CellPress
OPEN ACCESS

Trends in
Cancer

Review

Insights into the metastatic cascade through research autopsies

Sonya Hessey,[1,2,3] Petros Fessas,[2] Simone Zaccaria,[1,3] Mariam Jamal-Hanjani,[1,2,4,6,*] and Charles Swanton[1,4,5,6,*]

Metastasis is a complex process and the leading cause of cancer-related death globally. Recent studies have demonstrated that genomic sequencing data from paired primary and metastatic tumours can be used to trace the evolutionary origins of cells responsible for metastasis. This approach has yielded new insights into the genomic alterations that engender metastatic potential, and the mechanisms by which cancer spreads. Given that the reliability of these approaches is contingent upon how representative the samples are of primary and metastatic tumour heterogeneity, we review insights from studies that have reconstructed the evolution of metastasis within the context of their cohorts and designs. We discuss the role of research autopsies in achieving the comprehensive sampling necessary to advance the current understanding of metastasis.

Highlights

The identification of cells responsible for metastasis and their routes of dissemination is critical for understanding the characteristics that engender metastatic potential.

DNA sequencing of primary tumours and metastases from the same patient enables the reconstruction of the evolutionary steps of metastasis, but the accuracy of this strategy is reliant on comprehensive sampling of both.

Research autopsy studies overcome the limitations of tissue sampling in living patients, thereby providing the detailed view of the metastatic landscape needed to inform an understanding of the biological underpinnings of metastasis and approaches to prevent its onset.

Tracing the steps of metastasis using genomics

Metastasis is a multi-step process wherein primary tumour cells spread to a distant anatomical site [1]. During this process, cells leave the primary tumour by invading surrounding tissues, intravasate into blood and lymphatic systems, where they migrate until they arrest, extravasate into new organ sites, and proliferate into micro- then clinically detectable metastases [2]. Despite the inefficiency of this process, many patients with cancer suffer metastatic disease at diagnosis and, given that metastasis is the leading cause of cancer-related mortality [3], there remains a clinical need to better understand its biological underpinnings.

To successfully metastasise, cancer cells require a set of phenotypic traits that overcome biological barriers unique to each step of the process, subsequently referred to as 'metastatic potential'. These include, but are not limited to, the ability to permeate blood vessels, to withstand detachment from the extracellular matrix, to evade immune elimination, and to adapt to conditions of foreign tissues [4]. Understanding the genetic basis and molecular mechanisms of these phenotypes, as well as how they contribute to cancer dissemination, could inform strategies to prevent and treat this lethal condition [5].

Identifying cancer cells capable of metastasising and their routes of dissemination is complex, since it involves tracking the migration of cells from the primary tumour to metastases, between metastases [6], and even from metastases back to the primary tumour in cases where the primary remains *in situ* [7]. Recent studies demonstrated that DNA-sequencing data from primary tumour and metastasis samples from the same patient can be used to accomplish this task by allowing the evolutionary history of the subpopulations of genetically distinct cancer cells, or 'clones', in each sample to be reconstructed [8–10] (Box 1). However, the accuracy of this approach depends on how representative the samples are of a patient's primary tumour and metastatic burden. Heterogeneity within an individual primary tumour or metastasis (intratumour heterogeneity) implies that a single region sample may not detect all clones present in that tumour. Similarly,

[1]Cancer Research UK Lung Cancer Centre of Excellence, University College London Cancer Institute, London, UK
[2]Cancer Metastasis Laboratory, University College London Cancer Institute, London, UK
[3]Computational Cancer Genomics Research Group, University College London Cancer Institute, London, UK
[4]Department of Oncology, University College London Hospitals, London, UK
[5]Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK
[6]These authors jointly supervised this work.

*Correspondence:
m.jamal-hanjani@ucl.ac.uk
(M. Jamal-Hanjani) and
charles.swanton@crick.ac.uk
(C. Swanton).

heterogeneity between distinct metastases from the same patient (intermetastasis heterogeneity) implies that sampling only a subset of existing metastases may not detect clones unique to the unsampled sites. In most solid cancer types, primary tumours are recognised to have extensive intratumour heterogeneity [11] and, thus, require multiregion sampling to adequately characterise their clonal composition [12–14]. The degree of intratumour heterogeneity in metastases has not been systematically investigated, but studies that include multiple metastases from the same patient have demonstrated significant intermetastasis heterogeneity [6,9,10,15–19]. Since the identification of metastasising cells and approaches to reconstruct their dissemination patterns rely on the accurate identification of clones within the primary tumour and metastases (Box 1), deciphering the steps of metastasis using DNA sequencing necessitates studies that employ comprehensive multiregion and multisite sampling. Although this approach leverages the progressive accumulation of somatic mutations over time to infer the steps involved in the metastatic transition, its resolution is limited to the clones present at the time of sampling. Therefore, the additional collection of samples, such as circulating tumour DNA or circulating tumour cells, which can provide information about clonal dynamics in the interval between tissue sampling [21–23], has potential to further enhance the accuracy with which metastatic progression can be traced.

