Deciphering functional tumor states at single-cell resolution

Rolando Vegliante[1,†], Ievgenia Pastushenko[1,†] & Cédric Blanpain[1,2,*] (iD)

Abstract

Within a tumor, cancer cells exist in different states that are associated with distinct tumor functions, including proliferation, differentiation, invasion, metastasis, and resistance to anti-cancer therapy. The identification of the gene regulatory networks underpinning each state is essential for better understanding functional tumor heterogeneity and revealing tumor vulnerabilities. Here, we review the different studies identifying tumor states by single-cell sequencing approaches and the mechanisms that promote and sustain these functional states and regulate their transitions. We also describe how different tumor states are spatially distributed and interact with the specific stromal cells that compose the tumor microenvironment. Finally, we discuss how the understanding of tumor plasticity and transition states can be used to develop new strategies to improve cancer therapy.

Introduction

Solid cancers are composed of tumor cells (TCs) and their stroma, which includes cancer-associated fibroblasts (CAFs), vascular cells, extracellular matrix, and immune/inflammatory cells. TCs do not usually constitute a homogeneous cell population. They are rather composed of functionally heterogeneous populations that present different cellular states dynamically evolving over time. The notion that tumors are composed by heterogeneous subpopulations of TCs with different histology, karyotype, growth rates, enzymes, and response to cytotoxic drugs has been known for decades (Heppner, 1984). Initially, tumor heterogeneity has been attributed only to genetic diversity arising from clonal evolution, which is discussed in the review by Swanton and colleagues (Vendramin *et al*, 2021). This gene-centric hypothesis has been challenged by the discovery of functional diversity of TCs. Functional assays, such as *in vitro* clonogenic assays, transplantation, and *in vivo* lineage tracing, have

suggested that some tumors are hierarchically organized and present a population of cells called cancer stem cells (CSCs) that sustain tumor growth by giving rise to TCs with more restricted proliferative potential (Lapidot *et al*, 1994; Beck & Blanpain, 2013; Prager *et al*, 2019). These two concepts are not mutually exclusive, as populations of CSCs can exhibit substantial intra-tumoral genetic heterogeneity (Shipitsin *et al*, 2007). Historically, similar to the primary tumors, metastasis-initiating cells have been initially dominated by the genetic model of tumor evolution. Although additional driver mutations can be found during metastatic dissemination (Yates *et al*, 2017; Nayar *et al*, 2019), metastases usually do not present driver mutations exclusively found in metastasis and not in primary tumors, suggesting that other mechanisms besides genetic evolution can drive the metastatic dissemination (Birkbak & McGranahan, 2020; Massague & Ganesh, 2021). Growing evidence indicates that intra-tumoral heterogeneity in primary tumors and metastasis is not only determined by genetic and epigenetic features in cancer cells but can also be influenced by the tumor microenvironment.

Historically, researchers had studied functional heterogeneity in cancers using different functional assays. Inspired by developmental and stem cell biology, cancer biologists had identified cell surface markers that are heterogeneously expressed within a given tumor and separated TCs into distinct subpopulations and assessed their clonogenic and differentiation potential by transplantation experiments using limiting dilutions. Using this approach, Dick and colleagues demonstrated for the first time, the existence of a small population of leukemic cells that were much more efficient at forming secondary leukemia than the bulk of cancer (Lapidot *et al*, 1994; Bonnet & Dick, 1997) and called this subpopulation leukemic stem cells. This approach has been used to identify, characterize functionally and molecularly CSCs in many different cancers (Al-Hajj *et al*, 2003; Nassar & Blanpain, 2016). Whereas isolation followed by transplantation *in vivo* or clonogenic assays *in vitro* had been widely used to characterize CSCs, these assays have their limitations as cancer cells are dissociated from their surrounding microenvironment and transplanted very often in a heterotopic site. To study tumor heterogeneity within its native microenvironment, researchers have used lineage tracing to label TCs *in situ* and studied their ability to change fate and their clonogenic potential over time. These studies in different types of solid tumors have demonstrated that not all TCs are equal and clonogenic, and only a small

---

1 Laboratory of Stem Cells and Cancer, Université Libre de Bruxelles, Brussels, Belgium
2 WELBIO, Université Libre de Bruxelles, Brussels, Belgium
*Corresponding author. Tel: +32 2 555 4175; E-mail: cedric.blanpain@ulb.be
†These authors contributed equally to this work as first authors

population survives long term and fuels the tumor growth (Nassar & Blanpain, 2016). Other methods such as pulse-chase experiments have been used to characterize proliferation heterogeneity within the tumor (Fillmore & Kuperwasser, 2008; Pece et al, 2010; Roesch et al, 2010; Schober & Fuchs, 2011; Brown et al, 2017). The same approaches using cell surface markers have been used to identify TC populations with increased metastatic potential called metastasis-initiating cells (Celia-Terrassa & Kang, 2016; Pascual et al, 2017; Massague & Ganesh, 2021) or to deconvolute tumor heterogeneity occurring during epithelial-to-mesenchymal transition (EMT) and identify TC populations enriched for metastasis-initiating cells (Pastushenko et al, 2018).

However, these approaches are biased toward the availability of cell surface markers that recognize the different TC populations, or their physical or chemical characteristics such as the expression of aldehyde dehydrogenase (Moreb, 2008; Luo et al, 2012). Nevertheless, bulk RNA and DNA sequencing of human tumors allowed the identification of genetic and transcriptional heterogeneity between different tumors, giving rise to more clinically relevant, molecular-based classifications (Verhaak et al, 2010; Cancer Genome Atlas, 2012a, 2012b, 2015; Brennan et al, 2013). However, bulk sequencing approaches average the genetic and expression profiles of the different tumor subpopulations and are not very powerful in identifying the distinct tumor states that compose the tumors.

The development of new technologies based on single-cell sequencing opened new avenues to capture different tumor states and understand intra-tumoral heterogeneity with unprecedented resolution and scale. Single-cell RNA sequencing (scRNA-seq) allows to define the transcriptome of individual TCs and to identify clusters of cells (cell states) presenting similar gene expression profiles within a tumor. During the last years, several tumor states, including proliferative, differentiated, invasive, hypoxic, metastatic, and stress-like states, have been identified and functionally characterized. Each tumor state is associated with different hallmarks of cancer, such as tumor progression, metastasis, and resistance to therapy. Single-cell profiling also allows the identification of stromal cells including CAFs, immune cells, and endothelial cells that compose the tumor microenvironment (Fig 1). Single-cell "-omics"
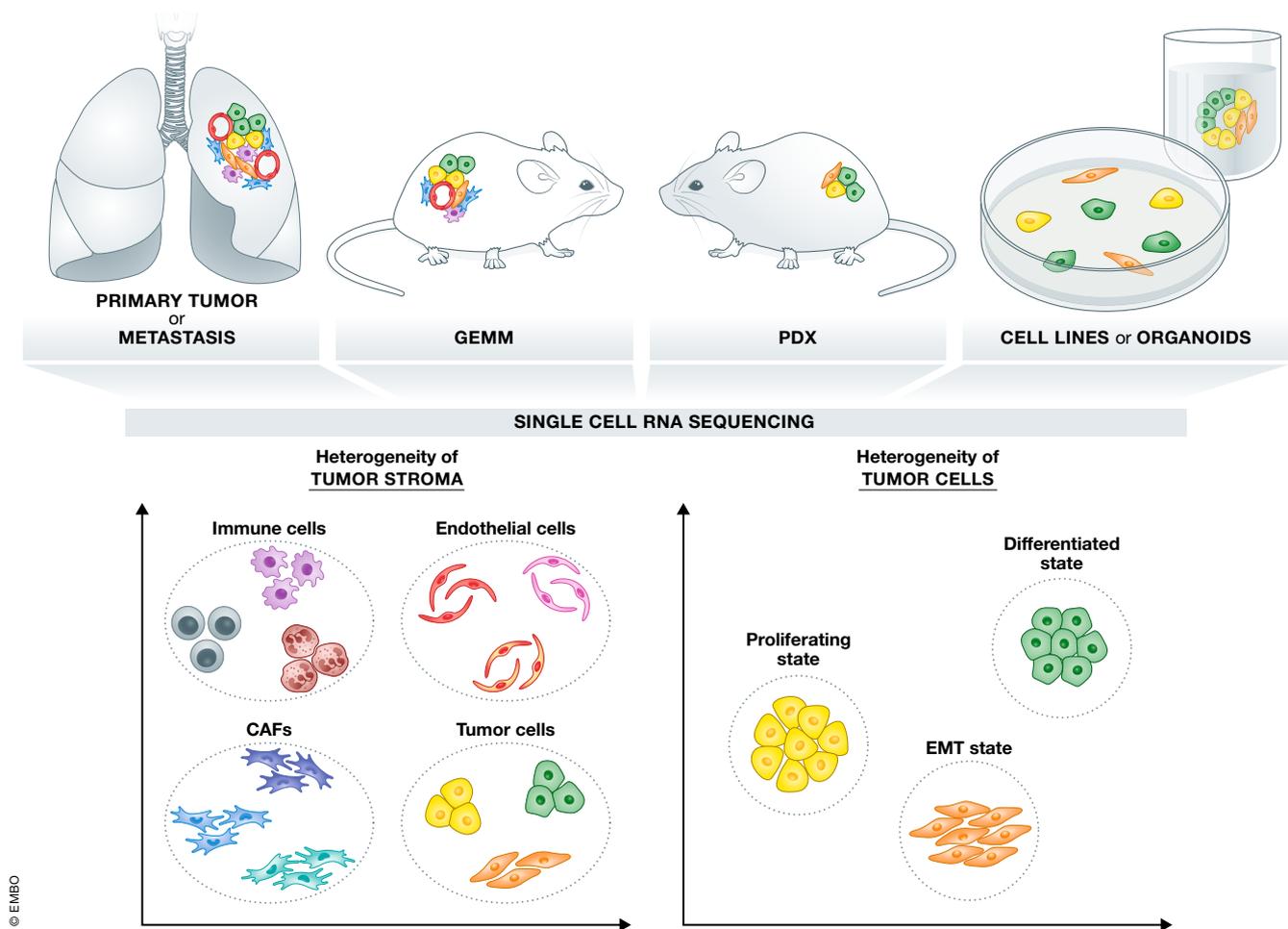

Figure 1. Single-cell RNA sequencing to unravel tumor heterogeneity.

Schematic showing how tumors derived from different models can be used for transcriptional analysis at the single-cell level of both malignant and stromal cells. CAF: cancer-associated fibroblasts; EMT: epithelial-to-mesenchymal transition; GEMM: genetically engineered mouse model; PDX: patient-derived xenograft.

(genomics, transcriptomics, epigenomics, and proteomics) technologies are evolving at a very rapid pace (see Box 1). In this review, we will summarize the most recent single-cell studies that allowed the identification of tumor functional states (Table 1), their spatial organization, and their dynamics during tumor progression, metastasis, and response to therapy.

Tumor states in solid tumors

Certain tumors hijack the developmental programs of normal tissues and mimic their cellular hierarchy leading to TCs with high self-renewing capacities known as CSCs and TCs that are more differentiated and have decreased tumor-initiating capacity. Single-cell sequencing studies have described the similarities of tumor subpopulations and developmental lineages.

Developmental programs in brain tumors

A pioneering work using scRNA-seq identified different tumor states in primary human glioblastomas. These states include cycling, quiescent, hypoxic, and a continuum of stem-like tumor states. Interestingly, malignant cells invading the surrounding tissue express a transcriptional program characterized by low hypoxia, low proliferation, and high migratory capacity (Patel *et al*, 2014). In glioblastoma, four distinct tumor states were found that recapitulate the different neural cell types including neural progenitor-like, oligodendrocyte-like, astrocyte-like, and mesenchymal-like. Each

state contains proliferative cells, although higher proliferation is observed in the neural- and oligodendrocyte progenitor-like states (Neftel *et al*, 2019). Mass cytometry, immunostaining, and xenografting experiments demonstrated that among the different glioblastoma subpopulations, glial progenitor cells are more proliferative and possess higher tumor formation capacity than TCs with astrocyte and neuronal lineage differentiation. This highly proliferative state is also the most resistant to the alkylating agent temozolomide (Couturier *et al*, 2020). Importantly, specific genetic alterations promote the relative abundance of the different tumor states. For example, mutations in *EGFR* promote the astrocyte-like state abundance, whereas TCs harboring chromosome 5q deletion preferentially differentiate into a mesenchymal-like state.

