HHS Public Access

Author manuscript

Nat Cell Biol. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

Nat Cell Biol. 2018 December ; 20(12): 1349–1360. doi:10.1038/s41556-018-0236-7.

Tumour heterogeneity and metastasis at single-cell resolution

Devon A. Lawson[1,2,*], **Kai Kessenbrock**[2,3,*], **Ryan T. Davis**[1], **Nicholas Pervolarakis**[3,4], and **Zena Werb**[5]

[1]Department of Physiology and Biophysics, University of California, Irvine, CA, USA.

[2]Chao Family Comprehensive Cancer Center, University of California, Irvine, CA, USA.

[3]Department of Biological Chemistry, University of California, Irvine, CA, USA.

[4]Center for Complex Biological Systems, University of California, Irvine, CA, USA.

[5]Department of Anatomy, and Helen Diller Comprehensive Cancer Center, University of California, San Francisco, CA, USA.

Abstract

Tumours comprise a heterogeneous collection of cells with distinct genetic and phenotypic properties that can differentially promote progression, metastasis and drug resistance. Emerging single-cell technologies provide a new opportunity to profile individual cells within tumours and investigate what roles they play in these processes. This Review discusses key technological considerations for single-cell studies in cancer, new findings using single-cell technologies and critical open questions for future applications.

Heterogeneity is pervasive in human cancer and manifests as morphological differences between cells or distinct karyotypic patterns, protein and biomarker expression levels and genetic profiles[1,2]. Tumours are complex ecosystems of malignant cells surrounded by non-malignant stroma, including fibroblasts, endothelial cells and infiltrating immune cells[3–5]. Intratumour heterogeneity arises through various mechanisms (Fig. 1). In clonal evolution models, stochastic accumulation of mutations through genomic instability results in increasing genetic diversity, with the tumour acquiring subclones with distinct genotypes over time[6]. Heterogeneity is also generated through cellular differentiation. In cancer stem cell (CSC) models, cancers are hierarchically organized with a stem cell-like population, sustaining tumour growth through self-renewal and differentiation[7]. The tumour microenvironment also generates intratumour heterogeneity by exerting different selective pressures in distinct regions of the tumour[8–11]. These models are not mutually exclusive and act together to create a complex system with multiple layers of heterogeneity established by the distinct genetic, epigenetic, transcriptomic, proteomic and functional properties of different cells.

Reprints and permissions information is available at www.nature.com/reprints.

*__Correspondence__ should be addressed to D.A.L. dalawson@uci.edu; or K.K. kai.kessenbrock@uci.edu.

Competing interests
The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Nevertheless, most cancer research and therapy decisions are carried out at the whole-population level. Standard treatment strategies target a single receptor or pathway, treating cancer as a homogenous disease. Even new precision medicine programmes, such as the NCI-MATCH (National Cancer Institute—Molecular Analysis for Therapy Choice) trial, which genetically profiles individual patient tumours to determine the most appropriate targeted therapy, do not consider the number of cells that express the targeted variant and only require it to be detectable above background[12]. This therapeutic approach may fail for many reasons: if the variant is not critical to drive tumour growth or not expressed in the tumour-promoting cell populations; if some cell populations have additional driver or resistance mutations; or if tumour growth, viability or resistance is encoded at the non-genetic level.

However, technologies for interrogating the whole genome, transcriptome, epigenome and proteome in single cells are maturing. Advances in accuracy, throughput, automation, computational analysis and cost provide the potential to profile thousands of cells from an individual tumour. A first goal in cancer is to characterize the extent of intratumour heterogeneity in individual tumours, at various regulatory levels, from genotype to phenotype, and to spatially localize cell populations within tumours. Subsequently, understanding the function and effect of different cell populations on tumorigenesis, including which features promote tumour initiation, progression or drug resistance, will also be key. Functional characterization will be particularly challenging, as there is no clear method for extrapolating cell function from large-scale 'omics' data aside from traditional experimental interrogation. In the long term, new insights may be translated to the clinic, for example, to enable tumour composition analysis for diagnostics and therapeutic assignment, or to identify pre-existing drug-resistant subclones prior to treatment. In this Review, we discuss important technological considerations for experimental designs in cancer research, review single-cell studies that have provided new insights in tumour biology and present open questions for future single-cell applications.

Technological considerations for single-cell studies of cancer

Single-cell technologies have advanced rapidly in the past several years. Currently available protocols vary in cell capture method, library preparation chemistry and throughput (Table 1) (reviewed in refs [13–16]). Most protocols require single-cell suspension, so the first critical consideration is optimizing tumour dissociation to generate a cell suspension that is fully representative of the intact tumour in terms of cell populations, their frequencies and expression programmes. Digestion of solid tumours eliminates spatial information and can obscure the true programme of individual cells[17,18]. Although there is no consensus for how to measure these profound effects, cellular diversity after dissociation can be analysed by flow cytometry for known cell types or markers for the specific tumour type. Populations are typically validated by follow-up analyses in situ, but this approach only confirms their existence and does not determine whether all cell populations in the tumour were accounted for after cell dissociation. Ultimately, identification of the same populations using different protocols would increase confidence in the results.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Technologies for transcriptome analysis are the most advanced and have been used to profile CSCs, map differentiation trajectories, describe drug resistance programmes and define the immune infiltrate in tumours[19–21]. The first single-cell transcriptome technologies utilized microfluidics to capture cells followed by multiplex quantitative PCR (qPCR) for selected genes, but most recent studies favour single-cell RNA sequencing (scRNA-seq) to enable assessment of the entire transcriptome. Selection of the most appropriate scRNA-seq protocol depends on the sample size, the number of cells to be sequenced and whether transcript counting or full-length mRNA sequencing is desired (Table 1). When large cell numbers must be sequenced, high-throughput, semi-automated droplet-based approaches (for example, inDrop[22] and Dropseq[23]) are optimal. However, these approaches typically achieve lower transcriptome coverage and detect fewer lowly expressed genes, and there is very limited transcript sequence coverage due to 3′ end counting, precluding single-nucleotide variant (SNV) and splicing analyses[17] (Table 1). Droplet-based protocols are also less amenable to studying small cell numbers, for instance, circulating tumour cells (CTCs), disseminated tumour cells or micrometastases. For such samples, cell isolation by flow cytometry or micromanipulation followed by manual library preparation in microwell plates is more tractable. These protocols typically amplify full-length mRNA by switching mechanism at 5′ end of RNA template (SMART) or alternative chemistries, which enable full-length mRNA sequencing for SNV, splicing and deeper transcriptome analyses[24–27]. For tissues that are preserved or cannot be readily dissociated, single-nucleus RNA-seq approaches, such as DroNc-seq[28] or microwell-based single-nucleus RNA-seq[29], may be optimal (Table 1).

Although less widely utilized, single-cell genome analysis has been used to track clonal dynamics, infer evolutionary trajectories and compare paired primary and metastatic tumours (as detailed below). A major challenge is that the DNA must be massively amplified with minimal error. Several alternative chemistries have been developed for whole-genome amplification, which are alternatively better suited for SNV or copy number variant (CNV) analyses (Table 1). Another major issue for single-cell genome studies is that, unlike whole-transcriptome approaches in which lower sequencing depth can still provide robust information about cell identity, the genome has a fixed, large size and there is effectively no cheap way to sequence it. Many single-cell studies are beginning to opt for lower breadth and sequence either the whole exome or a targeted panel of genes (reviewed in ref. [15]). Analysis of single-cell genomic data is also more challenging and less standardized owing to technical errors, such as uneven amplification and allelic dropout. Allelic dropout, in which a particular region of one chromosome is not amplified, is the main technical and analytic challenge and requires imputation of the missing data. Imputation of variants based on probability can be incorrect and confound conclusions. Experimental methods that increase chromosomal compliment, such as in vitro clonal amplification or selection of cells undergoing mitosis, may enable more accurate genotyping[30,31]. Comparison of whole-genome amplification approaches and selected analytic methods can be found in refs [15,32–36].

Protocols to measure other regulatory levels, such as the epigenome and proteome, are developing rapidly, as are multiscale approaches to analyse multiple regulatory levels in the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

same cell (that is, multi-omics) and methods for spatial analysis in intact tissues. Selected methods are highlighted in Box 1.

Genetic heterogeneity and subclonal dynamics

Most tumours comprise subpopulations of cells with distinct genotypes called subclones. Next-generation sequencing data have revealed that a tumour possesses on average over 10,000 somatic mutations: ~2–8 in 'driver' genes that confer a selective growth advantage, and ~30–60 protein-coding changes in 'passenger' genes that may alter other cellular functions[37]. Subclonal diversification arises through genomic instability and numerous mechanisms, including homologous recombination deficiency, chromosome instability, chromothripsis, misregulation of APOBEC enzyme activity and drug treatment[6,38–46]. Several models for subclonal diversification have been proposed based on next-generation sequencing data of bulk human tumour samples, the prevailing being linear, branching, neutral and punctuated models (reviewed in ref. [47]). Most next-generation sequencing studies report branching evolution in human cancers, including leukaemia, breast and liver cancers, colorectal cancer, ovarian cancer, prostate cancer, kidney cancer, melanoma and brain cancer (reviewed in ref. [47]).

