seq2 STRT-seq	Fluidigm C1 system (Fluidigm)	[16]
Microdroplet microfluidics	High( 1000–9000)	Drop-Seq InDrop-seq	Chromium system (10xGenomics) InDrop system (1cellBio) Nadia (Dolomite Bio)	[134]
Microwell platform	High (1000–9000)	Cyto-seq SEQ-well	None	[12,13]
In-situ barcoding	Very high (>10000)	SPLiT-seq Sci-RNA-seq	None	[14,15]

**Table 2.**

Summary of single-cell DNA sequencing technologies

scDNA-Seq technology	Genome coverage	Amplification method	Application	Ref
GenomePlex PCR	low	DOP-PCR	CNA	[17]
MDA	high	MDA	SNV	[18]
PicoPLEX	high	DOP-PCR and MDA	SNV and CNA	[19]
MALBAC	high	DOP-PCR and MDA	SNV and CNA	[20]

DOP: degenerate oligonucleotide primers; MDA: multiple displacement amplification; SNV: single-nucleotide variation; CNA: copy number alternations.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 3.**

Summary of single-cell RNA sequencing technologies

scRNA-Seq technology	Transcript coverage	Amplification method	UMI	ERCC	Ref
Tang method	Full length	PCR	No	Yes	[22]
Smart-seq / seq2	Full length	TS-PCR	No	Yes	[28,29]
CEL-seq /seq2	3'-end	IVT	Yes	Yes	[23,24]
STRT-seq / seq2	5'-end	TS-PCR	Yes	Yes	[26,27]
MARS-seq	3'-end	IVT	Yes	Yes	[25]
Cyto-seq	3'-end	GS-PCR	Yes	No	[12]
Drop-seq	3'-end	TS-PCR	Yes	No	[134]
InDrop-seq	3'-end	IVT	Yes	No	[135]
SPLiT-seq	3'-end	TS-PCR	Yes	No	[14]
sci-RNA-seq	3'-end	TS-PCR	Yes	No	[15]
SEQ-well	3'-end	TS-PCR	Yes	Yes	[13]
Quartz-seq1 / seq2	3'-end	PCR	No	Yes	[136]

TS: template switch; IVT: vitro transcription; ERCC: External RNA Control Consortium; UMI: unique molecular identifier. ERCC controls are RNA transcripts with known sequences and quantity that are applied to calibrate the measurements of RNA expression. UMIs are DNA sequences used to label each individual transcript within a cell during reverse transcription in order to estimate absolute molecular counts.

**Table 4.**

Summary of single-cell epigenome technologies

scEpi-Seq technology	Throughput (cells/run)	Feature Coverage	Application	Ref
scRRBS-seq	Low (10–100)	Low (10 %)	DNA methylation	[33]
scBS-seq	Low (10–100)	Low (18 %)	DNA methylation	[34]
scCGI-seq	Low (10–100)	High (70 %)	DNA methylation	[37]
scChIL-seq	Very low (5)	Very High (80 %)	Histone modification	[39]
scCUT&tag	High (1000)	Medium (50 %)	Histone modification	[40]
scChIC-seq	Medium (100–300)	Medium (50 %)	Histone modification	[41]
scATAC-seq	High (1000)	High (60 %)	Chromatin structure	[42]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5.**