In human *IDH1* or *IDH2* mutant oligodendroglioma and astrocytoma, the majority of cancer cells differentiate along two glial transcriptional programs: oligodendrocytes (characterized by *OLIG1, OLIG2,* or *OMG* expression) and astrocytes (characterized by *APOE, ALDOC,* or *SOX9* expression) (Tirosh *et al*, 2016b). In addition, a rare subpopulation of undifferentiated TCs, associated with a neural stem cell expression program, is enriched for proliferation signature, suggesting that this stem-cell-like population fuels tumor growth (Tirosh *et al*, 2016b). Further analysis of scRNA-seq data showed that these undifferentiated TCs exhibit a strong similarity in gene expression profile between the two tumor histotypes, raising the possibility of a shared cell of origin for IDH-mutant oligodendroglioma and astrocytoma (Venteicher *et al*, 2017). Similarly, in H3K27M glioma, the stem-like state proliferates more actively and

---

Box 1

Single-cell sequencing has rapidly evolved over the last 10 years and has revolutionized the ability to interrogate tumor heterogeneity. These approaches enable to investigate the transcriptional, genomic, epigenomic, and proteomic features of thousands of individual cells within a tumor. Although scRNA-seq has largely contributed to the identification of TC states, analysis of other layers of cell features, either on their own or integrated with transcriptomic interrogation, have greatly expanded our understanding of cancer biology. Single-cell whole-genome sequencing (scWGS) is a powerful tool to decipher cell heterogeneity in a biological sample. Several single-cell genome amplification techniques have been developed, such as DOP-PCR, multiple displacement amplification (MDA), multiple annealing and looping-based amplification cycles (MALBAC), and linear amplification via transposon insertion (LIANTI). These techniques can accurately call copy number variations, indels, and single-nucleotide variations (Navin, 2015; Mallory *et al*, 2020). High-dimensional single-cell DNA sequencing in a clinical context is limited by its high cost. One option is to apply single-cell targeted sequencing of the genetic aberrations of interest identified by bulk sequencing. Techniques of such type carry the bias of seeking pre-identified mutations leaving out the discovery of new ones (Rodriguez-Meira *et al*, 2019). Epigenetics plays a pivotal role in cell biology as different modifications on DNA (e.g., methylation) and histones, as well as varying chromatin accessibility and organization, tune gene expression (Kelsey *et al*, 2017). Single-cell methods that analyze the epigenetic landscapes of thousands of single cells have rapidly expanded. Single-cell chromatin immunoprecipitation sequencing (scChIPseq) detects direct binding of transcription factors onto DNA (Rotem *et al*, 2015); high-throughput chromosome conformation capture (Hi-C) determines high-order chromatin organization (Nagano *et al*, 2015); DNA modifications such as methylation can be identified by single-cell bisulfite sequencing (sc-BSseq) or inferred by the use of restriction enzymes whose activity depends on the methylation state of DNA (Guo *et al*, 2013; Cheow *et al*, 2015). Single-cell chemical-labeling-enabled C-to-T conversion sequencing (CLEVER-seq) enables identification of 5-formycytosine (5fc), where C is read as a T after specific chemical labeling (Zhu *et al*, 2017). A wide application in single-cell -omics has been found for the assay for transposase-accessible chromatin sequencing (ATAC-seq) that interrogates chromatin accessibility based on the ability of the Tn5 transposase to add sequencing adapters into open chromatin regions (Buenrostro *et al*, 2015; Cusanovich *et al*, 2015). As opposite to single-cell technologies detecting nucleic acids, analysis of the proteome in a cell progresses at a slower pace. Mass cytometry, commercialized as CyTOF (mass cytometry by time of flight), relies on metal-isotope-conjugated antibodies for immunolabeling, and analyses them by mass spectrometry (MS) and has emerged as a powerful tool in the field of single-cell proteomics as it can measure 40 to 100 parameters in each cell (Bandura *et al*, 2009; Bendall *et al*, 2011). Increased throughput and peptide quantification are the advantages of an improved mass cytometry-based platform called SCoPE-MS (Budnik *et al*, 2018). Based on the same technology, imaging CyTOF has been developed and can be applied to tissues on slides to combine proteomics and spatial architecture (Giesen *et al*, 2014). Similarly, MALDI imaging mass spectrometry relies on a matrix that coats a tissue sections, extracts molecule from the tissues, and generates mass spectra that can be matched with histological staining of the section (Norris & Caprioli, 2013). Importantly, many of the techniques described above have been integrated into single-cell multi-omics. Generally, scRNA-seq is included in most scMulti-omics studies as gene expression data represent a necessary tool to decipher cell processes. Simultaneous characterization of different layers within a single cell has already taken off and will continue to provide valuable information about cell identities. For a detailed and comprehensive discussion on the integrative approaches used for scMulti-omics, please refer to reviews (Ma *et al*, 2020; Longo *et al*, 2021).

Table 1. Selected studies that used single-cell RNA sequencing to understand tumor cell states.

Reference/Year	Species	Model	Tumor type	scRNA-seq analysis on Tumor cells	Stromal cells
Kinker *et al* (2020)	Human	Cancer cell lines	198 cell lines from 22 cancer subtypes	X	
Kumar *et al* (2018)	Mouse	Syngeneic tumors	Melanoma, breast, lung, colon carcinomas, and fibrosarcoma	X	X
Cook and Vanderhyden (2020)	Human	Cancer cell lines	Lung, prostate, breast, and ovarian cancer cell line	X	
Chung *et al* (2017)	Human	Primary, metastatic lymph nodes	Breast cancer	X	X
McFaline-Figueroa *et al* (2019)	Human	Cancer cell line		X	
Davis *et al* (2021)	Human	PDX, tumor, lymph node, and lung metastasis		X	
Deshmukh *et al* (2021)	Human	Cancer cell line		X	
Karaayvaz *et al* (2018)	Human	Primary tumors	Triple-Negative Breast Cancer	X	X
Pal *et al* (2021)	Human	(healthy, pre-neoplastic, tumoral) tissue + lymph nodes	Mammary duct and breast cancer	X	X
Li *et al* (2017)	Human	Primary tumors	CRC	X	X
Lee at al. (2020)	Human	Primary tumors		X	X
Gojo *et al* (2020)	Human	Primary tumors, PDX	Ependymoma	X	X
Gillen *et al* (2020)	Human	Primary tumors		X	
Neftel *et al* (2019)	Human	Primary tumors	Glioblastoma	X	
Patel *et al* (2014)	Human	Primary tumors		X	
Pine *et al* (2020)	Human	Organoids, PDX		x	
Filbin *et al* (2018)	Human	Primary tumors, PDX	Glioma	X	
Venteicher *et al* (2017)	Human	Primary tumors	Gliomas (Oligodendroglioma and Astrocytoma)	X	
Hovestadt *et al* (2019)	Human	Primary tumors	Medulloblastoma	X	
Tirosh *et al* (2016)	Human	Primary tumors	Oligodendroglioma	X	
Yao *et al* (2020)	Mouse	4-NQO induced primary tumors	Esophageal SCC	X	
Wu *et al* (2018)	Human	Primary tumors	Esophageal SCC and ADC	X	
Puram *et al* (2017)	Human	Primary tumors and metastatic lymph nodes	Head and Neck cancer	X	X
Chen *et al* (2020)	Human	Primary tumors	Nasopharyngeal carcinoma	X	X
Zhao *et al* (2020)	Human	Primary tumors		X	X
Kim *et al* (2020)	Human	Primary tumors, pleural fluids, lymph node, and brain metastasis	Lung adenocarcinoma	X	X
Laughney *et al* (2020)	human	primary tumors + metastasis		X	X
Quinn *et al* (2021)	Human	PDX cancer cell lines, primary tumors, and metastasis	Lung cancer	X	
Ireland *et al* (2020)	Mouse	Cell lines derived from primary tumors	Small Cell Lung Cancer	X	
Tirosh *et al* (2016)	Human	Primary tumors and lymph nodes + distant metastasis	Melanoma	X	X
Wouters *et al* (2020)	human	cell lines derived from primary tumors		X	
Pastushenko *et al* (2018)	Mouse	Primary tumors	Skin SCC	X	
Ji *et al* (2020)	Human	Primary tumors, PDX		X	X
Hu *et al* (2020)	Human	Primary tumors	Ovarian and endometrial cancer	X	X
Izar *et al* (2020)	Human	Cells isolated from ascites	Ovarian cancer	X	X
Peng *et al* (2019)	Human	Primary tumors	Prostate carcinoma	X	

Table 1 (continued)

Reference/Year	Species	Model	Tumor type	scRNA-seq analysis on	
				Tumor cells	Stromal cells
Chen *et al* (2021)	Human	Primary tumors, one lymph node metastasis	Prostate carcinoma	X	X
Young *et al* (2018)	Human	Primary tumors	Renal cancer	X	

presents higher clonogenic capacity upon transplantation as compared to the more differentiated states (Filbin *et al*, 2018). Medulloblastoma, a childhood cerebellar tumor, comprises four molecular subgroups associated with different oncogenic mutations and transcriptional landscapes (Liu *et al*, 2020). WNT, SHH, and Group 3 Medulloblastoma contain both stem-like and more differentiated tumor states. The differentiation of the SHH subgroup resembles cerebellar granule neurons, whereas Groups 3 and 4 resemble neuronal-like cells in different proportions (Hovestadt *et al*, 2019). In ependymoma, a similar hierarchical organization of cell states including stem-like and differentiated tumor states was found. The stem-like state is associated with a poorer prognosis compared to tumors with differentiated-like states (Gojo *et al*, 2020). Further investigation of ependymoma subgroups revealed the presence of two subpopulations with ependymal differentiation features (cilia function and cellular transport) and an undifferentiated subpopulation associated with clinical aggressivity (Gillen *et al*, 2020).

Overall, it appears that most brain tumors follow well-defined hierarchical architectures, with stem-like tumor states driving tumor growth and giving rise to more differentiated TCs. Understanding the molecular mechanisms underlying differentiation of TCs can be exploited as a valuable strategy to hinder tumor growth and lead to clinical benefits. Inspired by the treatment of acute promyelocytic leukemia with retinoic acid and arsenic trioxide that triggers leukemia cell differentiation and disease elimination (de The, 2018), new insights in the molecular basis of differentiated tumor states would help designing pro-differentiation therapies in solid tumors.

Normal differentiation and tumor states

As with brain tumors, pairwise scRNA-seq analyses of tumors and matched healthy tissues show that the different tumor states recapitulate the spectrum of differentiation found in the normal tissue. For example, in colorectal cancer, tumor states include enterocytes, goblet cells, and stem cells (Li *et al*, 2017). In normal fallopian tube epithelium, which is the cell of origin of ovarian cancer, four different secretory cell types, including EMT-like cluster and ciliated cell cluster, were identified by scRNA-seq. The ciliated subtype is enriched in low-grade ovarian cancer, whereas the EMT subtype is associated with high-grade serous ovarian cancer and poor overall survival (Hu *et al*, 2020). In triple-negative breast cancer (TNBC), clustering of single cancer cells using bulk RNA-seq-derived signatures of normal basal, luminal progenitors, and differentiated luminal cells of the mammary gland showed that most TCs express the luminal progenitor signature, consistent with these cells being the cell of origin of TNBC (Lim *et al*, 2009; Molyneux *et al*, 2010; Van Keymeulen *et al*, 2015). Similar results were obtained with unsupervised clustering showing the presence of a basal-like signature and that most tumors contain a subpopulation resembling luminal

differentiated cells (Karaayvaz *et al*, 2018). Another study analyzed the cellular composition and heterogeneity of non-pathologic mammary gland and breast cancers in a large cohort of patients. Epithelial cell populations do not differ across healthy patients, whereas the tumor microenvironment composition—particularly fibroblasts—changes between pre- and post-menopausal subjects. The largest proliferative population is found in TNBC rather than ER+ and Her2+ breast cancers. EMT-expressing TCs did not appear as a discrete cluster but were scattered throughout different subclusters across the three breast tumor subtypes. Among EMT-related genes, TNBCs and ER+ tumors expressed higher vimentin than HER2+ (Pal *et al*, 2021).

Pancreatic ductal adenocarcinomas display two types of ductal tumor states expressing typical ductal lineage markers. Nevertheless, the type 2 ductal tumor state expresses much higher levels of pancreatic adenocarcinoma markers, whereas type 1 ductal cells express higher level of genes regulating normal pancreatic functions, including digestion, pancreatic secretion, and bicarbonate transport. However, these cells can be further clustered into two groups, one resembling normal ductal cells and the other similar to malignant type 2 ductal cells (Peng *et al*, 2019). Lineage trajectory analysis suggests that type 1 ductal cells are the cells of origin giving rise to type 2 malignant TCs. The malignant ductal cell markers were used to cluster human pancreatic tumor samples. Proliferative ductal markers are more enriched in two of the three pancreatic adenocarcinoma clusters that are associated with lower survival rate. CDK1, PLK1, and AURKA are markers of proliferative ductal cells, and the pancreatic cancer cell line MIA PaCa-2 growth was suppressed by inhibitors targeting these three proteins (Peng *et al*, 2019). ScRNA-seq of normal lung epithelia and lung adenocarcinoma identified four common differentiated lineages (including alveolar epithelial cell types 1 and 2, ciliated and club cells). In lung adenocarcinoma, two new tumor states that are observed during the regeneration of severely injured lung can be identified (Laughney *et al*, 2020). These two states include *SOX2*-derived *KRT5+* basal-like cells, which exhibit increased RAS signaling and mesenchymal gene enrichment associated with wound response and the other one corresponds to *SOX9*-expressing alveolar epithelial progenitors. (Vaughan *et al*, 2015; Zuo *et al*, 2015; Laughney *et al*, 2020).