Although sophisticated statistical and mathematical models have been developed to infer subclonal dynamics and tumour evolution from bulk data, they rely on major assumptions and cannot extrapolate single-cell genotypes because overlapping mutation frequencies cannot be assigned to the same or different cells, technical errors may yield imprecise mutation frequencies and detection limits preclude identification of rare subclones[15] (Fig. 2). Single-cell genome analyses mitigate many of these limitations and, most importantly, can determine whether mutations are in the same or different cells. This can help to address important questions about subclonal dynamics, such as how specific subclones interact (for example, collaborate versus compete), which subclones can invade and metastasize, how subclonal composition affects clinical outcome and how drug resistance evolves (for example, from pre-existing clones versus the acquisition of new mutations).

Single-cell studies of subclonal heterogeneity have produced new details about subclonal frequencies and their evolution during tumour progression. Single-nucleus sequencing of breast tumours showed that copy number evolution occurred in short bursts early in tumour evolution, whereas point mutations evolved gradually over time to produce more extensive clonal diversity[48]. This provides support for a combined punctuated–branched model for tumour evolution in breast cancer. Genetically distinct clones with unique biological and clinical properties have been identified in acute lymphoblastic leukaemia, colon and breast tumours by single-cell genome analysis[49,50]. For example, targeted single-cell SNV analysis of patients with acute lymphoblastic leukaemia revealed codominant clones, and showed that *KRAS* mutations occurred late in disease development but were not sufficient for clonal dominance[49]. In a case study of colon cancer, the dominant clone possessed *APC* and *TP53* mutations typical of colon cancers, but a rare subclone with *CDC27* and *PABPC1* mutations was also identified[50]. These studies suggest that combinatorial therapies that target multiple subclones may produce better results against polyclonal disease.

Author Manuscript

A study investigating clonal evolution during breast cancer invasion showed a direct lineage relationship between non-invasive ductal carcinoma in situ and adjacent invasive ductal carcinoma lesions in individual patients with breast cancer using topographical single-cell sequencing[51]. These results indicate that most mutations and CNVs evolved within the ducts prior to invasion, with multiple clones escaping from the ducts and co-migrating into the adjacent tissues to establish invasive carcinomas. These findings contrast with models for cancer cell invasion proposing that distinct clones give rise to in situ and invasive tumour cells[52–54], and argue against the notion that extrinsic stimuli (that is, ‘field effects’) cause multifocal disease.

Single-cell genetic analysis has also provided new insights into the longstanding debate over whether drug resistance is caused by the selection of rare pre-existing clones or through acquired resistance by induction of new mutations. Single-cell DNA sequencing of longitudinal samples from patients with breast cancer before and after neoadjuvant chemotherapy showed that resistant genotypes are pre-existing and selected by chemotherapy[55]. Interestingly, scRNA-seq of the same samples showed that the transcriptome profiles of cells pre-treatment and post-treatment were entirely distinct. Cells with resistant programmes were undetectable before treatment, although subsets of cells expressing individual genes associated with chemoresistance could be identified. These data indicate that chemoresistance is conferred through both genetic selection and induction of new transcriptome programmes.

Single-cell analyses of CTCs have also revealed genetic mechanisms of drug resistance. Studies of CTCs from patients with breast cancer have shown heterogeneity for genes important for diagnosis and therapy response, such as *ERBB2* and *PIK3CA*[56,57]. Analysis of CTCs from patients with small-cell lung cancer revealed distinct molecular mechanisms for resistance to chemotherapy[58]. CTCs from chemoresistant patients displayed different CNV profiles than CTCs from chemorefractory patients, suggesting a different genetic basis for immediate resistance (chemorefractory) versus delayed resistance (chemoresistant). A classifier was also developed to distinguish between the two types of patients, supporting the application of single-cell genomics for CTC-based diagnostics[58].

There are still several limitations to single-cell genomic analyses that prevent their more general use in experimental and clinical practice. Technical limitations that hinder confident SNV calling due to allelic dropout are a major issue. Understanding of the biological function and clinical importance of specific genetic clones is also lacking. It will be critical to decipher how a single-cell genotype translates into cellular function, how individual clones affect tumour behaviour and how clones interact to promote tumour progression, metastasis and drug resistance. Innovative strategies are needed to address the lack of a universal approach for the isolation of specific genetic clones and experimental interrogation of their function.

Non-genetic heterogeneity and cellular differentiation

ScRNA-seq studies have shown that many normal tissues are maintained by a pool of adult stem cells that differentiate into multiple ‘cell types’ with distinct ‘cell states’ distinguished

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

by more subtle differences in differentiation, activation, metabolic state or stage of the cell cycle[59–65] (Fig. 2c). Similarly, tumours also contain CSCs, which behave like normal stem cells in their capacity to self-renew, differentiate and propagate the tumour upon transplant (reviewed in ref. [66]). However, some tumours do not follow this model[67,68]. One outstanding question is to what degree CSCs and other tumour cell populations maintain the developmental programme of their normal cell counterparts, versus dedifferentiating or assuming an aberrant cell state (Fig. 2d). Single-cell technologies provide the opportunity to measure cell states in individual cells and map tumour cells onto the normal spectrum of allowable cell types and states of their tissue of origin.

One of the first single-cell studies in cancer used a single-cell multiplex qPCR approach to show that human colon cancers contain distinct cell populations that mirror the cellular lineages of the normal colon[19]. This transcriptional heterogeneity was not due to underlying genetic heterogeneity, as injection of single CSCs into immune-deficient mice gave rise to monoclonal tumours as heterogeneous as the parental one. A study of human oligodendroglioma using scRNA-seq[20] showed that tumours were composed of astrocyte-like and oligodendrocyte-like cells and a rare subpopulation of undifferentiated cells resembling neural stem cells. Similar to the colon cancer study, CNV and SNV analysis showed that each subclonal lineage displayed similar hierarchies. Both studies found differences in the tumour cells relative to their normal counterparts. In the oligodendroglioma study, the two glial lineages seemed to originate from the stem cell-like population, without the discrete differentiation intermediates observed in the normal tissue. Rather, a continuum of differentiation profiles was observed along each lineage, indicating that tumour cells may exist in more dynamic and less discrete differentiation states than normal cells[20] (Fig. 2d). Another study of oligodendroglioma showed that progression to more advanced stages was associated with expansion of primitive cells, with tumours bearing less similarity to the cellular composition of normal tissues[69]. However, the lack of normal cell population profiling in these studies makes it difficult to discern how similar each population was to its healthy counterpart. Likewise, in the colon cancer study[19], several cell subpopulations typical of the normal colon were absent in the tumours, suggesting that there was a skewing or block in normal differentiation[19] (Fig. 2d).

Single-cell studies of acute myelogenous leukaemia (AML) have shown similar alterations in normal differentiation. A mouse model of AML was shown to contain two cell types: one resembling granulocyte/monocyte progenitors and the other macrophage and dendritic cells[70]. Normal and leukaemic counterparts showed significantly different gene network modules, suggesting an aberrant programme in leukaemic cells. Altered myeloid differentiation has also been observed in human paediatric AML by single-cell mass cytometry. In one report, two alternative approaches, PhenoGraph and statistical analysis of perturbation response (SARA), were used to compare signalling programmes in normal haematopoietic and leukaemic cells[71]. Although both primitive and mature monocytelike cells were identified in all patients, they displayed different degrees of aberrant myeloid differentiation, which did not correlate with the cell-surface marker phenotype. In AML, enumeration of primitive $CD34^{+}/CD45^{low}$ blasts is often used for clinical diagnosis and classification of leukaemias[72]. This study showed that intrinsic signalling programmes were often decoupled from the cell-surface phenotype, challenging the standard diagnostic

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

approach. Many patients possessed $CD34^{-}$ leukaemic cells displaying primitive signalling programmes, and the percentage of primitive cells rather than cell-surface phenotype was more predictive of overall survival[71].

These studies suggest that cancer cells often resemble normal cell types and states, but acquire aberrant programmes and display skewed differentiation towards more restricted lineages than their tissue of origin. Defining the aberrant programmes in single cells with relevance to stability and plasticity may reveal new mechanistic insight and therapeutic targeting strategies.

Heterogeneity in diagnostics and therapy response

Mechanisms of drug response have also been studied at the single-cell level. CTCs have been a major focus as they offer a non-invasive window into tumour response. One study reveals that the majority of patients with prostate cancer with varying stages of disease and resistance to androgen deprivation therapies harboured CTCs with at least one type of androgen receptor alteration[73]. Furthermore, patients on the second-line androgen receptor inhibitor enzalutamide (administered after the development of resistance to androgen deprivation therapy) often had CTCs with activated non-canonical Wnt signalling, showcasing how different mechanisms for generating non-genetic heterogeneity can produce drug resistance, and that tracking this through assaying CTCs is clinically feasible.

Single-cell mass cytometry has also been used to measure differential responses to chemotherapy in patients with AML. The percentage of proliferating stem and progenitor-like leukaemic cells was shown to be significantly correlated with therapy response, in addition to signalling abnormalities in primitive AML cell populations[74]. Patients with clinically favourable subtypes of AML demonstrated a higher fraction of cells in S-phase than other subtypes. This suggests that the quiescent state facilitates cancer cell protection and survival during therapy. In another study, chemotherapy increased phenotypic diversity away from traditional stem cell phenotypes in patients with AML[75]. This finding contrasts with classic resistance models that suggest selection of a more primitive cell population during treatment results in outgrowth of a more uniform population.