Here, we review genomic sequencing studies that have aimed to gain insight into the metastatic process through the reconstruction of its evolutionary steps. We first summarise the designs of studies that have enabled the acquisition of primary tumour and metastasis pairs, which are fundamental requirements to this approach. We focus on findings relating to two key aspects of metastasis examined in these studies: the characteristics that endow metastatic potential, and the modes and routes by which cancer disseminates. For each, we consider how tumour heterogeneity combined with constraints on sampling might influence the reliability of conclusions, and how this might be overcome. Finally, we describe our experience within the Posthumous Evaluation of Advanced Cancer Environment (PEACE; NCT03004755) autopsy programme and discuss ways that future studies could be designed to address outstanding questions. Although we focus here on the utility of DNA sequencing in studying metastatic progression, a complete picture of the metastatic cascade will require an understanding of the contribution of nongenetic characteristics of cancer cells [24,25] to metastatic potential, and the role of tumour extrinsic factors, including tumour microenvironment changes [26,27] and metastatic niche priming [28,29], as well as cell–cell interactions [30,31], in mediating dissemination and colonisation.

Research autopsies facilitate the study of primary-metastasis pairs

Establishing cohorts of patients from whom both primary and metastatic tumour tissues are available is challenging: unlike primary tumours, metastases are infrequently resected, and the clinical management of metastatic disease usually necessitates a sample of either a patient's primary tumour or metastasis, but rarely both.

One approach to overcoming this challenge is to identify opportune clinical pathways that enable the collection of paired primary and metastatic tumours, and embed studies with strategic designs within these pathways. In various cancer types, including colorectal [32–34], endometrial [35], melanoma [18], renal [12,36], breast [37], and, recently, lung [38], for which surgery is the mainstay of treatment for patients with localised disease, studies have collected resected primary tumour tissue at the time of surgery, often enabling multiregion sampling, and subsequent biopsy or metastasectomy samples in the event of disease relapse. Such cohorts enable the study of temporal dynamics (from diagnosis to relapse or progression), but are limited in their assessment of spatial variation between distinct metastatic sites, since few metastases are feasibly sampled in each patient.

Research autopsies overcome many of the limitations of collecting tissue for research from living patients. Research autopsy studies conducted on patients with breast [16,17,22,25,27,39–43], prostate [6,44–47], skin [48–50,121], pancreatic [51–55], lung [56–63], ovarian [64], brain [65], gastrointestinal [66–70], renal [36,71], and urothelial [72,73] cancers have demonstrated the feasibility of this approach, with a view to sampling both diseased and normal tissues to support basic or translational research [74]. They are particularly valuable in the metastatic cancer setting because macroscopic pathological examination enables the collection of a set of samples that are representative of a patient's full metastatic burden at the time of death. In patients whose primary tumour is still *in situ* at the time of death, research autopsies enable simultaneous sampling of both primary and metastatic tissue. Furthermore, larger tissue samples can be collected allowing for the integration of multiple parallel analyses, including cell line, organoid and xenograft model generation [75], and studies of the tumour microenvironment [71].

Study design influences cohort characteristics

The clinical setting from which a study acquires its tumour samples (surgical resection, biopsies, or research autopsies) will determine the type, quantity, timepoints, and breadth of tissue available for downstream analyses, and the biological questions that can be addressed. Consequently, studies of primary–metastasis pairs conducted over the past decade have varied substantially in terms of number of analysed samples and patients, as well as the type of sequencing performed, each of which might influence inferences about the genetic determinants of metastatic potential and routes of dissemination.