Different states of melanoma

Whereas using bulk sequencing, melanoma could be classified as MITF-high or AXL-high, at the single-cell level every tumor contains malignant cells corresponding to both states. Melanoma cells with the AXL program are selected and enriched following treatment with RAF and MEK inhibitors. CAF abundant tumors are enriched for the AXL-high signature, suggesting that CAFs promote the AXL-high tumor state (Tirosh *et al*, 2016a). Melanomas comprise cycling

states and non-cycling states, the latter showing high expression of the histone demethylase KDM5B. Within the cycling population, a high-cycling state with unique high expression of cyclinD3 differs from the low-cycling state. In cell lines characterized by a low percentage of AXL+ cells, the treatment with BRAF/MEK inhibitors promotes the enrichment of AXL+ cells leading to drug tolerance, suggesting that a rare tumor-resistant state exists before treatment (Tirosh et al, 2016a). Moreover, gene regulatory network and trajectory inference in patient-derived cancer cell lines indicate that the increase in AXL expression occurs during the transition from the melanocytic tumor state (characterized by the expression of SOX10 and MITF transcription factors) to the mesenchymal tumor state (characterized by the expression of SOX9 and AP-1) via an intermediate state expressing SOX6 and displaying simultaneous melanocytic (pigmentation) and mesenchymal-like (increased migration) phenotypes. Consistent with a role of SOX10 in regulating the melanocytic state, SOX10 knockdown results in the downregulation of makers of the melanocytic lineage and the increase in the expression of mesenchymal-like genes (Wouters et al, 2020). Altogether, these studies in melanoma show that scRNA-seq analysis allows the identification of dynamically regulated cellular states with different properties.

Insights from mouse models

Tumor mouse models can be helpful in deciphering functional heterogeneity, allowing longitudinal studies from tumor initiation to late stages of tumor progression. Single-cell analyses performed during tumor initiation in a chemically induced mouse model of esophagus squamous cell carcinoma (SCC) revealed different tumor states across different stages of tumor development including hyperplasia, dysplasia, and carcinoma. This study proposed a model where, during carcinogenesis, proliferating cells either switch to a malignant state progressively acquiring gene signatures typical of EMT, angiogenesis, immunosuppression, and invasiveness, or differentiate. The malignant switch is associated with the increase in transcription factors such as Snai3 and Ets1 and the decrease in tumor suppressor gene expression (Trp53, Pit1, and Bclaf1) (Yao et al, 2020). The progression through the malignant phenotype was accompanied by gradual decrease in Notch1 expression, consistent with the high proportion of Notch1 mutations in SCCs (Dotto, 2009; Sanchez-Danes & Blanpain, 2018; Yao et al, 2020).

A study in a mouse model of small cell lung carcinoma (SCLC)— a neuroendocrine tumor classified into four molecular subtypes (SCLC-A, SCLC-N, SCLC-P, and SCLC-Y)—provided deeper insights in the understanding of the mechanisms underpinning cell phenotype plasticity. The authors showed that mutated-MYC overexpression in a pulmonary neuroendocrine cell of origin, in the context of Trp53 and Rb1 deletion, promotes the SCLC-N and SCLC-Y, but not SCLC-P, molecular subtypes (Ireland et al, 2020). scRNA-seq and pseudotime ordering of four tumors developed in this mouse model showed that MYC drives the evolution of SCLC fate from neuroendocrine to non-neuroendocrine states. Mechanistically, MYC increases NOTCH signaling to destabilize the neuroendocrine identity during SCLC evolution. Therefore, this work sheds light on the mechanisms that determine SCLC subtypes, demonstrating that genetic (MYC mutation), cell of origin (neuroendocrine cell), and cell plasticity matter in SCLC evolution (Ireland et al, 2020).

In conclusion, mouse models allow to interrogate the tumor states at different stages of tumor progression. In addition, the genetic overexpression or downregulation of candidate genes and transcription factors in vivo allows the identification of precise molecular mechanisms regulating cell state transitions in the native tumor environment.

Spatial organization and interactions of tumor subpopulations

The identities and functions of different tumor states are regulated by the interactions between TCs and the various components of the TME, including endothelial cells, the immune system, CAFs, extracellular matrix, and various signaling molecules. Different tumor states have also been identified in cell cultures where the TME populations are missing. ScRNA-seq analysis of 198 cancer cell lines from 22 cancer types uncovered different tumor states that are found across multiple cancer cell lines, including proliferation, stress response, interferon response, senescence, and EMT. The tumor states found in cell lines faithfully recapitulate the transcriptional programs of human primary tumors (Kinker et al, 2020). Consistent with what is found in vivo (Puram et al, 2017; Aiello et al, 2018; Pastushenko et al, 2018), cancer cell lines show a continuum of states during EMT (McFaline-Figueroa et al, 2019). These data support the notion that soluble signals can be sufficient to drive tumor heterogeneity in the absence of direct interactions with the TME. However, this is context dependent as demonstrated in a study comparing tumor states in primary glioblastomas and matched in vitro models. The study showed that glioblastoma organoids co-cultured with brain-like cells (generated from the differentiation of human embryonic stem cells) better phenocopy the primary tumor state composition compared to 2D cultures and glioblastoma organoids alone. This demonstrates that the presence of an intact tissue architecture matters in the regulation of tumor states (Pine et al, 2020). Novel bioinformatic software such as CellPhoneDB (Efremova et al, 2020) or NicheNet (Browaeys et al, 2020) that use scRNA-seq data across multiple cell types to predict paracrine signaling pathways that operate within the tissue allow to predict the interactions between TCs and the TME subpopulations. ScRNA-seq-based ligand–receptor predictions revealed that macrophages are involved in the induction of the mesenchymal-like state in a glioblastoma mouse model. Multiplexed error-robust fluorescence in situ hybridization (MERFISH) confirmed the enrichment of macrophages in proximity of EMT-like cells (Hara et al, 2021). Other computational approaches enable to associate specific interaction scores (ligand–receptor) to phenotypes (e.g., tumor growth rate) (Kumar et al, 2018). Interestingly, besides MERFISH, other techniques that localize hundreds of genes in intact tissues such as in situ RNA sequencing and sequential fluorescence in situ hybridization (seqFISH) have been developed (Ke et al, 2013; Lubeck et al, 2014). They allow to quantify the transcriptome in tissue sections, integrating phenotypes with spatial information (Stahl et al, 2016; Eng et al, 2019). Laser-capture microdissection of human liver tumors, combined with bulk- and scRNA-seq allowed the localization of different subpopulations in distinct areas, such as high CAF and T-cell abundance in fibrotic areas and at the tumor borders, respectively (Massalha et al, 2020). Spatial transcriptomics and scRNA-seq integration facilitated the identification of specific cell population enrichment in different areas of pancreatic tumors.

For example, tissue resident M2 macrophages are abundant in the ducts, whereas M1 macrophages are more enriched in the stroma and the cancer regions characterized by an inflammatory environment. Cancer cells that are in physical proximity with inflammatory fibroblasts present an increase in a stress-like state (Moncada *et al*, 2020). Similarly, CSCs in skin SCC are localized in the immediate vicinity of endothelial cells. VEGF overexpression by CSCs sustains tumor stemness and symmetric CSC division, leading to CSC expansion by promoting the formation of the vascular niche and directly regulating tumor stemness in a cellular autonomous manner through Nrp1 signaling in TCs (Beck *et al*, 2011). Also, CSCs in SCCs display an heterogeneous response to TGFβ, with non-responsive TCs fueling tumor growth and responsive TCs acquiring malignant and invasive features. Responsive TCs proliferate slowly and promote glutathione metabolism, both mechanisms protecting cells from cytotoxic drugs such as cisplatin (Oshimori *et al*, 2015).

Overall, these methods enable to integrate cell phenotypes with spatial information. In conclusion, bioinformatic and molecular strategies can predict cell–cell interactions and identify the spatial distribution of tumor states within their ecosystem. These data can eventually be used to assess the functional relevance of these communications and interactions and how they impact TC states.

Tumor states during epithelial-to-mesenchymal transition at single-cell resolution

EMT is a reversible process whereby cells lose their differentiated epithelial characteristics such as polarity and adhesion and acquire mesenchymal features, including motility and invasive properties, as discussed in detail in the review by Brabletz published in the same issue (Brabletz *et al*, 2021). In mouse skin SCCs and mammary tumors, different EMT tumor states have been identified. These tumor states comprise epithelial, hybrid (i.e., showing both epithelial and mesenchymal characteristics), and more mesenchymal tumor states. Interestingly, these tumor states exhibit different tumor functions. The epithelial state is the most proliferative suggesting that epithelial cells drive tumor growth, whereas the mesenchymal cells are more invasive and the intermediate hybrid EMT cells are the most metastatic. Although all EMT subpopulations present a certain degree of plasticity upon transplantation, the early hybrid EMT state is primed toward the epithelial phenotype, while the most mesenchymal TCs preferentially give rise to late EMT tumor states. These states are regulated by different combinations of transcription factors and are located in different cellular niches with specific compositions of stromal cells (Pastushenko *et al*, 2018; Pastushenko & Blanpain, 2019). These different EMT tumor states were identified in other tumor types and the hybrid TCs are also associated with increased metastatic potential (Neftel *et al*, 2019; Wouters *et al*, 2020; Simeonov *et al*, 2021).

Cancer mouse models coupled with fluorescent lineage tracing allowed the tracking of TCs in the primary tumor, in the blood, and at the metastatic sites independently of dynamic changes in gene expression that accompanied EMT. Using a combination of CRE and DRE lineage tracing in MMTV-PyMT breast cancer model, it has been shown that N-cadherin lineage tracing marks the vast majority of metastatic initiating cells, whereas very few metastatic cells were labelled using vimentin lineage tracing. These data support the notion that partial, but not full, EMT is required for the initiation of the metastatic cascade (Li *et al*, 2020; Vieugue & Blanpain, 2020). In contrast to mouse models, in human cancers, it is more complicated to distinguish CAFs from TCs that undergo EMT. Importantly, the analysis of the transcriptomes of tumor and stromal cells at single-cell resolution allows the identification of copy number variations and thus exclusion of non-tumoral cells.

Single-cell transcriptomic analysis of primary and metastatic head and neck SCCs identified hybrid EMT programs. TCs exhibiting this partial EMT phenotype are spatially localized at the leading edge of the tumors (Puram *et al*, 2017). Hybrid EMT states have also been identified in human nasopharyngeal carcinomas (Zhao *et al*, 2020). scRNA-seq of primary human skin SCC showed that cutaneous SCCs contain TCs that recapitulate the major cell states of normal epidermis (basal, cycling, and differentiating) and identified a tumor-specific subpopulation (Ji *et al*, 2020). Interestingly, this tumor-specific cell state expresses EMT markers, such as vimentin, while maintaining high expression levels of epithelial genes, suggesting that, similar to the head and neck tumors, human skin SCC cells undergo partial EMT (Ji *et al*, 2020). Interestingly, SCC cells enriched for hybrid EMT signature were located at the leading edge of the tumor. This tumor-specific EMT state comprise a minority of cells and most of the leading cells are enriched for basal tumor genes. Single-cell spatial transcriptomic analysis identified the ligand–receptor interactions between EMT SCC cells at the leading edge and the surrounding stroma. EMT SCC cells at the leading-edge signal to CAF through MMP9-LRP1 and TNC-SDC1 ligand–receptor interactions. Additionally, this cellular state may modulate the endothelium through the ligand–receptor pairs PGF-FLT1, PGF-NRP2, and EFNB1-EPHB4. Conversely, endothelial cells and CAFs prominently co-express ligands predicted by Niche-Net such as TFPI, FN1, and THBS1, matching EMT SCC expressing receptors (Ji *et al*, 2020).

Combination of single-cell lineage tracing and single-cell transcriptomic analysis in orthotopic transplantation of Kras/p53 mouse pancreatic adenocarcinoma cells showed that cells occupy a continuum of EMT states during tumor progression (Simeonov *et al*, 2021). *In vivo*, non-aggressive clones that constitute the majority of the total number of clones express epithelial genes, while more mesenchymal states are restricted to large aggressive clones. Interestingly, PDAC cells cultured *in vitro* are mostly epithelial. These data suggest that the epithelial tumor state composes the baseline state of PDAC cells and that EMT is induced only *in vivo*. A more detailed analysis of *in vivo* single-cell data identified six tumor states: epithelial state, four hybrid EMT states, and a mesenchymal state. These states exhibit differential enrichment of epithelial and EMT markers and transcription factors. EMT master regulators such as Zeb1, Zeb2, or Snai1 are enriched in early clusters, while ETS-domain TFs are enriched in intermediate hybrid EMT states. Motifs bound by members of the Sox and Fox families are enriched in late hybrid EMT and mesenchymal states. Transitioning from early to late EMT states is accompanied by a strong shift from oxidative phosphorylation toward glycolysis (Simeonov *et al*, 2021). Single-cell analysis of primary human pancreatic cancer also revealed the presence of TCs with high EMT signature. Interestingly, primary tumors were very heterogeneous among different patients, and while 2 of 10 primary samples present very high percentage of TCs with high EMT signature (> 70% of TCs), the remaining samples

present low percentage of TCs with EMT features. All the metastatic lesions analyzed in this study are fully epithelial (Lin *et al*, 2020). scRNA-seq analysis of primary human colorectal cancers revealed that after exclusion of non-tumoral cells based on the copy number variation profile, TCs do not show downregulation of CDH1 or upregulation of known EMT markers and that only CAFs express EMT genes suggesting that EMT signature derived from bulk RNA-seq may sometimes come from CAF contamination (Li *et al*, 2017).