ScRNA-seq has also shown how standard diagnostic bulk tumour profiling can mask clinically relevant heterogeneity. Glioblastomas can be stratified into four distinct subtypes based on their mRNA expression signature: proneural, neural, classical and mesenchymal[76]. Interestingly, scRNA-seq analysis showed that glioblastomas contain cells that resemble each of the four subtypes, even though this is not observed in bulk analyses. Increased intratumour heterogeneity was also associated with decreased survival, perhaps because of improper patient stratification and treatment assignment, or enhanced genomic instability, tumour evolution or resistance[77].

Single-cell analyses may also eventually be valuable for predictive diagnostics. Single-cell analysis of targeted transcriptome (SCATTome), which predicts drug response for individual cells based on transcriptome signatures using an assortment of machine learning methods,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

was able to predict response to proteasome inhibitor treatment in patients with multiple myeloma[78].

Despite these new discoveries, several limitations of single-cell analyses are worth further consideration. As in single-cell genome studies, assigning function to cell populations is the biggest challenge for most non-genetic single-cell studies. Gene ontology and gene set enrichment analyses can provide clues for signalling pathways that are overrepresented in a cell population, but most of these tools are developed for bulk transcriptome analysis and handle missing data poorly (reviewed in refs [79,80]). As most scRNA-seq protocols only detect 10–40% of the transcriptome, they are not optimal for studying lowly expressed genes, such as those encoding transcription factors, or for the investigation of specific signalling pathways. Dropout of signalling intermediates makes determining whether a pathway is activated challenging. It is also often difficult to determine what constitutes a true cell population, as slight changes in clustering parameters can produce different results. How to decide what represents a cell type versus a cell state continues to be a major debate in the field, particularly for rare cell populations, which may represent cells transiently switching from one state to another or true rare cell types. The only way to assign function to a cell population is through conventional experimentation, which most single-cell reports lack. It is also important to consider that tumour behaviour is probably more than just the sum of its individual cell parts, with different cell populations synergizing or collaborating to take on new macroscale functions and behaviours that are not observable using reductionist approaches. If so, new integrated informatic, experimental and modelling approaches for analysing single cells will be necessary to truly understand tumour behaviour.

Heterogeneity and metastasis

Metastasis remains the cause of most patient mortality and continues to be challenging to treat clinically and investigate experimentally. The general dogma is that metastasis is carried out by rare cells with unique cellular and molecular properties[81]. Single-cell investigations now enable the identification and characterization of such cells, including their localisation in primary tumours, and the effect of genetic versus non-genetic and intrinsic versus extrinsic factors on metastasis (Fig. 3).

Next-generation sequencing studies of bulk tumour samples indicate that metastasis is initiated by a subclone of the primary tumour. In one study, whole-genome sequencing identified numerous point mutations and small indels represented in higher frequencies in the brain metastatic tumour than in the paired primary breast tumour[82]. Similar disparities in mutation frequencies have been reported in paired tumours and metastases from patients with pancreatic and renal cancers, in which the metastasis founder subclones localized to a specific region in the primary tumours, supporting a subclonal model for metastasis initiation[83,84]. One confounding issue in these studies is that the metastases often possessed unique mutations that were not found in paired primary tumours[83]. This makes it challenging to infer the genotype of the original metastasis founder cells, because it is not clear whether the metastasis-exclusive mutations were present in the primary tumour below the limit of detection or whether they arose after metastatic seeding via parallel evolution (Fig. 2b).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Single-cell genome studies of metastasis are beginning to resolve these challenges. Early single-cell studies of metastasis focused on disseminated tumour cells in the bone marrow of patients with breast cancer, and found that these cells disseminate early in tumour evolution by karyotype and subchromosomal variant analyses[85,86]. This is corroborated by reports showing that metastatic lesions similarly derive from early dissemination events[87,88]. Other single-cell genetic methods have generated contradictory findings. A single-cell CNV analysis of breast cancer metastasis showed that a single clonal expansion formed the patient primary tumour and seeded its metastasis[89]. Single-cell CNV and SNV analysis in patients with colon cancer showed monoclonal metastatic seeding in one patient and polyclonal in another[90]. The authors of that study also concluded that metastasis occurred late in primary tumour evolution in both patients, as the metastases harboured all of the trunk mutations that were present in the primary tumour cells[85–88,91–93].

Single-cell genomic analyses also provide evidence for collective cell migration as a mechanism for cancer cell invasion. Topographical single-cell sequencing analysis showed that multiple clones co-migrated through the basement membrane of breast ducts and into adjacent tissues to establish invasive breast carcinomas[51]. This is consistent with previous work showing collective cell migration at the invasive front of tumours, as well as observations that CTC clusters in the bloodstream are more effective than single CTCs for seeding metastasis[94–96].

Single-cell genome analyses of patient CTCs also provide insight into the genotype of potential metastasis-initiating cells. CTCs display substantial subclonal diversity, suggesting that cells of various genotypes are capable of entering the circulation[97]. This includes very rare subclones, as CTCs from patients with colon cancer were reported to carry SNVs that were only present in the paired primary tumours at very low frequencies[98]. Another study found that CTCs in patients with breast cancer often possess variants that are not found in the primary tumour, indicating that they either represent a rare subclone or occur after dissemination[57]. Importantly, the CTCs were heterogeneous for mutations in two common breast cancer drug targets, *ERBB2* and *PIK3CA*, so identifying which CTCs subsequently produce metastases has direct clinical relevance.

As it is clear that metastatic propensity is not encoded exclusively at the genetic level (reviewed in ref. [99]), it will be necessary to investigate other programmes (for example, transcriptomic and epigenetic) driving metastatic progression at single-cell resolution. Single-cell multiplex qPCR technology has shown that metastasis is initiated by cells with stem cell and epithelial–mesenchymal transition-like characteristics in patient-derived xenograft models of breast cancer[100]. This is consistent with results from the MMTV-PyMT breast cancer mouse model showing CSCs as the origin of metastasis[101], as well as reports implicating stem cell and epithelial–mesenchymal transition programmes in other breast cancer models[93,102]. A recent scRNA-seq study of human head-and-neck cancers further implicated epithelial–mesenchymal transition in metastasis[103]. Future studies should utilize single-cell technologies as a new opportunity to investigate other major outstanding questions about metastatic progression, such as what drives metastatic latency and reawakening, how metastatic cells interact with their microenvironment and how they develop resistance to anti-tumour immunity.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Heterogeneity in the microenvironment

The cellular composition of the tumour microenvironment is critical for disease progression and patient prognosis[104] and changes dramatically during tumorigenesis due to increased angiogenesis, tumour-associated inflammation and fibrosis[8]. Emerging single-cell tools bear great potential to profile the individual cell types comprising the tumour ecosystem, to determine how they differ from normal homeostasis and to reconstruct feedback mechanisms mediating cell-to-cell communications between tumour and stromal cells (Fig. 1).

A fundamental caveat in bulk cancer genomic analyses (for example, The Cancer Genome Atlas) is that the differences in transcriptomic signatures between tumours may arise due to differential infiltration of immune and other stromal cell types rather than differences in the tumour cells. This can be addressed with analyses at single-cell resolution. Indeed, scRNA-seq analysis has revealed that the bulk transcriptome differences between oligodendroglioma and astrocytoma driven by mutant IDH are primarily due to alterations within stromal macrophage and microglial populations[69]. Another study matched the transcriptomes of non-malignant cells to cell-type-specific expression profiles and inferred the signalling dialogue between cancer cells and their microenvironment[21]. This revealed distinct patterns of exhaustion or activation of tumour-infiltrating T cells in some tumours, suggesting they may exhibit differential responses to immunotherapies[21].

Breakthroughs in cancer immunotherapy have sparked intensive research into tumour–immune cell interactions using genomic tools, and single-cell analysis pipelines will be instrumental in these approaches[105]. ScRNA-seq analysis of human metastatic melanoma revealed specific homeostatic modules in monocytes and dendritic cells within the tumour microenvironment[106]. A recent systems-level approach using single-cell mass cytometry demonstrated that engagement of systemic and peripheral immunity is critical for tumour rejection after immunotherapy[107]. Future approaches would need to dissect the specificity of anti-tumour immunity before and after checkpoint inhibition, for example, using single-cell V(D)J sequencing to identify the T cell and B cell clones associated with anti-tumour immunity[108]. Direct comparison of single-cell expression signatures from stromal cells in the tumour microenvironment to those in other physiologically similar conditions (for example, inflammation or wound healing) will also determine whether the tumour context induces changes within the normal spectrum of states that a cell may adopt under different conditions or whether cells assume an aberrant tumour microenvironment-specific programme that is not found in normal physiological contexts (Fig. 3c,d).