Primary–metastasis studies that collect samples from surgeries or biopsies have included a greater number of patients (median 12, range 3–170) compared with research autopsy studies (median 5, range 2–63). Meanwhile, research autopsy studies have included a greater number of metastasis samples per patient (median 5, range 1–14) compared with studies of surgery or biopsy samples (median 2, range 1–7), highlighting the trade-off between representativeness of a patient population and representativeness of intermetastatic heterogeneity between these study designs (Figure 1A). Multiregion sequencing was performed on approximately half of primary tumours in the studies surveyed irrespective of the sample acquisition strategy (Figure 1B). By contrast, multiregion sequencing of metastases was rarely performed and notably infrequent in research autopsy studies despite tissue availability (Figure 1B,C). Furthermore, primary tumour tissue was not available for 37% of patients enrolled in research autopsy studies (Figure 1C), precluding reconstruction of primary to metastasis evolution and dissemination patterns in these patients.

It is not known how the number of metastases sequenced in these studies relates to the actual number of metastases and extent of heterogeneity present at the time of sampling. As such, how representative the sequenced samples are of a patient's disease burden, and the degree to which inferences about the metastasising cells and their route of dissemination recapitulate the full complexity of the metastatic process, cannot be determined. Therefore, considering the findings of primary–metastasis studies to date within the context of their cohort characteristics will help identify unmet needs in metastasis research and inform the design of future studies. The following sections summarise insights into the genetic features of metastatic potential and dissemination afforded by these studies.

Genomic characteristics that drive metastatic potential

Genomic alterations can be categorised according to their impact on cancer cell phenotypes: those that confer a fitness or proliferative advantage are termed 'drivers', while the rest are termed 'passengers'. The identification of drivers of tumorigenesis has led to the development of therapies that target these alterations and improved patient outcomes [76–78]. Similar approaches