Pseudo-temporal (pseudotime) ordering of single-cell data provides differentiation trajectories that may recapitulate transitions between cell states. This type of analysis in skin SCCs that undergo spontaneous EMT revealed that EMT occurring during tumor progression is not a unique continuum, but shows two different divergent trajectories (Pastushenko *et al*, 2018). Soft clustering analysis of single-cell EMT datasets of skin SCC and embryogenesis of intestine, liver, lung, and skin confirmed the existence of different EMT trajectories in cancer and during normal tissue development (Sha *et al*, 2020b). Different trajectories can also be observed upon TGFβ-induced EMT in lung cancer cells and in mesenchymal-to-epithelial transition occurring upon withdrawal of TGFβ *in vitro* (Karacosta *et al*, 2019). Although the existence of different trajectories during EMT seems to be conserved across different cancer types, the functional significance of these trajectories remains unclear. Further studies are needed to understand whether different EMT trajectories represent different EMT states (hybrid versus full EMT), have different metastatic potential or respond differently to therapy (Fig 2A). *In situ* identification and characterization of different trajectories and whether a specific trajectory is associated with a particular niche remain open questions.

In addition, the cell-autonomous and non-cell-autonomous mechanisms regulating different EMT states and trajectories remain poorly understood. ScRNA-seq performed on multiple cancer cell lines revealed EMT-related heterogeneity (McFaline-Figueroa *et al*, 2019; Cook & Vanderhyden, 2020; Deshmukh *et al*, 2021), suggesting that EMT heterogeneity is regulated, at least in part, by cancer cell-autonomous mechanisms such as the expression of Itgb3 and Itgav by TCs that were previously identified in different EMT states (Pastushenko *et al*, 2018). Single-cell analysis of primary murine and human tumors also revealed stromal cell heterogeneity (Kanzaki & Pietras, 2020; Cildir *et al*, 2021; Yuan *et al*, 2021). However, whether stromal cell heterogeneity can modulate EMT states and trajectories, and which are the precise ligand–receptor interactions or other mechanisms controlling this heterogeneity, remains poorly understood.

Recently, bioinformatic analysis of available datasets analyzed the cell–cell communications and the multilayer gene–gene regulation networks underlying gene regulatory dynamics along EMT. Analyses of scRNA-seq datasets of ovarian cancer cell lines exposed to different EMT-inducing agents at different time points revealed that EMT driven by TGFβ-1 is highly synchronized, while EMT promoted by EGF and TNF is asynchronous. Ligand–receptor interaction analysis between ovarian cancer cells with different degrees of EMT revealed that the large proportion of mesenchymal cells behaved as "receiver" of TGFβ released from the epithelial and hybrid EMT states (Sha *et al*, 2020a).

The metastatic and the dormant tumor states

During the metastatic cascade, TCs leave the primary tumor, invade the underlying matrix (invasion), penetrate in the bloodstream (intravasation), and travel to the target organ where they extravasate and infiltrate the stroma, giving rise to disseminated TCs (colonization) (Pantel & Brakenhoff, 2004; Lambert *et al*, 2017; Mohme *et al*, 2017) (Fig 2B). Depending on intrinsic or extrinsic factors, disseminated TCs start proliferating giving rise to macrometastasis, or remain dormant during variable periods of time (Risson *et al*, 2020). Single-cell profiling can be key to uncover regulators promoting each stage of the metastatic cascade.

Metastases are seeded by rare TCs with unique properties, the so-called metastasis-initiating cells (MICs). scRNA-seq of TNBC patient-derived xenografts (PDXs) revealed that primary tumors contain a rare subpopulation of stem-like cells, enriched in basal-like gene signature. Similarly, low-burden micrometastasis and circulating tumor cells (CTCs) are enriched in basal-stem like genes, suggesting that these stem-like cells may represent the true metastatic seeder cells (Lawson *et al*, 2015). Micrometastases express higher levels of quiescence- and dormancy-associated genes, while macrometastases show a shift toward a more proliferative signature, with higher levels of cell-cycle-promoting genes such as MYC and CDK2, as well as MMP1 and CD24, which have been previously associated with reactivation after dormancy. Interestingly, treatment with the CDK inhibitor Dinaciclib decreases the incidence of macrometastasis in this model (Lawson *et al*, 2015). scRNA-seq of human lung adenocarcinoma cells from primary tumors at different stages and from metastatic lesions identified two transcriptional tumor states. One trajectory represents normal lung differentiation including club cell, alveolar type 1, and type 2 cells, and the second trajectory is only found in tumors and not normal tissue. This second trajectory signature is associated with late-stage tumors and higher frequency of lymph node and brain metastasis, suggesting that it represents the metastatic initiating cells and correlates with poor survival (Kim *et al*, 2020). Lineage tracing using CRISPR-Cas9 in a human lung adenocarcinoma cell line transplanted into immunodeficient mice demonstrated that primary tumors contain different cell subpopulations with distinct metastatic potentials. The different clones are associated with distinct transcriptional signatures and give rise to metastasis in different organs, suggesting that the organotropism of each clone is pre-existing in the cell line before transplantation (Quinn *et al*, 2021). Different tumor types give rise to metastases that preferentially colonize particular tissues. For example, prostate cancer preferably metastasizes to the bones (Gao *et al*, 2019), while the most common site of metastasis of colon cancer is the liver (Engstrand *et al*, 2018). Breast cancer can metastasize to different sites, including bone, liver, lung, and brain. However, the luminal estrogen receptor (ER)-positive subtype has a higher propensity to metastasize to the bone (Chen *et al*, 2017). Interestingly, using a barcode system and single-cell analysis, the bone microenvironment has been shown to facilitate breast and prostate cancer cells to further metastasize and establish multiorgan secondary metastases. This metastasis-promoting effect is driven by epigenetic reprogramming, enhanced by EZH2 activity, which confers stem cell-like properties to cancer cells disseminated from bone lesions (Zhang *et al*, 2021). The osteogenic niche transiently reduces ER expression in bone micrometastasis of luminal ER-positive breast cancer cells through epigenetic reprogramming regulated by EZH2, leading to intra-metastatic heterogeneity and endocrine resistance (Bado *et al*, 2021). Lineage tracing combined with single-cell transcriptomic analysis of Kras/p53 mouse pancreatic
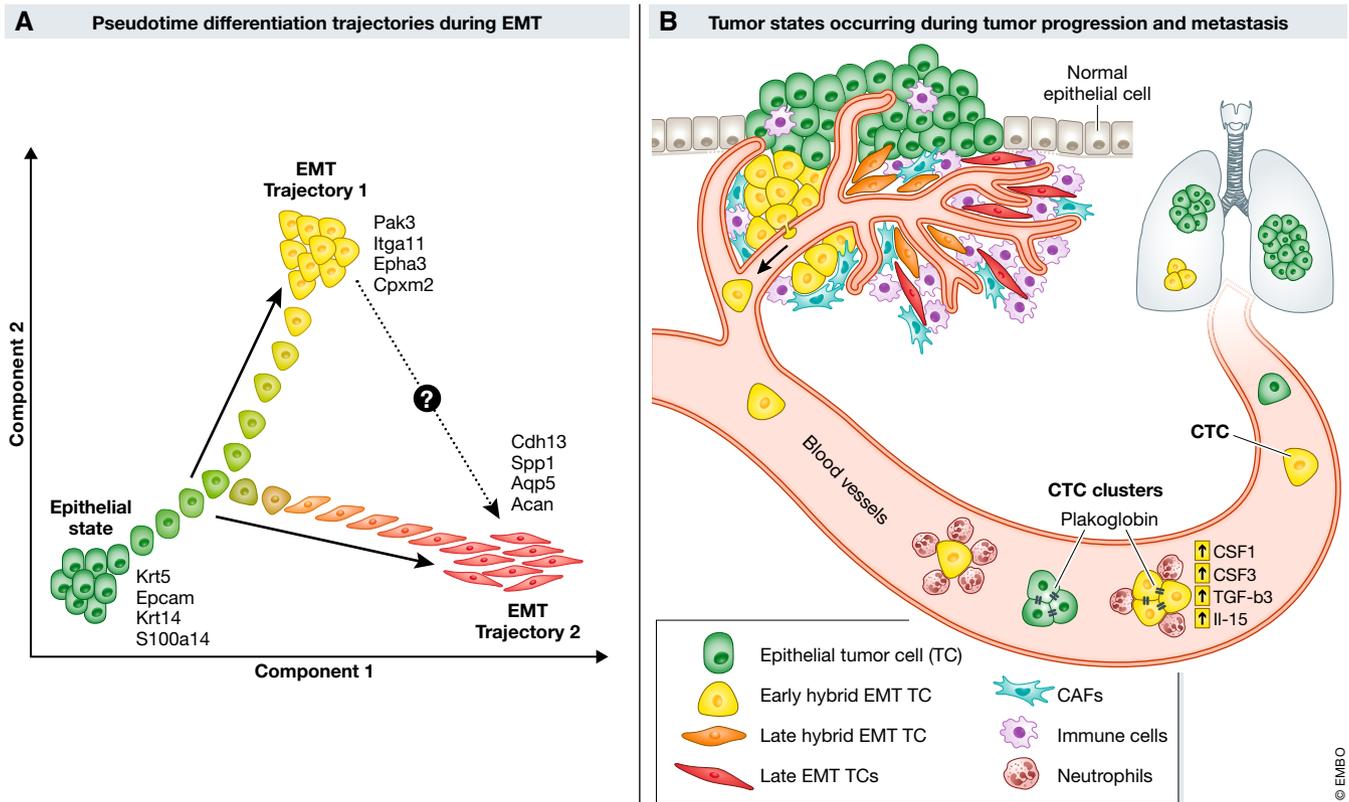
Figure 2. Dynamic differentiation trajectories and states during EMT.

(A) Schematic representation of EMT trajectories described in skin squamous cell carcinoma (SCC) undergoing spontaneous EMT. Epithelial cells give rise to two different EMT trajectories. It is still an open question whether these two EMT trajectories are divergent or whether hybrid EMT trajectory (EMT1) can give rise to late EMT (EMT2). EMT: epithelial-to-mesenchymal transition. (B) Tumor progression occurs through different tumor transition states. These tumor states are located in different niches, characterized by increasing number of immune cells and vascular density. Cells in the hybrid EMT state have the highest metastatic capacity. Circulating TCs can circulate individually or in clusters. CTCs: circulating tumor cells.

adenocarcinoma cells at primary and metastatic sites demonstrated that, in the primary tumor, more than 50% of cells come from a minority of clones. At metastatic sites, 80–90% of cells typically come from a single clone and both mice analyzed have one clearly dominant clone across all disseminated sites. Interestingly, 51% of clones fail to metastasize, suggesting that mutations in Kras and p53 alone do not ensure metastatic potential (Simeonov *et al*, 2021). Interestingly, metastatic dissemination of TCs peaks on late-hybrid EMT state and then sharply declines at highly mesenchymal states (Simeonov *et al*, 2021).

The detection, isolation, and transcriptome analysis of single CTCs offer an appealing minimally invasive approach that allows the characterization of tumor heterogeneity that can help to better understand the biology of tumor evolution and metastasis. The CTCs are very rare events as compared to the total number of blood cells in the circulation. This makes the identification, enumeration, and isolation of CTCs a very challenging process. The use of anti-Epcam antibodies to capture CTCs from blood specimens of patients with epithelial cancers has opened a new field of cancer diagnostics and the use of an automated CellSearch platform in clinical trials has demonstrated the prognostic value of Epcam$^+$ CTCs in several cancer types (Andree *et al*, 2016). Transplantation of Epcam$^+$ CTCs

isolated from patients with primary luminal breast cancer in mice gives rise to bone, lung, and liver metastasis, demonstrating that Epcam$^+$ CTCs contain metastasis-initiating cells. Further analysis revealed that the metastasis-initiating cells are defined by the co-expression of Epcam, CD44, CD47, and MET markers. Interestingly, in a cohort of patients with breast cancer metastasis, the number of Epcam$^+$CD44$^+$CD47$^+$ MET$^+$ CTCs, but not of bulk Epcam$^+$ CTCs, correlates with lower overall survival and increased number of metastatic sites (Baccelli *et al*, 2013). Besides the use of Epcam as a marker for detection of CTCs, the expression of Epcam has been shown to be mechanistically associated with increased cancer cell adhesion, proliferation, migration, and stemness during cancer progression (Keller *et al*, 2019). However, Epcam expression varies according to the tumor type and its expression is also different in primary and metastatic tumors (Eslami *et al*, 2020). CTCs isolated from patients with breast cancer using a label-free microfluidic flow device are heterogeneous regarding Epcam expression (Hyun *et al*, 2016). Similar to primary TCs, CTCs can be plastic and dynamically change their gene expression pattern. CTCs can undergo EMT and lose part of the epithelial characteristics, including Epcam expression. Serial monitoring of CTCs in breast cancer patients revealed enrichment of mesenchymal phenotype in CTCs, while only few

cells expressing EMT markers are detected in primary tumors. In addition, mesenchymal CTCs are associated with disease progression. Interestingly, reversible shifts between epithelial and mesenchymal phenotype accompany each cycle of response to therapy and disease progression (Yu *et al*, 2013). In the mouse model of skin SCC, the majority of YFP+ CTCs were Epcam- (Pastushenko *et al*, 2018; Revenco *et al*, 2019). In contrast, in the syngeneic mouse model of breast cancer, epithelial type CTCs (defined by high levels of Epcam, E-Cadherin, and Grhl2 expression) have the strongest lung metastasis formation as compared to mesenchymal CTCs (characterized by low Epcam, E-cadherin expression, and high levels of Vimentin, Slug, Zeb1, and Zeb2 expression) (Liu *et al*, 2019).