Future directions

Given its effect on tumour behaviour and clinical outcome, measurements of intratumour heterogeneity should be increasingly incorporated into standard clinical practice. Single-cell targeted DNA sequencing has been commercialized for cancer diagnostics in AML[109]. In addition, commercial scRNA-seq platforms are simple, automated and fast enough to be immediately amenable for diagnostic applications. The continuous release of higher-throughput sequencers will also continue to drive cost down[110]. Thus, the main hurdles to

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

clinical implementation will probably be upstream and downstream of the single-cell technology itself. First, analysis of single-cell sequencing data requires computational skill, so pipelines for demultiplexing, alignment and cell population analysis should be packaged into software and integrated into instrumentation in a single workflow. Furthermore, we lack automated instrumentation or consensus protocols to generate single-cell suspensions from an excised tumour without human intervention. Typical laboratory workflows are manual, low-throughput and not sufficiently reproducible for clinical application. Third, sampling bias is a notable hindrance and it may be necessary to catalogue whole tumours for accurate assessment of heterogeneity. In many cases, only small biopsies are available or patients are pre-treated with neoadjuvant therapies before sample procurement, which may alter the cellular composition of the tumour. However, the main hurdle to clinical implementation will be data interpretation. To harness the power of single-cell assessment of intratumour heterogeneity beyond descriptive cataloguing, we need to first elucidate the biological and clinical functions of different cell populations and develop new ways to specifically target them.

Acknowledgements

We thank those whose work informed the writing of this manuscript and apologize to those authors whose elegant studies we were unable to acknowledge in this Review. We thank K. Blake and J. Wu for thoughtful discussion and suggestions regarding the content of this Review. This work was supported by NIH grants (U01CA199315 to Z.W., K22 CA190511 to D.A.L. and R00 CA181490 to K.K.) and the Chan/Zuckerberg Initiative (HCA-A-1704-01668 to K.K. and D.A.L.). N.P was supported by the National Institute of Biomedical Imaging and Bioengineering, National Research Service Award T32 EB009418 from the University of California, Irvine, Center for Complex Biological Systems. R.T.D. was supported by the NIH, NCI Award T32CA009054 through matched funds.

References

1. Heppner GH Tumor heterogeneity. Cancer Res. 44, 2259–2265 (1984). [PubMed: 6372991]
2. Welch DR Tumor heterogeneity+a ‘contemporary concept’ founded on historical insights and predictions. Cancer Res. 76, 4–6 (2016). [PubMed: 26729788]
3. Mroz EA et al. High intratumor genetic heterogeneity is related to worse outcome in patients with head and neck squamous cell carcinoma. Cancer 119, 3034–3042 (2013). [PubMed: 23696076]
4. Marusyk A, Almendro V & Polyak K Intra-tumour heterogeneity: a looking glass for cancer? Nat. Rev. Cancer 12, 323–334 (2012). [PubMed: 22513401]
5. McGranahan N et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science 351, 1463–1469 (2016). [PubMed: 26940869]
6. McGranahan N & Swanton C Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. Cancer Cell 27, 15–26 (2015). [PubMed: 25584892]
7. Kreso A & Dick JE Evolution of the cancer stem cell model. Cell Stem Cell 14, 275–291 (2014). [PubMed: 24607403]
8. Quail DF & Joyce JA Microenvironmental regulation of tumor progression and metastasis. Nat. Med 19, 1423–1437 (2013). [PubMed: 24202395]
9. Pistollato F et al. Intratumoral hypoxic gradient drives stem cells distribution and MGMT expression in glioblastoma. Stem Cells 28, 851–862 (2010). [PubMed: 20309962]
10. Widmer DS et al. Hypoxia contributes to melanoma heterogeneity by triggering HIF1α-dependent phenotype switching. J. Invest. Dermatol 133, 2436–2443 (2013). [PubMed: 23474946]
11. Black JC et al. Hypoxia drives transient site-specific copy gain and drug-resistant gene expression. Genes Dev. 29, 1018–1031 (2015). [PubMed: 25995187]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

12. Lih CJ et al. Analytical validation of the next-generation sequencing assay for a nationwide signal-finding clinical trial: Molecular Analysis for Therapy Choice Clinical Trial. J. Mol. Diagn 19, 313–327 (2017). [PubMed: 28188106]
13. Nguyen QH, Pervolarakis N, Nee K & Kessenbrock K Experimental considerations for single-cell RNA sequencing approaches. Front. Cell Dev. Biol 6, 108 (2018). [PubMed: 30234113]
14. Kalisky T et al. A brief review of single-cell transcriptomic technologies. Brief Funct. Genomics 17, 64–76 (2018). [PubMed: 28968725]
15. Gawad C, Koh W & Quake SR Single-cell genome sequencing: current state of the science. Nat. Rev. Genet 17, 175–188 (2016). [PubMed: 26806412]
16. Van Loo P & Voet T Single cell analysis of cancer genomes. Curr. Opin. Genet. Dev 24, 82–91 (2014). [PubMed: 24531336]
17. Svensson V et al. Power analysis of single-cell RNA-sequencing experiments. Nat. Methods 14, 381–387 (2017). [PubMed: 28263961]
18. van den Brink SC et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. Nat Methods 14, 935–936 (2017). [PubMed: 28960196]
19. Dalerba P et al. Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat. Biotechnol 29, 1120–1127 (2011). [PubMed: 22081019]
20. Tirosh I et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature 539, 309–313 (2016). [PubMed: 27806376]
21. Tirosh I et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352, 189–196 (2016). [PubMed: 27124452]
22. Klein AM et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161, 1187–1201 (2015). [PubMed: 26000487]
23. Macosko EZ et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. Cell 161, 1202–1214 (2015). [PubMed: 26000488]
24. Picelli S et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat. Methods 10, 1096–1098 (2013) [PubMed: 24056875]
25. Picelli S et al. Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc 9, 171–181 (2014). [PubMed: 24385147]
26. Bhargava V, Head SR, Ordoukhanian P, Mercola M & Subramaniam S Technical variations in low-input RNA-seq methodologies. Sci. Rep 4, 3678 (2014). [PubMed: 24419370]
27. Bhargava V, Ko P, Willems E, Mercola M & Subramaniam S Quantitative transcriptomics using designed primer-based amplification. Sci. Rep 3, 1740 (2013). [PubMed: 23624976]
28. Habib N et al. Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat Methods 14, 955–958 (2017). [PubMed: 28846088]
29. Gao R et al. Nanogrid single-nucleus RNA sequencing reveals phenotypic diversity in breast cancer. Nat. Commun 8, 228 (2017). [PubMed: 28794488]
30. Leung ML et al. Highly multiplexed targeted DNA sequencing from single nuclei. Nat. Protoc 11, 214–235 (2016). [PubMed: 26741407]
31. Leung ML, Wang Y, Waters J & Navin NE SNES: single nucleus exome sequencing. Genome Biol. 16, 55 (2015). [PubMed: 25853327]
32. Huang L, Ma F, Chapman A, Lu S & Xie XS Single-cell whole-genome amplification and sequencing: methodology and applications. Annu. Rev. Genomics Hum. Genet 16, 79–102 (2015). [PubMed: 26077818]
33. Borgstrom E, Paterlini M, Mold JE, Frisen J & Lundeberg J Comparison of whole genome amplification techniques for human single cell exome sequencing. PLoS ONE 12, e0171566 (2017). [PubMed: 28207771]
34. Navin NE & Chen K Genotyping tumor clones from single-cell data. Nat. Methods 13, 555–556 (2016). [PubMed: 27355792]
35. Schwartz R & Schaffer AA The evolution of tumour phylogenetics: principles and practice. Nat. Rev. Genet 18, 213–229 (2017). [PubMed: 28190876]
36. Salehi S et al. ddClone: joint statistical inference of clonal populations from single cell and bulk tumour sequencing data. Genome Biol. 18, 44 (2017). [PubMed: 28249593]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

37. Vogelstein B et al. Cancer genome landscapes. Science 339, 1546–1558 (2013). [PubMed: 23539594]
38. Alexandrov LB et al. Signatures of mutational processes in human cancer. Nature 500, 415–421 (2013). [PubMed: 23945592]
39. Burns MB, Temiz NA & Harris RS Evidence for *APOBEC3B* mutagenesis in multiple human cancers. Nat. Genet. 45, 977–983 (2013). [PubMed: 23852168]
40. Roberts SA et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat. Genet 45, 970–976 (2013). [PubMed: 23852170]
41. Abkevich V et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. Br. J. Cancer 107, 1776–1782 (2012). [PubMed: 23047548]
42. Popova T et al. Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. Cancer Res. 72, 5454–5462 (2012). [PubMed: 22933060]
43. Stephens PJ et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144, 27–40 (2011). [PubMed: 21215367]
44. Zack TI et al. Pan-cancer patterns of somatic copy number alteration. Nat. Genet 45, 1134–1140 (2013). [PubMed: 24071852]
45. Meier B et al. *C. elegans* whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. Genome Res. 24, 1624–1636 (2014). [PubMed: 25030888]
46. Johnson BE et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science 343, 189–193 (2014). [PubMed: 24336570]
47. Davis A, Gao R & Navin N Tumor evolution: Linear, branching, neutral or punctuated? Biochim. Biophys. Acta 1867, 151–161 (2017).
48. Wang Y et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 512, 155–160 (2014). [PubMed: 25079324]
49. Gawad C, Koh W & Quake SR Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. Proc. Natl Acad. Sci. USA 111, 17947–17952 (2014). [PubMed: 25425670]
50. Yu C et al. Discovery of biclonal origin and a novel oncogene *SLC12A5* in colon cancer by single-cell sequencing. Cell Res. 24, 701–712 (2014). [PubMed: 24699064]
51. Casasent AK et al. Multiclonal invasion in breast tumors identified by topographic single cell sequencing. Cell 172, 205–217.e12 (2018) [PubMed: 29307488]
52. Gerlinger M et al. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. Nat. Genet 46,225–233 (2014). [PubMed: 24487277]
53. Miron A et al. *PIK3CA* mutations in in situ and invasive breastcarcinomas. Cancer Res. 70, 5674–5678 (2010). [PubMed: 20551053]
54. Yates LR et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat. Med 21, 751–759 (2015). [PubMed: 26099045]
55. Kim C et al. Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. Cell 173, 879–893.e13 (2018). [PubMed: 29681456]
56. Pestrin M et al. Heterogeneity of *PIK3CA* mutational status at the single cell level in circulating tumor cells from metastatic breast cancer patients. Mol. Oncol 9, 749–757 (2015). [PubMed: 25539732]
57. Polzer B et al. Molecular profiling of single circulating tumor cells with diagnostic intention. EMBO Mol. Med 6, 1371–1386 (2014). [PubMed: 25358515]
58. Carter L et al. Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer. Nat. Med 23, 114–119 (2017). [PubMed: 27869802]
59. Trapnell C Defining cell types and states with single-cell genomics. Genome Res. 25, 1491–1498 (2015). [PubMed: 26430159]
60. Treutlein B et al. Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509, 371–375 (2014). [PubMed: 24739965]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