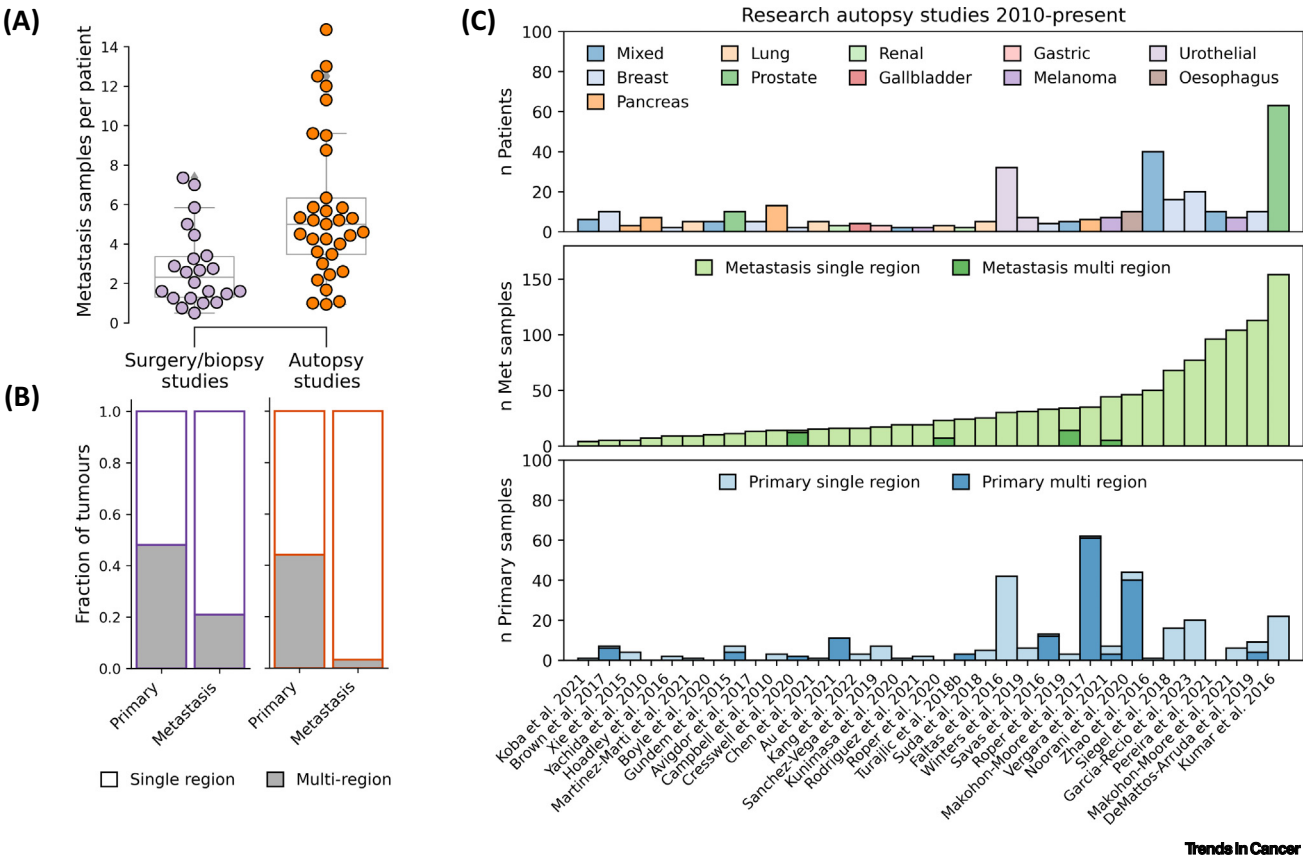

Figure 1. Research autopsies enable multisite sampling of metastases but vary in the availability of paired primary tumours and multiregion sequencing. The data presented in (A–C) were collated from studies of primary-metastasis pairs and autopsy studies that included genomic sequencing data conducted from 2010 to present. Studies including a single patient, patients with nonsolid tumour types, or paediatric patients and studies from which cohort information was not publicly available were excluded. (A) Comparison of the number of individual metastases sampled per patient in studies that collected tissue from surgeries or biopsies (purple) versus studies that collected tissue from research autopsies (orange). Each dot represents a study. (B) Comparison of the fraction of primary tumours or metastases that were multiregion sequenced (grey) or single-region sequenced (white) in studies that collected tissue from surgeries or biopsies (purple) or research autopsies (orange). (C) Summary of the number of patients and tumour type (top panel), the number of single region (light green) and multiregion (dark green) metastasis samples (middle panel), and the number of single region (light blue) and multiregion (dark blue) primary tumour samples (bottom panel) in the research autopsy studies performed from 2010 to present. Corresponding references: [6,16,17,22,25,27,36,39–42,45,48–52,54–56,58–63,66,68,70–73,114,115].

have been taken to identify the genomic alterations that enable cancer cells to metastasise, in the hope that these alterations could be exploited to prevent or treat metastasis.

To identify genes that may be implicated in metastasis, large pan-cancer studies have compared the frequency of somatic mutations and copy number alterations in individual genes in metastases to the frequency of these alterations in unpaired primary tumours [79–83]. The two largest of these studies to date (the MSK-MET cohort, comprising 10 143 panel sequenced metastases, and the Hartwig Medical Foundation cohort, comprising 2250 whole-genome sequenced metastases) both found *TP53* and *CDKN2A* to be the most frequently altered genes in metastases compared with unpaired primary tumours [79,82]. These and other studies consistently report that the frequency of alterations in previously described cancer driver genes, including *PIK3CA*, *APC*, *PTEN*, and MYC, and the overall burden of copy number alterations, including whole-genome duplication events, exceed the rates observed in primary tumours [79–84]. In their analysis of the Hartwig Medical Foundation cohort, Priestley and colleagues also identified genes previously uncharacterised in primary tumours [85], including *MLK4*, a mixed-lineage

Box 1. Identifying metastasising cells and their routes of dissemination

Subpopulations of cancer cells, or 'clones', can be identified within primary or metastatic tumour tissue by their unique complement of somatic mutations and genomic alterations identified using bulk DNA sequencing [116,117]. The ancestral relationships between these clones can then be reconstructed using existing tumour phylogenetic methods [118,119]. The resulting phylogenetic tree encodes information about the ancestral state of these clones, which can be leveraged to infer which clones have metastasised and their route of dissemination. For example, knowing that the ancestor of a clone found in metastasis M2 was previously present in a different metastasis M1 leads to the hypothesis that M2 might have been seeded by a cell or cells of a clone originating from M1 (Figure I). To resolve the most plausible routes of dissemination, computational methods that use statistical or combinatorial approaches have recently been developed [8–10,38,112].