These apparent contradictory results could be due, at least in part, to the different isolation methods of CTCs. Detection methods based on the expression of specific markers do not take into account the heterogeneity of CTCs and thus exclude CTCs that do not express or downregulate this specific antigen. In addition, the expression of an antigen is not always specific of TCs, even circulating Epcam+ cells, thus the quantification of CTCs based only on specific markers should be interpreted with caution. To be able to detect and isolate a pure population of CTCs and to exclude non-tumoral cell contamination, other methods, such as the detection of driver mutation or specific copy number variations, should be implemented. Immunohistochemistry or qPCR allows the detection of a limited number or markers, thus allowing only a partial view on their gene expression profile. In these settings, scRNA-seq provided information on the expression of thousands of genes per cell in a completely unbiased manner, offering important insights regarding the cellular heterogeneity, as well as identifying key pathways in these cells that can have potential clinical utility (Keller & Pantel, 2019). In addition, scRNA-seq data can be used to infer copy number variations to exclude the non-tumoral cells.

Isolation of CTCs from breast cancer patients using Chips coated with Epcam, Her, and EGFR antibodies, followed by scRNA-seq revealed that CTCs derived from lobular type cancers are mostly epithelial, whereas those from triple-negative subtype and Her2+ breast cancer are predominantly mesenchymal (Yu *et al*, 2013). Primary tumors are mainly epithelial and only rare cells co-expressing both epithelial and mesenchymal markers are found. In addition, reversible shifts between epithelial and mesenchymal phenotype can be observed following each cycle of response to therapy and disease progression (Yu *et al*, 2013). scRNA-seq of CTCs derived from patients with advanced breast cancer using Hydro-seq platform that utilizes size-based single-cell capture identified heterogeneous expression of ER, PR, and HER2 receptors in the detected CTCs. The analysis of EMT markers in these CTCs identified two populations, epithelial HER2$^+$ CTCs and mesenchymal HER2$^-$ CTCs. The expression of CSC markers was also heterogeneous, being ALDH highly expressed in epithelial CTCs and CD44 expressed in EMT vimentin$^+$ CTCs. Both inter- and intra-patient heterogeneity in CTCs was observed in this study (Cheng *et al*, 2019). scRNA-seq of CTCs cultures derived from patients with melanoma showed that CTCs upregulate lipogenesis and iron homeostasis pathways that are associated with resistance to BRAF inhibitors and poor outcome (Hong *et al*, 2021). Size-based isolation and scRNA-seq of CTCs in a pancreatic cancer mouse model identified three different subpopulations of CTCs. The biggest CTC cluster is characterized by expression of classic epithelial markers such as Krt7, Krt19, or Epcam. The

second cluster is enriched for platelet markers and the third cluster is characterized by a strong cellular proliferation signature. Interestingly, keratin-expressing classic CTCs exhibit high levels of expression of ECM gene transcripts, such as SPARC. Indeed, SPARC knockdown decreases PDAC proliferation *in vitro* and metastatic capacity *in vivo* upon intravenous injection of TCs in mice (Ting *et al*, 2014). CTC clusters have been shown to have higher metastatic capacity as compared to single CTCs (Aceto *et al*, 2014). Analysis of CTC clusters and primary tumors in mouse models with tagged mammary tumors at single-cell resolution demonstrated that CTC clusters arise from oligoclonal TC groupings and not from intravascular aggregation events. Interestingly, the CTC clusters exhibited high expression of plakoglobin and KD of this cell junction component prevented CTC cluster formation and suppressed lung metastasis (Aceto *et al*, 2014). Moreover, association of CTCs clusters with neutrophils promoted metastasis (Szczerba *et al*, 2019). scRNA-seq of CTCs derived from patients with breast cancer and mouse models uncovered that CTCs that are circulating in clusters and are surrounded by neutrophils exhibit a marked enrichment in positive regulators of cell cycle and DNA replication programs compared to CTCs alone. Interestingly, CTCs from CTC-neutrophil clusters expressed high levels of CSF1, CSF3, TGF-b3, and Il-15, probably involved in neutrophil stimulation, whereas associated neutrophils expressed their receptors. Depletion of neutrophils using anti-Ly6G antibody prevented the formation of CTC-neutrophil clusters and metastasis formation (Szczerba *et al*, 2019). Intra-cardiac injection of CTCs derived from patients with breast cancer into immunodeficient mice gave rise to metastasis in the bones, lungs, and brain similar to the sites where patients with breast cancer develop metastasis. Interestingly, 2 of 4 CTC-derived cell lines give rise to metastasis in the bones, lung, and ovaries, while the other two give rise to brain metastasis and low metastasis in other organs, suggesting a CTC-dependent organotropism. This pattern of metastasis reflects the evolution of patients from which the CTCs are isolated. RNA- and ATAC-sequencing and functional experiments identified semaphorin 4D as a regulator of TC transmigration through the blood–brain barrier and MYC as a regulator of the adaptation of disseminated TC to the brain microenvironment (Klotz *et al*, 2020).

Single-cell sequencing of the TCs at the metastatic sites can provide new insights on the programs required for the extravasation of TCs and metastatic colonization. For example, lymph node metastases of lung adenocarcinoma present higher myeloid infiltration and changes in fibroblasts and endothelial cell composition, suggesting that changes in the microenvironment are key for metastatic colonization (Kim *et al*, 2020). scRNA-seq of TNBC PDXs reveals high transcriptional heterogeneity in primary tumors and micrometastatic lesions. However, micrometastases harbor a distinct transcriptome program that is predictive of poor survival. Pathway analysis revealed mitochondrial oxidative phosphorylation as the top pathway upregulated in micrometastases, in contrast to higher levels of glycolytic enzymes in primary TCs. Importantly, pharmacological inhibition of oxidative phosphorylation dramatically decreases metastatic seeding in the lungs (Davis *et al*, 2020). ScRNA-seq of disseminated TCs and macrometastasis in mouse model of lung adenocarcinoma metastasis (inoculation of H2087 cells into the arterial circulation of athymic mice) led to the identification of specific transcription signatures and states associated with quiescence or spontaneous macrometastasis outgrowth (Laughney

et al, 2020). Disseminated TC that do not form macrometastasis express low levels of Sox2 and Sox9 and a transcriptional signature of the quiescence state found in the primary tumors. Clusters derived from spontaneous macrometastasis showed enrichment of Sox9 and were correlated with regenerative and proliferating stem-like states found in the primary tumor (Laughney et al, 2020).

Tumor relapse after therapy despite apparent complete response remains a major clinical problem that is attributed to the reactivation of dormant TCs that disseminate before primary tumor resection. How dormancy is established, what are the differences between TCs that remain dormant and the cells that actively proliferate after colonization of distant organs and which mechanisms ultimately drive the reactivation of dormant cells and growth of macroscopic metastasis remains poorly understood (Linde et al, 2016). Using label retention experiments in mouse prostate cancer injected intra-cardiacally, the labeled dormant and proliferating TCs were isolated from bone marrow. scRNA-seq revealed that interferon signaling is suppressed in proliferating TCs from bone marrow metastases as compared to dormant cells. The authors further demonstrated that administration of HDAC inhibitors restores the intrinsic interferon signaling and blocks bone metastatic progression (Owen KL EMBO REP 2020). Mouse pancreatic TCs labeled with fluorescent protein were orthotopically transplanted into the mouse pancreas. Few months after transplantation, the pancreas and the spleen were removed, which led to liver and/or peritoneal metastases in 65% of the mice, whereas 35% presented disseminated TCs without macrometastasis detectable by whole-body MRI. Dormant cells exhibit decreased expression of genes related to proliferation, upregulation of CDK inhibitors p27 and p21, and genes involved in immune regulation, metabolism, transport of lipids, and related to angiogenesis. Importantly, the mouse signature of dormant cells is enriched in cells expressing pancreatic cancer marker found in biopsies of normal-appearing liver of patients with pancreatic cancer (preprint: Dudgeon et al, 2020).
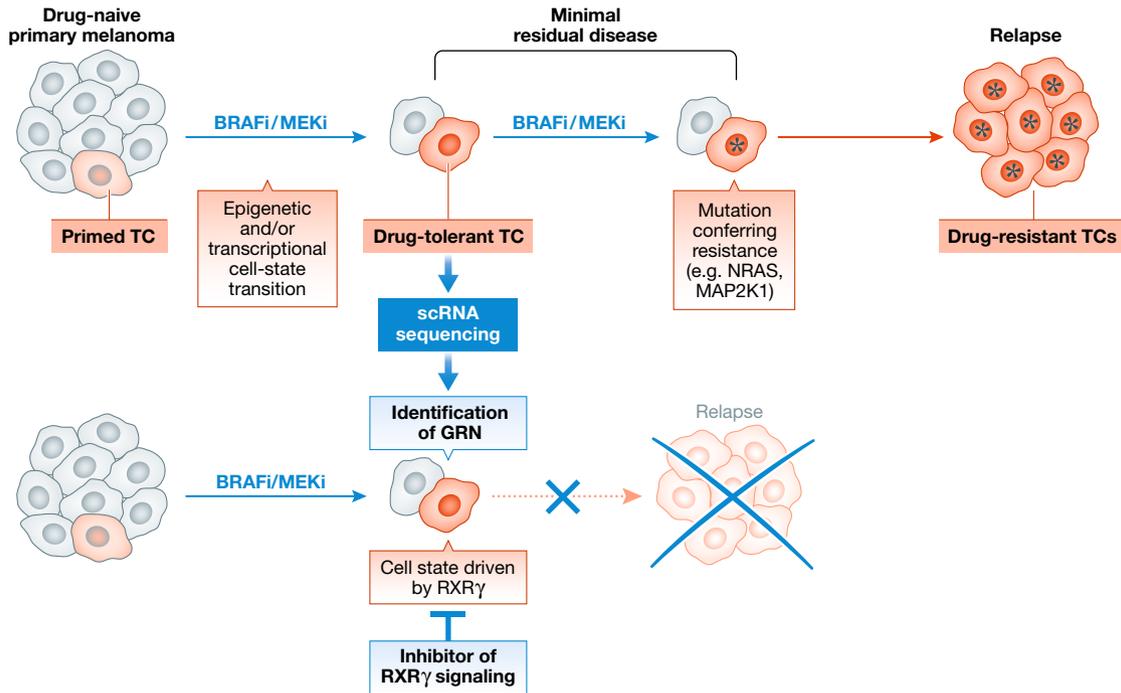
Response to therapy

Despite the progress that has been achieved in cancer therapy, thanks to the development of targeted therapies (e.g., directed against oncoproteins such as EGFR, BRAF, and ABL), resistance to therapy is still a major problem in clinical oncology. Under drug pressure, selection or acquisition of de novo mutations that provide TCs with survival/growth advantages has been proposed to be the main mechanisms underpinning acquired resistance. However, it has become clear that, along with mutational events, non-genetic adaptations are contributing to resistance to therapy. Like mutational events, some transcriptional programs, including EMT, enhanced tyrosine kinase pathways, and reduced immune response, are associated with resistance to therapy. These programs can pre-exist in tumors and be selected by the treatment or de novo acquired (Brady et al, 2017; Kim et al, 2018) (Marine et al, 2020; Shen et al, 2020). Within a drug-exposed tumor, a small fraction of cells, called minimal residual disease (MRD), can persist and remain undetected for long time, even years, before giving rise to tumor relapse. MRD is often composed of drug-tolerant persister (DTP) cells that rewire epigenetic and transcriptional programs to escape cell death. DTP cells were first identified in vitro as a small subpopulation of cells that enter a quiescent state upon drug exposure.

In EGFR-mutated NSCLC cell lines that are sensitive to the EGFR inhibitor erlotinib, rare TCs survive and enter a reversible IGF-1R-mediated drug-tolerant state that does not rely on increased drug efflux. Interestingly, this state is promoted by complex epigenetic events as suggested by increased histone demethylase KDM5A activity and sensitivity to the histone deacetylase HDAC inhibition (Sharma et al, 2010). KDM5A/B are involved in fulvestrant (anti-estrogen therapy) resistance in ER$^+$ breast cancer cells, as CRISPR-Cas9-mediated knockout of these genes and the use of KDM5 family inhibitors restore drug sensitivity to endocrine therapy. Moreover, KDM5B elevated expression correlates with higher transcriptional heterogeneity, likely contributing to therapy resistance (Hinohara et al, 2018).

Induction of a reversible slow-cycling drug-tolerant state seems to be a general survival mechanism in response to a wide range of treatments in different cancers. For example, basal cell carcinomas (BCC), whose survival relies on aberrant activation of the Hedgehog pathway, can be treated with HH inhibitors such as vismodegib. It has been observed in mouse models and patients that, despite clinical regression of BCC, some persisters survive HH inhibitor by a non-genetic switch to a quiescent cell state. These cells are the root of relapse as they resume proliferation after treatment discontinuation. Moreover, they display activation of the Wnt pathway and the combination of vismodegib and Wnt inhibitors eradicates DTCs and prevents tumor relapse (Biehs et al, 2018; Sanchez-Danes et al, 2018).