61. Pollen AA et al. Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat. Biotechnol 32, 1053–1058 (2014). [PubMed: 25086649]

62. Ahn RS et al. Transcriptional landscape of epithelial and immune cell populations revealed through FACS-seq of healthy human skin. Sci. Rep 7, 1343 (2017). [PubMed: 28465541]

63. Jaitin DA et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 343, 776–779 (2014). [PubMed: 24531970]

64. Bjorklund AK et al. The heterogeneity of human $CD127^{+}$ innate lymphoid cells revealed by single-cell RNA sequencing. Nat. Immunol 17, 451–460 (2016). [PubMed: 26878113]

65. Chen L et al. Transcriptomes of major renal collecting duct cell types in mouse identified by single-cell RNA-seq. Proc. Natl Acad. Sci. USA 114, E9989–E9998 (2017). [PubMed: 29089413]

66. Batlle E & Clevers H Cancer stem cells revisited. Nat. Med 23, 1124–1134 (2017). [PubMed: 28985214]

67. Quintana E et al. Efficient tumour formation by single human melanoma cells. Nature 456, 593–598 (2008). [PubMed: 19052619]

68. Ball CR et al. Succession of transiently active tumor-initiating cell clones in human pancreatic cancer xenografts. EMBO Mol. Med 9, 918–932 (2017). [PubMed: 28526679]

69. Venteicher AS et al. Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. Science 355, eaai8478 (2017). [PubMed: 28360267]

70. Saadatpour A, Guo G, Orkin SH & Yuan GC Characterizing heterogeneity in leukemic cells using single-cell gene expression analysis. Genome Biol. 15, 525 (2014). [PubMed: 25517911]

71. Levine JH et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell 162, 184–197 (2015). [PubMed: 26095251]

72. Craig FE & Foon KA Flow cytometric immunophenotyping for hematologic neoplasms. Blood 111, 3941–3967 (2008). [PubMed: 18198345]

73. Miyamoto DT et al. RNA-seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 349, 1351–1356 (2015). [PubMed: 26383955]

74. Behbehani GK et al. Mass cytometric functional profiling of acute myeloid leukemia defines cell-cycle and immunophenotypic properties that correlate with known responses to therapy. Cancer Discov. 5, 988–1003 (2015). [PubMed: 26091827]

75. Ferrell PB Jr et al. High-dimensional analysis of acute myeloid leukemia reveals phenotypic changes in persistent cells during induction therapy. PLoS ONE 11, e0153207 (2016). [PubMed: 27074138]

76. Verhaak RG et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in *PDGFRA, IDH1, EGFR*, and *NF1*. Cancer Cell 17, 98–110 (2010). [PubMed: 20129251]

77. Patel AP et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 344, 1396–1401 (2014). [PubMed: 24925914]

78. Mitra AK et al. Single-cell analysis of targeted transcriptome predicts drug sensitivity of single cells within human myeloma tumors. Leukemia 30, 1094–1102 (2016). [PubMed: 26710886]

79. Poirion OB, Zhu X, Ching T & Garmire L Single-cell transcriptomics bioinformatics and computational challenges. Front. Genet 7, 163 (2016). [PubMed: 27708664]

80. Wagner A, Regev A & Yosef N Revealing the vectors of cellular identity with single-cell genomics. Nat. Biotechnol 34, 1145–1160 (2016). [PubMed: 27824854]

81. Massague J & Obenauf AC Metastatic colonization by circulating tumour cells. Nature 529, 298–306 (2016). [PubMed: 26791720]

82. Ding L et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature 464, 999–1005 (2010). [PubMed: 20393555]

83. Gerlinger M et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N. Engl. J. Med 366, 883–892 (2012). [PubMed: 22397650]

84. Yachida S et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 467, 1114–1117 (2010). [PubMed: 20981102]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

85. Schardt JA et al. Genomic analysis of single cytokeratin-positive cells from bone marrow reveals early mutational events in breast cancer. Cancer Cell 8, 227–239 (2005). [PubMed: 16169467]

86. Schmidt-Kittler O et al. From latent disseminated cells to overt metastasis: genetic analysis of systemic breast cancer progression. Proc. Natl Acad. Sci. USA 100, 7737–7742 (2003). [PubMed: 12808139]

87. Husemann Y et al. Systemic spread is an early step in breast cancer. Cancer Cell 13, 58–68 (2008). [PubMed: 18167340]

88. Hosseini H et al. Early dissemination seeds metastasis in breast cancer. Nature 540, 552–558 (2016). [PubMed: 27974799]

89. Navin N et al. Tumour evolution inferred by single-cell sequencing. Nature 472, 90–94 (2011). [PubMed: 21399628]

90. Leung ML et al. Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. Genome Res. 27, 1287–1299 (2017). [PubMed: 28546418]

91. Riethmuller G & Klein CA Early cancer cell dissemination and late metastatic relapse: clinical reflections and biological approaches to the dormancy problem in patients. Semin. Cancer Biol 11, 307–311 (2001). [PubMed: 11513566]

92. Linde N et al. Macrophages orchestrate breast cancer early dissemination and metastasis. Nat. Commun 9, 21 (2018). [PubMed: 29295986]

93. Harper KL et al. Mechanism of early dissemination and metastasis in $Her2^+$ mammary cancer. Nature 540, 588–592 (2016). [PubMed: 27974798]

94. Cheung KJ, Gabrielson E, Werb Z & Ewald AJ Collective invasion in breast cancer requires a conserved basal epithelial program. Cell 155, 1639–1651 (2013). [PubMed: 24332913]

95. Liotta LA, Saidel MG & Kleinerman J The significance of hematogenous tumor cell clumps in the metastatic process. Cancer Res. 36, 889–894 (1976). [PubMed: 1253177]

96. Aceto N et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. Cell 158, 1110–1122 (2014). [PubMed: 25171411]

97. Ni X et al. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc. Natl Acad. Sci. USA 110, 21083–21088 (2013). [PubMed: 24324171]

98. Heitzer E et al. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. Cancer Res. 73, 2965–2975 (2013). [PubMed: 23471846]

99. Nguyen DX, Bos PD & Massague J Metastasis: from dissemination to organ-specific colonization. Nat. Rev. Cancer 9, 274–284 (2009). [PubMed: 19308067]

100. Lawson DA et al. Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature 526, 131–135 (2015). [PubMed: 26416748]

101. Malanchi I et al. Interactions between cancer stem cells and their niche govern metastatic colonization. Nature 481, 85–89 (2012).

102. Malladi S et al. Metastatic latency and immune evasion through autocrine inhibition of WNT. Cell 165, 45–60 (2016). [PubMed: 27015306]

103. Puram SV et al. Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. Cell 171, 1611–1624.e24 (2017). [PubMed: 29198524]

104. Fridman WH, Zitvogel L, Sautes-Fridman C & Kroemer G The immune contexture in cancer prognosis and treatment. Nat. Rev. Clin. Oncol 14, 717–734 (2017). [PubMed: 28741618]

105. Hackl H, Charoentong P, Finotello F & Trajanoski Z Computational genomics tools for dissecting tumour–immune cell interactions. Nat. Rev. Genet 17, 441–458 (2016). [PubMed: 27376489]

106. Nirschl CJ et al. IFNγ-dependent tissue-immune homeostasis is co-opted in the tumor microenvironment. Cell 170, 127–141.e15 (2017). [PubMed: 28666115]

107. Spitzer MH et al. Systemic immunity is required for effective cancer immunotherapy. Cell 168, 487–502.e15 (2017). [PubMed: 28111070]

108. Redmond D, Poran A & Elemento O Single-cell TCRseq: paired recovery of entire T-cell α and β chain transcripts in T-cell receptors from single-cell RNAseq. Genome Med. 8, 80 (2016). [PubMed: 27460926]

109. Eastburn DJ et al. High-throughput single-cell DNA sequencing of AML tumors with droplet microfluidics. Blood 130, 3965 (2017).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