Studies that use single-cell sequencing to characterize human tumors

Study	Malignancy	Single-cell Technology	Key Findings
Puram et al 2017	HNSCC	scRNA-seq	<ul style="list-style-type: none"> <li>Classified HNSCC to 3 subtypes: basal-mesenchymal, classical, atypical</li> <li>p-EMT program at tumor edge with CAF-TGFβ signaling</li> <li>p-EMT program recapitulated in LNs</li> <li>Higher p-EMT score associated with worse prognosis</li> </ul>
<b>Cyril Nefitel et al 2019</b>	GBM	scRNA-seq	<ul style="list-style-type: none"> <li>Four cellular states drive glioblastoma malignant cells heterogeneity</li> <li>In vivo single-cell lineage tracing supports plasticity between these four states</li> <li>Genetics and the microenvironment influence the frequency of cells in each state</li> <li>TCGA subtypes reflect the highest-frequency malignant states and the microenvironment</li> </ul>
<b>Mariella G. Filbin et al 2018</b>	GBM	scRNA-seq	<ul style="list-style-type: none"> <li>Gliomas with histone H3 lysine27-to-methionine mutations primarily contain cells that resemble oligodendrocyte precursor cells (OPC-like), whereas more differentiated malignant cells are a minority</li> <li>OPC-like cells exhibit greater proliferation and tumor-propagating potential than their more differentiated counterparts and are at least in part sustained by PDGFRA signaling.</li> </ul>
<b>Andrew S Venteicher et al 2017</b>	GBM	scRNA-seq	<ul style="list-style-type: none"> <li>differences in bulk expression profiles between IDH-A and IDH-O are primarily explained by the impact of signature genetic events and TME composition, but not by distinct expression programs of glial lineages in the malignant cells.</li> <li>higher-grade tumors present enhanced proliferation, larger pools of undifferentiated glioma cells, and an increase in macrophage over microglia programs in the TME</li> </ul>
Patel et al 2014	GBM	scRNA-seq	<ul style="list-style-type: none"> <li>Each tumor expressed at least 2 signatures of subtypes previously believed to be exclusive based on WES</li> </ul>
Casasent et al 2018	BrCa	scDNA-seq TSCS	<ul style="list-style-type: none"> <li>Genome evolution occurred in ducts prior to invasion from single clone(s)</li> <li>Multiple clones able to invade independently</li> </ul>
Kim et al 2018	TNBC	scDNA-seq scRNA-seq	<ul style="list-style-type: none"> <li>Treatment-resistant tumors harbored resistant clones that were undetectable with WES but detectable on single-cell level prior to initiation of chemotherapy</li> <li>Transcriptional changes within treatment resistant samples identified throughout course of treatment</li> </ul>
Guo et al 2018	Lung	scRNA-seq	<ul style="list-style-type: none"> <li>Identified “pre-exhausted” T cells</li> <li>Patients with high ratio of “pre-exhausted” to exhausted T cells had improved prognosis</li> </ul>
Yost et al 2019	BCC/cuSCC	scRNA-seq TCR-seq	<ul style="list-style-type: none"> <li>Included patients before/after anti-PD-1 therapy</li> <li>Clonal expansion of CD8+ T cells with novel clonotypes not present in pre-treatment tissue samples</li> </ul>
Jerby-Amon et al 2018	Melanoma	scRNA-seq	<ul style="list-style-type: none"> <li>Identified ICI resistance program associated with T cell exclusion and immune evasion</li> <li>Resistance program predicted clinical response to anti-PD-1 therapy in independent 112 patient cohort</li> </ul>
Sade-Feldman et al 2018	Melanoama	scRNA-seq	<ul style="list-style-type: none"> <li>Two distinct states of CD8+ T cells were identified and relative proportion of each was linked to regression or progression after ICI</li> <li>TCF7 expression in CD8+ T cells predictive of ICI response in independent cohort</li> </ul>
<b>Peter van Glen et al 2019</b>	AML	scRNA-seq	<ul style="list-style-type: none"> <li>Variable cell-type composition of AML correlates to genetics and outcome</li> <li>Primitive AML cells aberrantly co-express stemness and myeloid priming genes</li> <li>Differentiated AML cells express immunomodulatory factors and suppress T cells</li> </ul>
Li et al 2017	CRC	scRNA-seq	<ul style="list-style-type: none"> <li>Defined 2 subtypes of CAFs in tumor samples</li> <li>EMT-related genes upregulated in CAF subpopulations</li> <li>Improved prognostication based on single-cell expression compared to subtypes based on bulk transcriptomics</li> </ul>

Study	Malignancy	Single-cell Technology	Key Findings
Leung et al 2017	CRC	scDNA-seq	<ul style="list-style-type: none"> <li>• Supported linear theory of metastases in progression toward liver metastasis</li> <li>• One patient with two distinct metastatic sites able to be traced to separate mutation events in primary tumor through phylogenetic analysis</li> </ul>

BCC: basal cell carcinoma; BrCa: breast cancer; CAF: cancer-associated fibroblast; CRC: colorectal carcinoma; cuSCC: cutaneous squamous cell carcinoma; GBM: glioblastoma multiforme; HNSCC: head and neck squamous cell carcinoma; ICI: immune checkpoint inhibition; LN: lymph node; p-EMT: partial epithelial-to-mesenchymal transition; scDNA-seq: single cell DNA sequencing; scRNA-seq: single cell RNA sequencing; TCR-seq: T cell receptor sequencing; TNBC: triple negative breast cancer; TSCS: topographic single cell sequencing; WES: whole exome sequencing.