Recently, it has been demonstrated that persister cells derived from a variety of cancers challenged with different drugs depend on glutathione peroxidase 4 (GPX4) to escape lipid peroxides-mediated ferroptosis, an oxidative stress-mediated type of programmed cell death. It is noteworthy that GPX4 ablation elicits ferroptosis in vitro and prevents relapse in mice. Thus, GPX4 dependency observed in both drug-tolerant and high mesenchymal cell states represents a promising approach to achieve definitive cure. In prostate cancer, it has been demonstrated that androgen receptor (AR)-inhibiting therapy results in a drug-tolerant state characterized by an increased accumulation of phospholipids that underlies enhanced peroxidation and sensitivity to ferroptosis. Current GPX4 inhibitors display poor pharmacokinetic properties in vivo and more effective molecules are needed (Hangauer et al, 2017; Tousignant et al, 2020).

The advent of single-cell technologies has prompted researchers to isolate and characterize TCs from residual tissue in both human biopsies and patient-derived models. In a BRAF-mutant PDX model, melanoma cells can adopt different strategies to tolerate the cytotoxic effects of targeted therapy (BRAF/MEK inhibition) by acquiring new transcriptional identities that lead to the emergence of differentiated, starvation-like and de-differentiated states. A proportion of de-differentiated cells display neural crest stem cell (NCSC) traits and are likely to be the TC population that mediate tumor relapse (Fig 3). Importantly, the different tumor states within the residual tissue display also spatial heterogeneity which might reflect the reciprocal interactions between tumor subpopulations and stromal cells. Importantly, the NCSC state is controlled by the retinoid X receptor-γ (RXRG), which is pharmacologically targeted resulting in delayed tumor relapse (Rambow et al, 2018). Following the NCSC state depletion, all tumors that relapse acquired resistance-conferring mutations (Marin-Bejar et al, 2021). Another study in BRAF-mutated melanoma patient-derived cells treated with a BRAF inhibitor showed that the acquisition of the resistant state is a multi-step process. This work demonstrated that rare subpopulations are

Figure 3. Single-cell approaches enable the identification and targeting of tumor cell states responsible for therapy resistance.

Anti-cancer therapy induces epigenetic and/or transcriptional transition states that lead to the emergence of drug-tolerant cells (DTCs), which upon prolonged drug exposure eventually develop mutations conferring resistance. Single-cell RNA sequencing allows to identify and target these drug-tolerant states, thus delaying/hampering tumor relapse. Here we show the mechanisms that have been described in melanoma. GRN: gene regulatory network; RXRγ: Retinoid X Receptor gamma; TC: tumor cell.

in a transient transcriptionally "primed" state characterized by high resistance marker expression (e.g., EGFR, NGFR) before drug exposure, and become stably resistant upon treatment through epigenetic reprogramming. Loss-of-function CRISPR-Cas9 genetic screening targeting around two-thousand genes led to the identification of modulators of the primed state. Some of these factors, such as SOX10 and MITF, are well-known inducers of resistance to BRAF inhibitors in melanoma, while some others are unknown, among which the H3K79 methyltransferase DOT1L, whose inhibition drives higher propensity to resistance. Therefore, a slow-cycling population can even exist within a highly proliferating melanoma and its enrichment upon exposure to therapy can be at the basis of resistance. These apparently contrasting observations are most probably the consequence of high intra-tumor and inter-patient heterogeneity (Shaffer et al, 2017; Torre et al, 2021). Analysis of the MRD following targeted therapy in NSCLC biopsies revealed the existence of a low proliferating population presenting alveolar cell traits (Maynard et al, 2020). In the lungs, both type I and II alveolar cells respond to different types of stress and injury and participate in tissue repair (Desai et al, 2014). Thus, MRD TCs hijack alveolar cell repair programs to escape cell death. This specific state is driven by genes associated with the WNT/β-catenin signaling including SUSD2 and CAV1.

Isolation and profiling of CTCs provide useful insights in the understanding of resistance mechanisms in patients undergoing therapy. Retrospective analyses of CTCs from prostate cancer patients showed high intra-patient heterogeneity of the AR pathway, with CTCs harboring one or more AR gene splice variants associated with resistance to anti-androgen therapies. Moreover, comparison of CTCs isolated from drug-naïve versus treated patients (exhibiting tumor progression during therapy) identified non-canonical Wnt signaling as one of the most enriched signatures associated with resistance onset (Miyamoto et al, 2015). CTCs isolated from ER+/HER2- breast cancer patients treated with ER antagonist are either HER2$^+$ or HER2$^-$, suggesting that therapy induces the acquisition of HER2 expression which is not linked to gene amplifications or mutations. These HER2$^+$ cells are not sensitive to antibody-mediated HER2 inhibition but show increased proliferation, whereas HER2$^-$ respond to inhibition of NOTCH signaling but do not respond to chemotherapy, suggesting that the combination of chemotherapy and NOTCH inhibitors could help to simultaneously target these two tumor states (Jordan et al, 2016).

Conclusions and perspectives

This review discusses the great power of single cell-based technologies in providing novel insights into tumor functional heterogeneity.

These studies reveal that tumors are composed of different states that are likely to accomplish different functions. Tumors invariably present a proliferative tumor state, which corresponds to the cells that actively proliferate, and which are probably the cells referred as CSCs that sustain tumor growth. The differentiated tumor state is also very frequently identified in the various tumor that have been profiled, and that resemble normal differentiated cells found in the

tissues from where the tumor arises, as tumor differentiation reuses the physiological programs of tissue differentiation. Invasive tumor states or EMT-like tumor states are frequently found and likely correspond to metastasis-initiating cells. Chemotherapy or targeted therapy is also associated with new tumor states, sometimes pre-existing or induced by the therapy and that correspond to the persister states that drive tumor relapse after therapy. Finally, other tumor states are also identified in a more tumor-specific manner.

While these states strongly suggest their functions, these assumptions have rarely been tested experimentally and directly. To get some clues, new experimental strategies are needed. One needs to identify new cell surface markers or fluorescent transgenic reporters to isolate such cells and assess their functions in different biological assays such as transplantation *in vivo*, or clonogenic and invasive assays *in vitro*. Lineage tracing experiments using an inducible recombinase coupled with fluorescent or barcoding reporters should be developed to assess the fate of these tumor states within their native microenvironment. With these new techniques, it will be possible to test experimentally the different bioinformatic predictions of the lineage trajectories that these states are giving rise to during the natural growth of the tumor, the course of the metastatic process, and following therapy.

Whereas different tools to predict the transcription factors and the target genes that regulate each tumor state have been developed, more studies will be needed to test these predictions experimentally. This can be done by individual gain- and loss-of-function studies using hypothesis-driven approaches or more unbiased genome-wide loss-of-function screenings followed by single-cell analysis, allowing to identify genes that are enriched or depleted in the different tumor states.

Another key unresolved question from these studies is the lineage plasticity of the different tumor states. Transplantation, lineage tracing, and lineage ablation studies will reveal how stable the states are, how they transit from one state to the other, what is the role of the microenvironment in regulating these states or their transition, and what are the impact of anti-cancer therapy on these states.

There is a need for a much better understanding of how the different populations of stromal cells that compose the TME control the different tumor states. New computational methods predicting ligands, receptors, and signaling pathways expressed in the various tumor states and their neighboring stromal cells will allow to predict how the different populations of stromal cells may regulate the different tumor states. These predictions will need to be tested experimentally by *in situ* characterization and by inhibiting or activating the signaling pathways predicted to regulate the tumor states. Novel methods of multiplex immunomarking or *in situ* hybridization, spatial transcriptomics, or single-cell -omics methods assessing the proteome and their post-translational modifications will be useful to visualize the interactions between stromal cells and their associated tumor states. These experiments will be key to define the specific niches controlling the different functional states.

Ultimately, these studies should lead to the identification of tumor state vulnerabilities and pharmacological interventions that block the essential functions of these various tumor states and inhibit tumor plasticity.

Single-cell approaches have paved the way to a revolutionary approach in tumor biology that will be important to improve precision medicine to tailor treatments to the specific features of each patient's tumor.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Aceto N, Bardia A, Miyamoto D, Donaldson M, Wittner B, Spencer J, Yu M, Pely A, Engstrom A, Zhu H *et al* (2014) Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 158: 1110–1122

Aiello NM, Maddipati R, Norgard RJ, Balli D, Li J, Yuan S, Yamazoe T, Black T, Sahmoud A, Furth EE *et al* (2018) EMT subtype influences epithelial plasticity and mode of cell migration. *Dev Cell* 45: 681–695.e4

Akbani R, Akdemir K, Aksoy B, Albert M, Ally A, Amin S, Arachchi H, Arora A, Auman J, Ayala B *et al* (2015) Genomic classification of cutaneous melanoma. *Cell* 161: 1681–1696

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003) Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci USA* 100: 3983–3988

Andree KC, van Dalum G, Terstappen LW (2016) Challenges in circulating tumor cell detection by the cell search system. *Mol Oncol* 10: 395–407

Baccelli I, Schneeweiss A, Riethdorf S, Stenzinger A, Schillert A, Vogel V, Klein C, Saini M, Bäuerle T, Wallwiener M *et al* (2013) Identification of a population of blood circulating tumor cells from breast cancer patients that initiates metastasis in a xenograft assay. *Nat Biotechnol* 31: 539–544

Bado IL, Zhang W, Hu J, Xu Z, Wang H, Sarkar P, Li L, Wan Y-W, Liu J, Wu W *et al* (2021) The bone microenvironment increases phenotypic plasticity of ER(+) breast cancer cells. *Dev Cell* 56: 1100–1117.e9

Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick JE, Tanner SD (2009) Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem* 81: 6813–6822

Beck B, Blanpain C (2013) Unravelling cancer stem cell potential. *Nat Rev Cancer* 13: 727–738

Beck B, Driessens G, Goossens S, Youssef KK, Kuchnio A, Caauwe A, Sotiropoulou PA, Loges S, Lapouge G, Candi A *et al* (2011) A vascular niche and a VEGF-Nrp1 loop regulate the initiation and stemness of skin tumours. *Nature* 478: 399–403

Bendall SC, Simonds EF, Qiu P, Amir E-AD D, Krutzik PO, Finck R, Bruggner RV, Melamed R, Trejo A, Ornatsky OI *et al* (2011) Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 332: 687–696

Biehs B, Dijkgraaf GJP, Piskol R, Alicke B, Boumahdi S, Peale F, Gould SE, de Sauvage FJ (2018) A cell identity switch allows residual BCC to survive Hedgehog pathway inhibition. *Nature* 562: 429–433

Birkbak NJ, McGranahan N (2020) Cancer genome evolutionary trajectories in metastasis. *Cancer Cell* 37: 8–19

Bonnet D, Dick JE (1997) Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat Med* 3: 730–737

Brabletz S, Schuhwerk H, Brabletz T, Stemmler MP (2021) Dynamic EMT: a multi-tool for tumor progression. *EMBO J* 40: e108647

Brady SW, McQuerry JA, Qiao YI, Piccolo SR, Shrestha G, Jenkins DF, Layer RM, Pedersen BS, Miller RH, Esch A *et al* (2017) Combating subclonal evolution of resistant cancer phenotypes. *Nat Commun* 8: 1231

Brennan C, Verhaak R, McKenna A, Campos B, Noushmehr H, Salama S, Zheng S, Chakravarty D, Sanborn J, Berman S *et al* (2013) The somatic genomic landscape of glioblastoma. *Cell* 155: 462–477

Browaeys R, Saelens W, Saeys Y (2020) NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat Methods* 17: 159–162

Brown JA, Yonekubo Y, Hanson N, Sastre-Perona A, Basin A, Rytlewski JA, Dolgalev I, Meehan S, Tsirigos A, Beronja S *et al* (2017) TGF-beta-induced quiescence mediates chemoresistance of tumor-propagating cells in squamous cell carcinoma. *Cell Stem Cell* 21: 650–664. e658

Budnik B, Levy E, Harmange G, Slavov N (2018) SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol* 19: 161

Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature* 523: 486–490

Cancer Genome Atlas N (2012a) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 487: 330–337

Cancer Genome Atlas N (2012b) Comprehensive molecular portraits of human breast tumours. *Nature* 490: 61–70

Celia-Terrassa T, Kang Y (2016) Distinctive properties of metastasis-initiating cells. *Genes Dev* 30: 892–908

Chen MT, Sun HF, Zhao Y, Fu WY, Yang LP, Gao SP, Li LD, Jiang HL, Jin W (2017) Comparison of patterns and prognosis among distant metastatic breast cancer patients by age groups: a SEER population-based analysis. *Sci Rep* 7: 9254

Cheng Y-H, Chen Y-C, Lin E, Brien R, Jung S, Chen Y-T, Lee W, Hao Z, Sahoo S, Min Kang H *et al* (2019) Hydro-Seq enables contamination-free high-throughput single-cell RNA-sequencing for circulating tumor cells. *Nat Commun* 10: 2163

Cheow LF, Quake SR, Burkholder WF, Messerschmidt DM (2015) Multiplexed locus-specific analysis of DNA methylation in single cells. *Nat Protoc* 10: 619–631

Cildir G, Yip KH, Pant H, Tergaonkar V, Lopez AF, Tumes DJ (2021) Understanding mast cell heterogeneity at single cell resolution. *Trends Immunol* 42: 523–535

Cook DP, Vanderhyden BC (2020) Context specificity of the EMT transcriptional response. *Nat Commun* 11: 2142