110. Zheng GX et al. Massively parallel digital transcriptional profiling of single cells. Nat. Commun 8, 14049 (2017). [PubMed: 28091601]
111. Tirosh I et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352, 189–196 (2016). [PubMed: 27124452]
112. Chung W et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat. Commun. 8, 15081 (2017). [PubMed: 28474673]
113. Darmanis S et al. Single-cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. Cell Rep. 21, 1399–1410 (2017). [PubMed: 29091775]
114. Hayashi T et al. Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. Nat. Commun 9, 619 (2018). [PubMed: 29434199]
115. Klein AM et al. Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161, 1187–1201 (2015). [PubMed: 26000487]
116. Muller S et al. Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. Genome Biol. 18, 234 (2017). [PubMed: 29262845]
117. Bell JM et al. Chromosome-scale mega-haplotypes enable digital karyotyping of cancer aneuploidy. Nucleic Acids Res. 45, e162 (2017). [PubMed: 28977555]
118. Savage P et al. A targetable EGFR-dependent tumor-initiating program in breast cancer. Cell Rep. 21, 1140–1149 (2017). [PubMed: 29091754]
119. Yuan J & Sims PA An automated microwell platform for large-scale single cell RNA-seq. Sci. Rep. 6, 33883 (2016). [PubMed: 27670648]
120. Gierahn TM et al. Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nat. Methods 14, 395–398 (2017). [PubMed: 28192419]
121. Rosenberg AB et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. Science 360, 176–182 (2018). [PubMed: 29545511]
122. Cao J et al. Comprehensive single-cell transcriptional profiling of a multicellular organism. Science 357, 661–667 (2017). [PubMed: 28818938]
123. Taniguchi K, Kajiyama T & Kambara H Quantitative analysis of gene expression in a single cell by qPCR. Nat. Methods 6, 503–506 (2009). [PubMed: 19525960]
124. Lee JH et al. Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. Nat. Protoc 10, 442–458 (2015). [PubMed: 25675209]
125. Shah S, Lubeck E, Zhou W & Cai L seqFISH accurately detects transcripts in single cells and reveals robust spatial organization in the hippocampus. Neuron 94, 752–758.e1 (2017). [PubMed: 28521130]
126. Chen KH, Boettiger AN, Moffitt JR, Wang S & Zhuang X Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348, aaa6090 (2015). [PubMed: 25858977]
127. Habib N et al. Div-Seq: single-nucleus RNA-seq reveals dynamics of rare adult newborn neurons. Science 353, 925–928 (2016). [PubMed: 27471252]
128. Lake BB et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science 352, 1586–1590 (2016). [PubMed: 27339989]
129. Grindberg RV et al. RNA-sequencing from single nuclei. Proc. Natl Acad. Sci. USA 110, 19802–19807 (2013). [PubMed: 24248345]
130. Faridani OR et al. Single-cell sequencing of the small-RNA transcriptome. Nat. Biotechnol 34, 1264–1266 (2016). [PubMed: 27798564]
131. Klein CA et al. Comparative genomic hybridization, loss of heterozygosity, and DNA sequence analysis of single cells. Proc. Natl Acad. Sci. USA 96, 4494–4499 (1999). [PubMed: 10200290]
132. Lohr JG et al. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat. Biotechnol 32, 479–484 (2014). [PubMed: 24752078]
133. Lu S et al. Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. Science 338, 1627–1630 (2012). [PubMed: 23258895]
134. Chen C et al. Single-cell whole-genome analyses by linear amplification via transposon insertion (LIANTI). Science 356, 189–194 (2017). [PubMed: 28408603]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

135. Lodato MA et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. Science 350, 94–98 (2015). [PubMed: 26430121]
136. Jin W et al. Genome-wide detection of DNase I hypersensitive sites in single cells and FFPE tissue samples. Nature 528, 142–146 (2015). [PubMed: 26605532]
137. Buenrostro JD et al. Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523, 486–490 (2015). [PubMed: 26083756]
138. Cusanovich DA et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. Science 348, 910–914 (2015). [PubMed: 25953818]
139. Ramani V et al. Massively multiplex single-cell Hi-C. Nat. Methods 14, 263–266 (2017). [PubMed: 28135255]
140. Rotem A et al. Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. Nat. Biotechnol 33, 1165–1172 (2015). [PubMed: 26458175]
141. Smallwood SA et al. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat. Methods 11, 817–820 (2014). [PubMed: 25042786]
142. Han L et al. Bisulfite-independent analysis of CpG island methylation enables genome-scale stratification of single cells. Nucleic Acids Res. 45, e77 (2017). [PubMed: 28126923]
143. Zhu C et al. Single-cell 5-formylcytosine landscapes of mammalian early embryos and ESCs at single-base resolution. Cell Stem Cell 20, 720–731.e5 (2017). [PubMed: 28343982]
144. Mooijman D, Dey SS, Boisset JC, Crosetto N & van Oudenaarden A Single-cell 5hmC sequencing reveals chromosome-wide cell-to-cell variability and enables lineage reconstruction. Nat. Biotechnol 34, 852–856 (2016). [PubMed: 27347753]
145. Macaulay IC et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. Nat. Methods 12, 519–522 (2015). [PubMed: 25915121]
146. Macaulay IC et al. Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq. Nat. Protoc 11, 2081–2103 (2016). [PubMed: 27685099]
147. Dey SS, Kester L, Spanjaard B, Bienko M & van Oudenaarden A Integrated genome and transcriptome sequencing of the same cell. Nat. Biotechnol 33, 285–289 (2015). [PubMed: 25599178]
148. Angermueller C et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. Nat. Methods 13, 229–232 (2016). [PubMed: 26752769]
149. Hou Y et al. Single-cell triple omics sequencing reveals genetic, epigenetic, and transcriptomic heterogeneity in hepatocellular carcinomas. Cell Res. 26, 304–319 (2016). [PubMed: 26902283]
150. Clark SJ et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. Nat. Commun 9, 781 (2018). [PubMed: 29472610]
151. Stoeckius M et al. Simultaneous epitope and transcriptome measurement in single cells. Nat. Methods 14, 865–868 (2017). [PubMed: 28759029]
152. Peterson VM et al. Multiplexed quantification of proteins and transcripts in single cells. Nat. Biotechnol 35, 936–939 (2017). [PubMed: 28854175]
153. Bandura DR et al. 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 (2009). [PubMed: 19601617]
154. Ornatsky O et al. Highly multiparametric analysis by mass cytometry. J. Immunol. Methods 361, 1–20 (2010). [PubMed: 20655312]
155. Bendall SC et al. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 332, 687–696 (2011). [PubMed: 21551058]
156. Goltsev Y et al. Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. Preprint at 10.1101/203166 (2018).
157. Lee JH et al. Highly multiplexed subcellular RNA sequencing in situ. Science 343, 1360–1363 (2014). [PubMed: 24578530]
158. Coskun AF & Cai L Dense transcript profiling in single cells by image correlation decoding. Nat Methods 13, 657–660 (2016). [PubMed: 27271198]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Box 1 |

Emerging technologies for single-cell analyses in cancer

Single-cell transcriptome and genome studies have emerged as vital tools for investigating mechanisms of cancer, and more technologies are being adapted to operate in single-cell resolution to explore the disease in ways that were previously unavailable. This is exemplified by several common protocols for epigenetic analysis that now can be performed at the single-cell level, including DNase-seq[136] and ATAC-seq[137,138] for defining regions of open chromatin, Hi-C for investigation of chromosomal contacts[139], ChIP–seq for histone position mapping[140], bisulfite sequencing (single-cell bisulfite-seq)[141] or bisulfite-free methods (single-cell CGI-seq)[142] for measuring DNA methylation state, and CLEVER-seq[143] and scAba-seq[144] to measure active DNA demethylation by 5-fluorocytosine sequencing and 5-hydroxymethylcytosine sequencing, respectively. Multiscale analysis of multiple regulatory levels in the same cell is also an area of rapid development and provides the potential to comprehensively understand how and why malignant cells produce a particular phenotype, function or behaviour. Combinatorial methods currently exist for analysis of the genome and transcriptome (G&T-seq and DR-seq)[145–147], epigenome and transcriptome (scMT-seq, scTrio-seq and scNMT-seq)[148–150], and techniques for studying the proteome and transcriptome are in development[151,152]. Single-cell proteomics are another area of paramount interest as protein expression is the ultimate functional output of the cell. Although there are no commercial methods for whole-proteome level analysis, several technologies for high-parameter protein analysis in single cells have been developed. Time of flight mass cytometry (CyTOF), which utilizes heavy metal-conjugated antibodies to quantify protein expression by time-of-flight inductively coupled plasma mass spectrometry, can theoretically multiplex up to 135 parameters in single cells[153–155]. The co-detection by indexing (CODEX) platform also enables high-parameter protein expression analysis using an in situ polymerization-based indexing procedure and fluorescently barcoded antibodies[156]. This provides the added advantage of spatial localization of cell populations within the native tissue context. Spatial genomics technologies for single-cell analysis, such as FISSEQ[157], seqFISH[158] or MERFISH[126], are another area of rapid development, which reveal cellular neighbourhoods that are specific to the tumour microenvironment. Ultimately, spatial approaches will allow investigators to define changes in cell populations that associate with specific histological and pathological tissue phenotypes.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

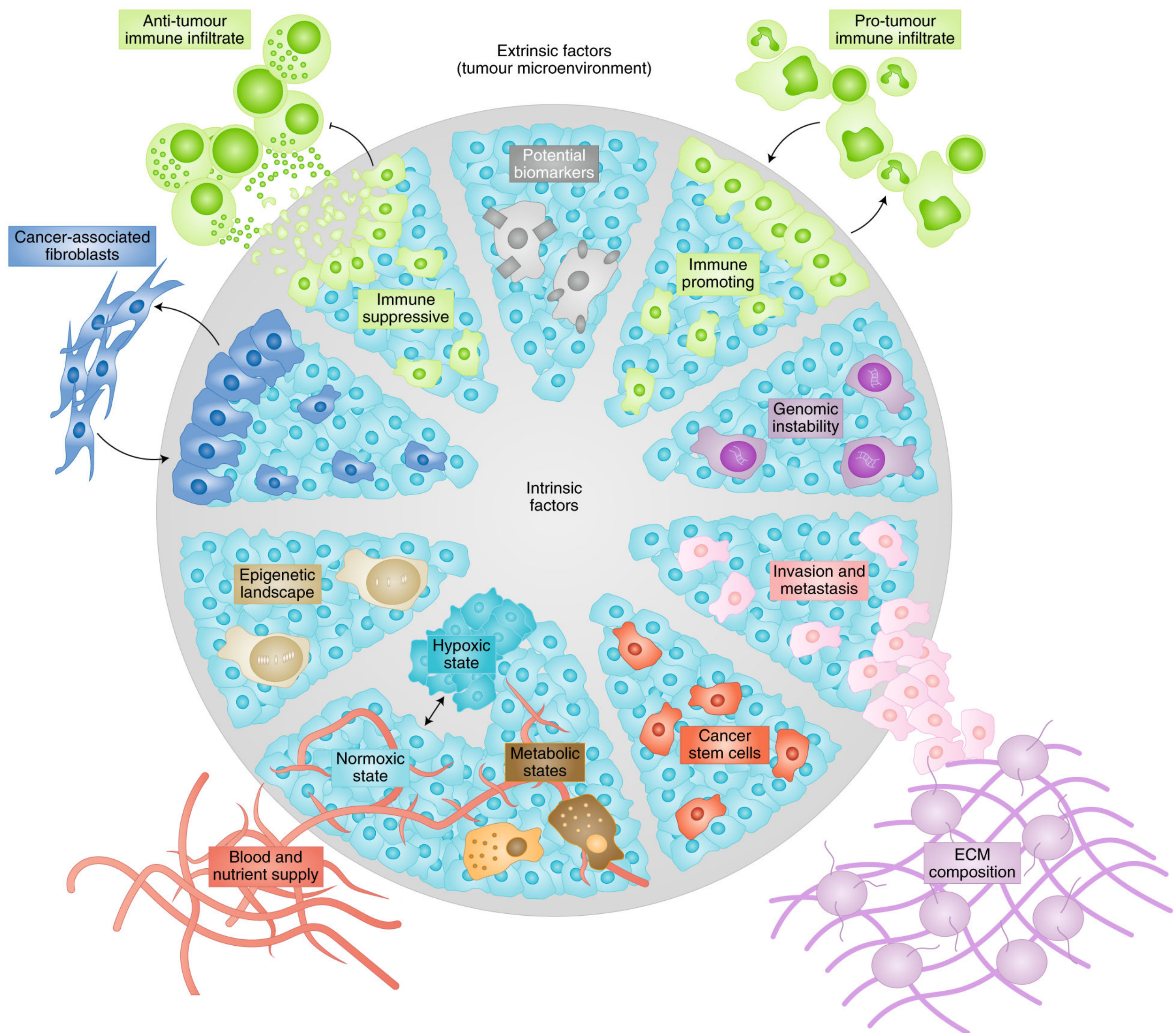

Fig. 1 |. Common types of intratumour heterogeneity and its regulation by intrinsic and extrinsic factors.

Tumours comprise a heterogeneous population of cells, which is regulated by both intrinsic and extrinsic factors. Tumour cells vary in biomarker expression, epigenetic landscape, hypoxic state, metabolic state, stage of differentiation, invasive potential and genotype due to genomic instability. The tumour microenvironment can also be heterogeneous, in which different types of fibroblasts, pro-tumour and anti-tumour immune infiltrate, vascular and lymphatic vessel density and extracellular matrix (ECM) composition affect tumour cell heterogeneity and function.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

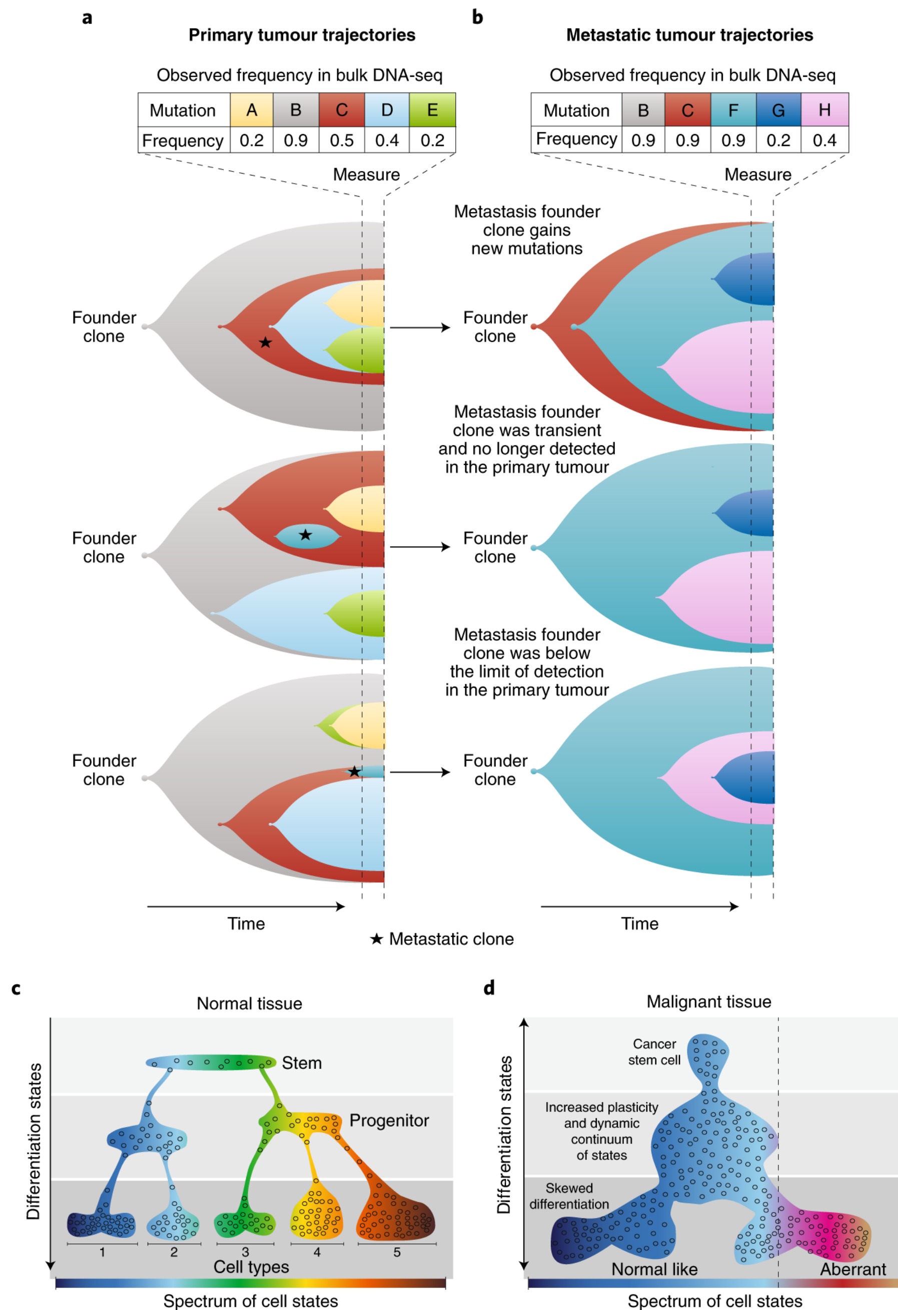

Fig. 2 |. Deciphering subclonal composition and cell types and states in single-cell omics data.
a, Inferring clonal trajectories and subclonal heterogeneity from bulk primary tumour genome sequencing data. In this example experiment, a tumour is sampled at a single time point (dotted lines). The table shows the frequency of each detected mutation. The panels show three (of many) possible clonal trajectories that can be inferred. The nodes represent points at which a mutation occurred, and overlapping coloured regions indicate that each of the mutations is present within any cell that is part of that population. **b**, Challenges associated with deciphering the genotype of a metastatic founder clone and subclonal trajectories from bulk genome sequencing of paired metastatic and primary tumours. The

table shows the observed mutation frequencies in an example experiment in which a metastatic tumour from the individual in **a** was sequenced. The panels show three possible explanations for the observed frequencies. **c**, Cell types and states found in normal tissues. Tissues comprise different mature 'cell types' (labelled 1–5), which carry out specified functions. Cells within a 'type' can exist in a spectrum of allowable 'cell states' depending on the physiological status of the tissue. Mature cell types are derived from stem cells through a series of discrete differentiation intermediates or progenitors. The circles represent single cells, and the colour clouds represent the spectrum of allowable states. The density of circles represents the probability of observing a cell with that phenotype in a scRNA-seq experiment. **d**, Tumour cell types and states differ from normal tissue. Single-cell studies have shown that tumours contain stem-like cells (CSCs) and that differentiation is often noisy, skewed towards specific cell lineages.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