Couturier CP, Ayyadhury S, Le PU, Nadaf J, Monlong J, Riva G, Allache R, Baig S, Yan X, Bourgey M *et al* (2020) Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat Commun* 11: 3406

Cusanovich DA, Daza R, Adey A, Pliner HA, Christiansen L, Gunderson KL, Steemers FJ, Trapnell C, Shendure J (2015) Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* 348: 910–914

Davis RT, Blake K, Ma D, Gabra MBI, Hernandez GA, Phung AT, Yang Y, Maurer D, Lefebvre AEYT, Alshetaiwi H *et al* (2020) Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing. *Nat Cell Biol* 22: 310–320

Desai TJ, Brownfield DG, Krasnow MA (2014) Alveolar progenitor and stem cells in lung development, renewal and cancer. *Nature* 507: 190–194

Deshmukh AP, Vasaikar SV, Tomczak K, Tripathi S, den Hollander P, Arslan E, Chakraborty P, Soundararajan R, Jolly MK, Rai K *et al* (2021) Identification of EMT signaling cross-talk and gene regulatory networks by single-cell RNA sequencing. *Proc Natl Acad Sci USA* 118: e2102050118

Dotto GP (2009) Crosstalk of Notch with p53 and p63 in cancer growth control. *Nat Rev Cancer* 9: 587–595

Dudgeon C, Harris CR, Chen Y, Ghaddar B, Sharma A, Shah MM, Roberts AI, Casabianca A, Collisson EA, Balachandran VP *et al* (2020) A novel model of pancreatic cancer dormancy reveals mechanistic insights and a dormancy gene signature with human relevance. *bioRxiv* https://doi.org/10.1101/2020.04.13.037374 [PREPRINT]

Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R (2020) Cell PhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 15: 1484–1506

Eng C-H, Lawson M, Zhu Q, Dries R, Koulena N, Takei Y, Yun J, Cronin C, Karp C, Yuan G-C *et al* (2019) Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* 568: 235–239

Engstrand J, Nilsson H, Stromberg C, Jonas E, Freedman J (2018) Colorectal cancer liver metastases - a population-based study on incidence, management and survival. *BMC Cancer* 18: 78

Eslami SZ, Cortes-Hernandez LE, Alix-Panabieres C (2020) Epithelial cell adhesion molecule: an anchor to isolate clinically relevant circulating tumor cells. *Cells* 9: 1836

Filbin MG, Tirosh I, Hovestadt V, Shaw ML, Escalante LE, Mathewson ND, Neftel C, Frank N, Pelton K, Hebert CM *et al* (2018) Developmental and oncogenic programs in H3K27M gliomas dissected by single-cell RNA-seq. *Science* 360: 331–335

Fillmore CM, Kuperwasser C (2008) Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res* 10: R25

Gao Y, Bado I, Wang H, Zhang W, Rosen JM, Zhang XH (2019) Metastasis organotropism: redefining the congenial soil. *Dev Cell* 49: 375–391

Giesen C, Wang HAO, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B, Schüffler PJ, Grolimund D, Buhmann JM, Brandt S *et al* (2014) Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* 11: 417–422

Gillen AE, Riemondy KA, Amani V, Griesinger AM, Gilani A, Venkataraman S, Madhavan K, Prince E, Sanford B, Hankinson TC *et al* (2020) Single-cell RNA sequencing of childhood ependymoma reveals neoplastic cell subpopulations that impact molecular classification and etiology. *Cell Rep* 32: 108023

Gojo J, Englinger B, Jiang LI, Hübner JM, Shaw ML, Hack OA, Madlener S, Kirchhofer D, Liu I, Pyrdol J *et al* (2020) Single-cell RNA-seq reveals cellular hierarchies and impaired developmental trajectories in pediatric ependymoma. *Cancer Cell* 38: 44–59.e9

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

Hangauer MJ, Viswanathan VS, Ryan MJ, Bole D, Eaton JK, Matov A, Galeas J, Dhruv HD, Berens ME, Schreiber SL *et al* (2017) Drug-tolerant persister cancer cells are vulnerable to GPX4 inhibition. *Nature* 551: 247–250

Hara T, Chanoch-Myers R, Mathewson ND, Myskiw C, Atta L, Bussema L, Eichhorn SW, Greenwald AC, Kinker GS, Rodman C *et al* (2021) Interactions between cancer cells and immune cells drive transitions to mesenchymal-like states in glioblastoma. *Cancer Cell* 39: 779–792.e11

Heppner GH (1984) Tumor heterogeneity. *Cancer Res* 44: 2259–2265

Hinohara K, Wu H-J, Vigneau S, McDonald TO, Igarashi KJ, Yamamoto KN, Madsen T, Fassl A, Egri SB, Papanastasiou M *et al* (2018) KDM5 histone demethylase activity links cellular transcriptomic heterogeneity to therapeutic resistance. *Cancer Cell* 34: 939–953.e9

Hong X, Roh W, Sullivan RJ, Wong KHK, Wittner BS, Guo H, Dubash TD, Sade-Feldman M, Wesley B, Horwitz E *et al* (2021) The lipogenic regulator SREBP2 induces transferrin in circulating melanoma cells and suppresses ferroptosis. *Cancer Discov* 11: 678–695

Hovestadt V, Smith KS, Bihannic L, Filbin MG, Shaw ML, Baumgartner A, DeWitt JC, Groves A, Mayr L, Weisman HR *et al* (2019) Resolving

medulloblastoma cellular architecture by single-cell genomics. *Nature* 572: 74–79

Hu Z, Artibani M, Alsaadi A, Wietek N, Morotti M, Shi T, Zhong Z, Santana Gonzalez L, El-Sahhar S, KaramiNejadRanjbar M *et al* (2020) The repertoire of serous ovarian cancer non-genetic heterogeneity revealed by single-cell sequencing of normal fallopian tube epithelial cells. *Cancer Cell* 37: 226–242.e7

Hyun KA, Koo GB, Han H, Sohn J, Choi W, Kim SI, Jung HI, Kim YS (2016) Epithelial-to-mesenchymal transition leads to loss of EpCAM and different physical properties in circulating tumor cells from metastatic breast cancer. *Oncotarget* 7: 24677–24687

Ireland AS, Micinski AM, Kastner DW, Guo B, Wait SJ, Spainhower KB, Conley CC, Chen OS, Guthrie MR, Soltero D *et al* (2020) MYC drives temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate. *Cancer Cell* 38: 60–78. e12

Ji AL, Rubin AJ, Thrane K, Jiang S, Reynolds DL, Meyers RM, Guo MG, George BM, Mollbrink A, Bergenstrahle J *et al* (2020) Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* 182: 497–514. e422

Jordan NV, Bardia A, Wittner BS, Benes C, Ligorio M, Zheng YU, Yu M, Sundaresan TK, Licausi JA, Desai R *et al* (2016) HER2 expression identifies dynamic functional states within circulating breast cancer cells. *Nature* 537: 102–106

Kanzaki R, Pietras K (2020) Heterogeneity of cancer-associated fibroblasts: opportunities for precision medicine. *Cancer Sci* 111: 2708–2717

Karaayvaz M, Cristea S, Gillespie SM, Patel AP, Mylvaganam R, Luo CC, Specht MC, Bernstein BE, Michor F, Ellisen LW (2018) Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat Commun* 9: 3588

Karacosta LG, Anchang B, Ignatiadis N, Kimmey SC, Benson JA, Shrager JB, Tibshirani R, Bendall SC, Plevritis SK (2019) Mapping lung cancer epithelial-mesenchymal transition states and trajectories with single-cell resolution. *Nat Commun* 10: 5587

Ke R, Mignardi M, Pacureanu A, Svedlund J, Botling J, Wahlby C, Nilsson M (2013) In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* 10: 857–860

Keller L, Pantel K (2019) Unravelling tumour heterogeneity by single-cell profiling of circulating tumour cells. *Nat Rev Cancer* 19: 553–567

Keller L, Werner S, Pantel K (2019) Biology and clinical relevance of EpCAM. *Cell Stress* 3: 165–180

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

Kim C, Gao R, Sei E, Brandt R, Hartman J, Hatschek T, Crosetto N, Foukakis T, Navin NE (2018) Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. *Cell* 173: 879–893.e13

Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, Lee J-I, Suh Y-L, Ku BM, Eum HH *et al* (2020) Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat Commun* 11: 2285

Kinker GS, Greenwald AC, Tal R, Orlova Z, Cuoco MS, McFarland JM, Warren A, Rodman C, Roth JA, Bender SA *et al* (2020) Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat Genet* 52: 1208–1218

Klotz R, Thomas A, Teng T, Han SM, Iriondo O, Li L, Restrepo-Vassalli S, Wang A, Izadian N, MacKay M *et al* (2020) Circulating tumor cells exhibit metastatic tropism and reveal brain metastasis drivers. *Cancer Discov* 10: 86–103

Kumar MP, Du J, Lagoudas G, Jiao Y, Sawyer A, Drummond DC, Lauffenburger DA, Raue A (2018) Analysis of single-cell RNA-seq identifies cell-cell communication associated with tumor characteristics. *Cell Rep* 25: 1458–1468.e4

Lambert AW, Pattabiraman DR, Weinberg RA (2017) Emerging biological principles of metastasis. *Cell* 168: 670–691

Lapidot T, Sirard C, Vormoor J, Murdoch B, Hoang T, Caceres-Cortes J, Minden M, Paterson B, Caligiuri MA, Dick JE (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367: 645–648

Laughney AM, Hu J, Campbell NR, Bakhoum SF, Setty M, Lavallée V-P, Xie Y, Masilionis I, Carr AJ, Kottapalli S *et al* (2020) Regenerative lineages and immune-mediated pruning in lung cancer metastasis. *Nat Med* 26: 259–269

Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K, Zhou A, Eyob H, Balakrishnan S, Wang C-Y *et al* (2015) Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* 526: 131–135

Li H, Courtois ET, Sengupta D, Tan Y, Chen KH, Goh JJL, Kong SL, Chua C, Hon LK, Tan WS *et al* (2017) Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. *Nat Genet* 49: 708–718

Li Y, Lv Z, Zhang S, Wang Z, He L, Tang M, Pu W, Zhao H, Zhang Z, Shi Q *et al* (2020) Genetic Fate mapping of transient cell fate reveals N-cadherin activity and function in tumor metastasis. *Dev Cell* 54: 593–607.e5

Lim E, Vaillant F, Wu DI, Forrest NC, Pal B, Hart AH, Asselin-Labat M-L, Gyorki DE, Ward T, Partanen A *et al* (2009) Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* 15: 907–913

Lin W, Noel P, Borazanci EH, Lee J, Amini A, Han IW, Heo JS, Jameson GS, Fraser C, Steinbach M *et al* (2020) Single-cell transcriptome analysis of tumor and stromal compartments of pancreatic ductal adenocarcinoma primary tumors and metastatic lesions. *Genome Med* 12: 80

Linde N, Fluegen G, Aguirre-Ghiso JA (2016) The relationship between dormant cancer cells and their microenvironment. *Adv Cancer Res* 132: 45–71

Liu X, Ding C, Tan W, Zhang A (2020) Medulloblastoma: Molecular understanding, treatment evolution, and new developments. *Pharmacol Ther* 210: 107516

Liu X, Li J, Cadilha BL, Markota A, Voigt C, Huang Z, Lin PP, Wang DD, Dai J, Kranz G *et al* (2019) Epithelial-type systemic breast carcinoma cells with a restricted mesenchymal transition are a major source of metastasis. *Sci Adv* 5: eaav4275

Longo SK, Guo MG, Ji AL, Khavari PA (2021) Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet* 22: 627–644

Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L (2014) Single-cell in situ RNA profiling by sequential hybridization. *Nat Methods* 11: 360–361

Luo Y, Dallaglio K, Chen Y, Robinson WA, Robinson SE, McCarter MD, Wang J, Gonzalez R, Thompson DC, Norris DA *et al* (2012) ALDH1A isozymes are markers of human melanoma stem cells and potential therapeutic targets. *Stem Cells* 30: 2100–2113

Ma A, McDermaid A, Xu J, Chang Y, Ma Q (2020) Integrative methods and practical challenges for single-cell multi-omics. *Trends Biotechnol* 38: 1007–1022

Mallory XF, Edrisi M, Navin N, Nakhleh L (2020) Methods for copy number aberration detection from single-cell DNA-sequencing data. *Genome Biol* 21: 208

Marin-Bejar O, Rogiers A, Dewaele M, Femel J, Karras P, Pozniak J, Bervoets G, Van Raemdonck N, Pedri D, Swings T *et al* (2021) Evolutionary

predictability of genetic versus nongenetic resistance to anticancer drugs in melanoma. *Cancer Cell* 39: 1135–1149.e8

Marine JC, Dawson SJ, Dawson MA (2020) Non-genetic mechanisms of therapeutic resistance in cancer. *Nat Rev Cancer* 20: 743–756

Massague J, Ganesh K (2021) Metastasis-initiating cells and ecosystems. *Cancer Discov* 11: 971–994

Massalha H, Bahar Halpern K, Abu-Gazala S, Jana T, Massasa EE, Moor AE, Buchauer L, Rozenberg M, Pikarsky E, Amit I *et al* (2020) A single cell atlas of the human liver tumor microenvironment. *Mol Syst Biol* 16: e9682

Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, Yu EA, Schenk EL, Tan W, Zee A *et al* (2020) Therapy-induced evolution of human lung cancer revealed by single-cell RNA sequencing. *Cell* 182: 1232–1251.e22