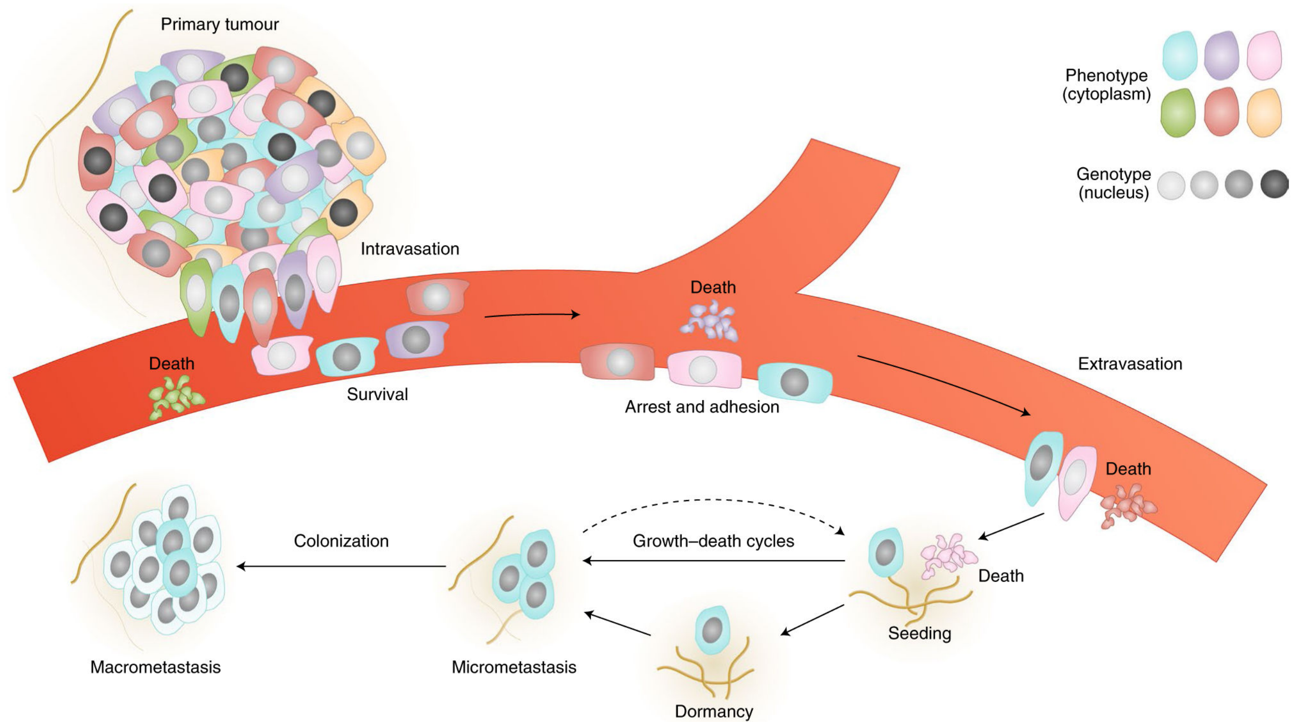

Fig. 3 |. Genetic and phenotypic properties of metastasis-initiating cells at the single-cell level.

Metastasis is a rare event, in which most cancer cells cannot progress through major bottlenecks associated with invasion, intravasation, extravasation, seeding and colonization to produce a malignant macrometastatic tumour. In this model, cancer cells are heterogeneous in genotype (nuclei) and phenotype (cytoplasm), and metastasis-initiating cells possess a distinct combination of both. Dashed arrow indicates that cancer cells within micrometastases can die. Death rates within micrometastases can balance proliferation rates, and thereby prevent progression to macrometastasis by the failure to produce net positive growth.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1 |

Technical characteristics, advantages and limitations of single-cell technologies

Transcriptome					
Protocol	**Library type**	**Throughput**	**Advantages**	**Limitations**	**Ref.**
SMART-seq	Full-length transcriptome	Low	•Ability to profile rare cell populations directly •Cell capture visualization •Splicing variant analysis and SNV profiles possible	•Limited scalability •Higher degree of manual inputs •Limited multiplexing possibilities	25,69,77, 111-113
RamDA-seq	Full-length transcriptome	Low	•Ability to profile rare cell populations directly •Splicing variant analysis and SNV profiles possible •Ability to capture non-poly(A) transcripts	•Limited scalability •Higher degree of manual inputs •Limited multiplexing possibilities	114
Droplet based	3′ Transcriptome	High	•Reduction in PCR amplification bias •High cell yield	•Decreased gene coverage per cell •Sequencing required to estimate cell capture •No splicing or full-transcript SNV information •High cell input volumes are not suitable to rarer cell populations	23,110, 115-118
Microwell based	3′ Transcriptome	Medium	•Cell capture visualization •Reduction in PCR amplification bias •Parallel processing of experimental conditions on the same chip	•No splicing or full-transcript SNV information •High cell input volumes are not suitable to rarer cell populations	119,120
Combinatorial indexing	3′ Transcriptome	High	•High cell yield •Ability to run multiple conditions or cell types while maintaining identity •Reduction in PCR amplification bias	•Decreased gene coverage per cell •No splicing or full-transcript SNV information	121,122
Single-cell qPCR	Targeted transcriptome	Medium	•Analysis of ‘small RNA’ targets (miRNA, snoRNA and piRNA, among others) •High sensitivity for lowly expressed transcripts •Ability to selectively probe for targets of interest	•Limited cell numbers per run •Limited target transcripts per run	100,123
FISSEQ	In situ sequencing	High	•Spatial localization of transcripts in tissue	•High imaging system costs •Extended protocol length (owing to imaging requirements) •No rRNA depletion	124
seqFISH	In situ hybridization	Medium	•Spatial localization of transcripts in tissue •Exact transcript counts per cell	•High imaging system costs •Limited scalability of transcript targets	125
MERFISH	In situ hybridization	Medium	•Spatial localization of transcripts in tissue •Reduced misidentification rate owing to a unique encoding scheme	•High imaging system costs •Potentially high imaging times, resulting in sample degradation	126
Single -nucleus-seq	3′ Nuclear transcriptome	High	•Fragile cell or tissue processing	•Potential bias of nuclear-retained transcripts over exported transcripts	28,127-129
Small RNA-seq	3’ Small RNA Transcriptome	Low	•Analysis of ‘small RNA’ targets (miRNA, snoRNA and piRNA, among others)	•Limited scalability •Potential 3′ end bias	130

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Transcriptome

Protocol	Library type	Throughput	Advantages	Limitations	Ref.
			•Reduction in PCR amplification bias		

Genome

Amplification protocol	Library type	Amplification	Advantages	Limitations	Ref.
Ampli1 WGA	Ligation-mediated PCR following a site-specific DNA digestion	Exponential	•Use of non-random primers results in a more even coverage	•Long and time-consuming protocol	131
DOP-PCR	Primer-based amplification	Exponential	•Accurate detection of CNVs	•The polymerase used has a higher error rate •Difficult to identify SNVs	89
MDA	Phi29 looping amplification	Exponential	•High-fidelity polymerase •Suitable for SNV analysis	•Uneven coverage of the total genome •Difficult to identify CNVs	30,132
MALBAC	Combination of looping and primer based	Quasi-linear	•Accurate enough for SNVs and large CNVs •Most even read distribution	•Lower confidence SNV identification than MDA	133
LIANTI	Linear amplification by transposon insertion	Linear	High genome coverage, reduced amplification bias and errors •Increased accuracy for SNVs and CNVs •Detection of micro CNV at kilobyte resolution	•Despite reduced errors, the false-positive rate (1.7×10^{-6}) still prevents exact detection of SNVs	134

Capture method	Library type	Cost per cell	Advantages	Limitations	Ref.
Whole genome	Coverage of the entire genome	$$$	•Unbiased •Powerful for phylogenetic analysis	•Highly dimensional data •Expensive •Computationally intense	131,135
Whole exome	Coverage of exonic and some regulatory regions	$$	•Emphasis on actionable mutations in protein-coding regions	•Potential dropout of exonic regions that are relevant owing to inefficient capture	132
Targeted	Coverage of specific genomic sites of interest	$	•Focused analysis on highly relevant regions •Cheaper	•Biased and potentially restrictive •Sacrifice of information	30

DOP-PCR, degenerate oligonucleotide-primed PCR; FISSEQ, fluorescent in situ sequencing; MALBAC, multiple annealing and looping-based amplification cycle; MDA, multiple displacement amplification; MERFISH, multiplexed error-robust fluorescence in situ hybridization; miRNA, microRNA; piRNA, Piwi-interacting RNA; RamDA-seq, random displacement amplification sequencing; seqFISH, sequential fluorescence in situ hybridization; snoRNA, small nucleolar RNA; WGA, whole-genome amplification.