McFaline-Figueroa JL, Hill AJ, Qiu X, Jackson D, Shendure J, Trapnell C (2019) A pooled single-cell genetic screen identifies regulatory checkpoints in the continuum of the epithelial-to-mesenchymal transition. *Nat Genet* 51: 1389–1398

Miyamoto DT, Zheng YU, Wittner BS, Lee RJ, Zhu H, Broderick KT, Desai R, Fox DB, Brannigan BW, Trautwein J *et al* (2015) RNA-seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. *Science* 349: 1351–1356

Mohme M, Riethdorf S, Pantel K (2017) Circulating and disseminated tumour cells - mechanisms of immune surveillance and escape. *Nat Rev Clin Oncol* 14: 155–167

Molyneux G, Geyer FC, Magnay F-A, McCarthy A, Kendrick H, Natrajan R, MacKay A, Grigoriadis A, Tutt A, Ashworth A *et al* (2010) BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* 7: 403–417

Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, Hajdu CH, Simeone DM, Yanai I (2020) Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat Biotechnol* 38: 333–342

Moreb JS (2008) Aldehyde dehydrogenase as a marker for stem cells. *Curr Stem Cell Res Ther* 3: 237–246

Nagano T, Lubling Y, Yaffe E, Wingett SW, Dean W, Tanay A, Fraser P (2015) Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. *Nat Protoc* 10: 1986–2003

Nassar D, Blanpain C (2016) Cancer stem cells: basic concepts and therapeutic implications. *Annu Rev Pathol* 11: 47–76

Navin NE (2015) The first five years of single-cell cancer genomics and beyond. *Genome Res* 25: 1499–1507

Nayar U, Cohen O, Kapstad C, Cuoco MS, Waks AG, Wander SA, Painter C, Freeman S, Persky NS, Marini L *et al* (2019) Acquired HER2 mutations in ER(+) metastatic breast cancer confer resistance to estrogen receptor-directed therapies. *Nat Genet* 51: 207–216

Neftel C, Laffy J, Filbin MG, Hara T, Shore ME, Rahme GJ, Richman AR, Silverbush D, Shaw ML, Hebert CM *et al* (2019) An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* 178: 835–849.e21

Norris JL, Caprioli RM (2013) Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chem Rev* 113: 2309–2342

Oshimori N, Oristian D, Fuchs E (2015) TGF-beta promotes heterogeneity and drug resistance in squamous cell carcinoma. *Cell* 160: 963–976

Pal B, Chen Y, Vaillant F, Capaldo BD, Joyce R, Song X, Bryant VL, Penington JS, Di Stefano L, Tubau Ribera N *et al* (2021) A single-cell RNA expression atlas of normal, preneoplastic and tumorigenic states in the human breast. *EMBO J* 40: e107333

Pantel K, Brakenhoff RH (2004) Dissecting the metastatic cascade. *Nat Rev Cancer* 4: 448–456

Pascual G, Avgustinova A, Mejetta S, Martín M, Castellanos A, Attolini C-O, Berenguer A, Prats N, Toll A, Hueto JA *et al* (2017) Targeting metastasis-initiating cells through the fatty acid receptor CD36. *Nature* 541: 41–45

Pastushenko I, Blanpain C (2019) EMT transition states during tumor progression and metastasis. *Trends Cell Biol* 29: 212–226

Pastushenko I, Brisebarre A, Sifrim A, Fioramonti M, Revenco T, Boumahdi S, Van Keymeulen A, Brown D, Moers V, Lemaire S *et al* (2018) Identification of the tumour transition states occurring during EMT. *Nature* 556: 463–468

Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL *et al* (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344: 1396–1401

Pece S, Tosoni D, Confalonieri S, Mazzarol G, Vecchi M, Ronzoni S, Bernard L, Viale G, Pelicci PG, Di Fiore PP (2010) Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell* 140: 62–73

Peng J, Sun B-F, Chen C-Y, Zhou J-Y, Chen Y-S, Chen H, Liu L, Huang D, Jiang J, Cui G-S *et al* (2019) Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res* 29: 725–738

Pine AR, Cirigliano SM, Nicholson JG, Hu Y, Linkous A, Miyaguchi K, Edwards L, Singhania R, Schwartz TH, Ramakrishna R *et al* (2020) Tumor microenvironment is critical for the maintenance of cellular states found in primary glioblastomas. *Cancer Discov* 10: 964–979

Prager BC, Xie Q, Bao S, Rich JN (2019) Cancer stem cells: the architects of the tumor ecosystem. *Cell Stem Cell* 24: 41–53

Puram SV, Tirosh I, Parikh AS, Patel AP, Yizhak K, Gillespie S, Rodman C, Luo CL, Mroz EA, Emerick KS *et al* (2017) Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* 171: 1611–1624.e24

Quinn JJ, Jones MG, Okimoto RA, Nanjo S, Chan MM, Yosef N, Bivona TG, Weissman JS (2021) Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science* 371: eabc1944

Rambow F, Rogiers A, Marin-Bejar O, Aibar S, Femel J, Dewaele M, Karras P, Brown D, Chang YH, Debiec-Rychter M *et al* (2018) Toward minimal residual disease-directed therapy in melanoma. *Cell* 174: 843–855.e19

Revenco T, Nicodème A, Pastushenko I, Sznurkowska MK, Latil M, Sotiropoulou PA, Dubois C, Moers V, Lemaire S, de Maertelaer V *et al* (2019) Context dependency of epithelial-to-mesenchymal transition for metastasis. *Cell Rep* 29: 1458–1468.e3

Risson E, Nobre AR, Maguer-Satta V, Aguirre-Ghiso JA (2020) The current paradigm and challenges ahead for the dormancy of disseminated tumor cells. *Nat Cancer* 1: 672–680

Rodriguez-Meira A, Buck G, Clark S-A, Povinelli BJ, Alcolea V, Louka E, McGowan S, Hamblin A, Sousos N, Barkas N *et al* (2019) Unravelling intratumoral heterogeneity through high-sensitivity single-cell mutational analysis and parallel RNA sequencing. *Mol Cell* 73: 1292–1305.e8

Roesch A, Fukunaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PA, Vultur A, Basu D, Gimotty P, Vogt T, Herlyn M (2010) A temporally distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* 141: 583–594

Rotem A, Ram O, Shoresh N, Sperling RA, Goren A, Weitz DA, Bernstein BE (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat Biotechnol* 33: 1165–1172

Sanchez-Danes A, Blanpain C (2018) Deciphering the cells of origin of squamous cell carcinomas. *Nat Rev Cancer* 18: 549–561

Sánchez-Danés A, Larsimont J-C, Liagre M, Muñoz-Couselo E, Lapouge G, Brisebarre A, Dubois C, Suppa M, Sukumaran V, del Marmol V et al (2018) A slow-cycling LGR5 tumour population mediates basal cell carcinoma relapse after therapy. *Nature* 562: 434−438

Schober M, Fuchs E (2011) Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF-beta and integrin/focal adhesion kinase (FAK) signaling. *Proc Natl Acad Sci USA* 108: 10544−10549

Sha Y, Wang S, Bocci F, Zhou P, Nie Q (2020a) Inference of intercellular communications and multilayer gene-regulations of epithelial-mesenchymal transition from single-cell transcriptomic data. *Front Genet* 11: 604585

Sha Y, Wang S, Zhou P, Nie Q (2020b) Inference and multiscale model of epithelial-to-mesenchymal transition via single-cell transcriptomic data. *Nucleic Acids Res* 48: 9505−9520

Shaffer SM, Dunagin MC, Torborg SR, Torre EA, Emert B, Krepler C, Beqiri M, Sproesser K, Brafford PA, Xiao M et al (2017) Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature* 546: 431−435

Sharma SV, Lee DY, Li B, Quinlan MP, Takahashi F, Maheswaran S, McDermott U, Azizian N, Zou L, Fischbach MA et al (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 141: 69−80

Shen S, Vagner S, Robert C (2020) Persistent cancer cells: the deadly survivors. *Cell* 183: 860−874

Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J, Nikolskaya T, Serebryiskaya T, Beroukhim R, Hu M et al (2007) Molecular definition of breast tumor heterogeneity. *Cancer Cell* 11: 259−273

Simeonov KP, Byrns CN, Clark ML, Norgard RJ, Martin B, Stanger BZ, Shendure J, McKenna A, Lengner CJ (2021) Single-cell lineage tracing of metastatic cancer reveals selection of hybrid EMT states. *Cancer Cell* 39: 1150−1162.e9

Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, Giacomello S, Asp M, Westholm JO, Huss M et al (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353: 78−82

Szczerba BM, Castro-Giner F, Vetter M, Krol I, Gkountela S, Landin J, Scheidmann MC, Donato C, Scherrer R, Singer J et al (2019) Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature* 566: 553−557

de The H (2018) Differentiation therapy revisited. *Nat Rev Cancer* 18: 117−127

Ting D, Wittner B, Ligorio M, Vincent Jordan N, Shah A, Miyamoto D, Aceto N, Bersani F, Brannigan B, Xega K et al (2014) Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep* 8: 1905−1918

Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G et al (2016a) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352: 189−196

Tirosh I, Venteicher AS, Hebert C, Escalante LE, Patel AP, Yizhak K, Fisher JM, Rodman C, Mount C, Filbin MG et al (2016b) Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. *Nature* 539: 309−313

Torre EA, Arai E, Bayatpour S, Jiang CL, Beck LE, Emert BL, Shaffer SM, Mellis IA, Fane ME, Alicea GM et al (2021) Genetic screening for single-cell variability modulators driving therapy resistance. *Nat Genet* 53: 76−85

Tousignant KD, Rockstroh A, Poad BLJ, Talebi A, Young RSE, Taherian Fard A, Gupta R, Zang T, Wang C, Lehman ML et al (2020) Therapy-induced lipid uptake and remodeling underpin ferroptosis hypersensitivity in prostate cancer. *Cancer Metab* 8: 11

Van Keymeulen A, Lee MY, Ousset M, Brohée S, Rorive S, Giraddi RR, Wuidart A, Bouvencourt G, Dubois C, Salmon I et al (2015) Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* 525: 119−123

Vaughan AE, Brumwell AN, Xi Y, Gotts JE, Brownfield DG, Treutlein B, Tan K, Tan V, Liu FC, Looney MR et al (2015) Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. *Nature* 517: 621−625

Vendramin R, Litchfield K, Swanton C (2021) Cancer evolution: darwin and beyond. *EMBO J* 40: e108389

Venteicher AS, Tirosh I, Hebert C, Yizhak K, Neftel C, Filbin MG, Hovestadt V, Escalante LE, Shaw ML, Rodman C et al (2017) Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. *Science* 355: eaai8478

Verhaak RGW, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, Miller CR, Ding LI, Golub T, Mesirov JP et al (2010) Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 17: 98−110

Vieugue P, Blanpain C (2020) Recording EMT activity by lineage tracing during metastasis. *Dev Cell* 54: 567−569

Wouters J, Kalender-Atak Z, Minnoye L, Spanier KI, De Waegeneer M, Bravo González-Blas C, Mauduit D, Davie K, Hulselmans G, Najem A et al (2020) Robust gene expression programs underlie recurrent cell states and phenotype switching in melanoma. *Nat Cell Biol* 22: 986−998

Yao J, Cui Q, Fan W, Ma Y, Chen Y, Liu T, Zhang X, Xi Y, Wang C, Peng L et al (2020) Single-cell transcriptomic analysis in a mouse model deciphers cell transition states in the multistep development of esophageal cancer. *Nat Commun* 11: 3715

Yates LR, Knappskog S, Wedge D, Farmery JHR, Gonzalez S, Martincorena I, Alexandrov LB, Van Loo P, Haugland HK, Lilleng PK et al (2017) Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* 32: 169−184.e7

Yu M, Bardia A, Wittner BS, Stott SL, Smas ME, Ting DT, Isakoff SJ, Ciciliano JC, Wells MN, Shah AM et al (2013) Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* 339: 580−584

Yuan X, Wang J, Huang Y, Shangguan D, Zhang P (2021) Single-Cell profiling to explore immunological heterogeneity of tumor microenvironment in breast cancer. *Front Immunol* 12: 643692

Zhang W, Bado IL, Hu J, Wan Y-W, Wu L, Wang H, Gao Y, Jeong H-H, Xu Z, Hao X et al (2021) The bone microenvironment invigorates metastatic seeds for further dissemination. *Cell* 184: 2471−2486.e20

Zhao J, Guo C, Xiong F, Yu J, Ge J, Wang H, Liao Q, Zhou Y, Gong Q, Xiang BO et al (2020) Single cell RNA-seq reveals the landscape of tumor and infiltrating immune cells in nasopharyngeal carcinoma. *Cancer Lett* 477: 131−143

Zhu C, Gao Y, Guo H, Xia B, Song J, Wu X, Zeng H, Kee K, Tang F, Yi C (2017) Single-cell 5-formylcytosine landscapes of mammalian early embryos and ESCs at single-base resolution. *Cell Stem Cell* 20: 720−731.e5

Zuo W, Zhang T, Wu DZ, Guan SP, Liew A-A, Yamamoto Y, Wang X, Lim SJ, Vincent M, Lessard M et al (2015) p63(+)Krt5(+) distal airway stem cells are essential for lung regeneration. *Nature* 517: 616−620