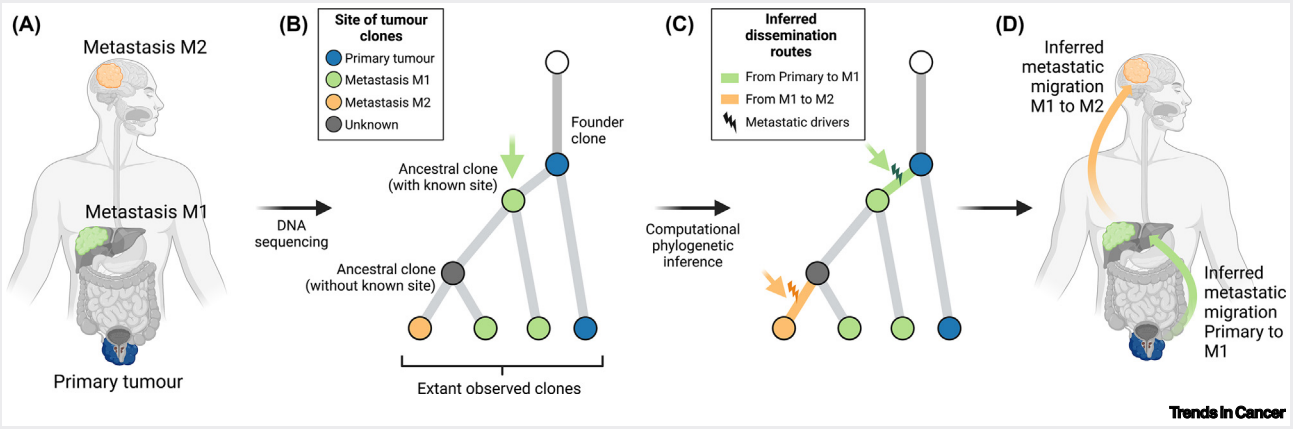

Figure I. Inferring metastasising cells and their routes of dissemination from DNA-sequencing data. (A) DNA sequencing is performed on one or multiple regions from the primary tumour (blue) and metastasis (green for metastasis M1 and orange for metastasis M2). (B) Tumour phylogenetic trees are reconstructed from DNA-sequencing data and describe tumour ancestral relationships such that: leaf nodes correspond to tumour clones observed in sequenced samples (coloured by the anatomical site in which they have been observed); internal nodes correspond to ancestral unobserved clones; and edges represent parental relationships. In addition, the root of the tree (white) corresponds to normal cells and the founder cancer clone is assumed to have been present in the primary tumour (coloured in blue). The site of ancestral clones is generally unknown but can be inferred by evaluating different plausible scenarios (an example is depicted by the green arrow for one ancestral clone). (C) Routes of metastatic cell disseminations (coloured edges) are inferred by existing computational methods (in the example, two inferred migrations correspond to the most parsimonious scenario). Drivers of metastatic dissemination (lightning symbols) can be identified by analysing the genomic alterations identified in the inferred metastasising clones. (D) Routes of disseminations (arrows) explain which cells have seeded each metastasis and whether these cells have disseminated from primary tumours (green arrow) or from other metastases (orange arrow).

kinase involved in the regulation of JNK, P38, and ERK signalling pathways, and *ZFPM1*, encoding a zinc-finger transcription factor protein, to be enriched in pan-cancer and breast cancer metastases, respectively [79].

The identification of recurrently altered genes in metastases in these studies could result from several biological scenarios. First, many of the alterations enriched in metastases will confer resistance to therapy, because a greater proportion of metastases are sampled after treatment. For example, *ESR1* is frequently mutated or epigenetically altered in breast cancer metastases treated with endocrine therapies [22,25,86] and alterations in the androgen receptor pathway are frequent in prostate cancer metastases treated with androgen blockade [45,87]. Second, genomic alterations enriched in metastases may provide a survival advantage to disseminated clones faced with tissue site-specific selective pressures. Nguyen and colleagues described differences in the prevalence of genomic alterations across metastatic sites within cancer types, such as *TP53* mutations, *TERT* amplification, and *EGFR* mutations in brain metastases arising from primary lung adenocarcinoma, suggesting that certain alterations may equip clones to thrive in specific tissue types [82].

A third explanation for the enrichment of certain genomic alterations in metastases is that they confer the ability to metastasise, but do not confer a growth advantage within the primary tumour

Box 2. Influence of sampling on metastatic migration inference

Failing to sample an existing metastatic site that is involved in a metastasis-to-metastasis migration would result in an underestimation of the prevalence of metastasis-to-metastasis migrations and overestimation of primary-to-metastasis migrations (Figure IB). A similar consequence would occur if, due to the illusion of clonality [12,120], a single region biopsy failed to identify all the founding clones within a metastasis (Figure IC). Since the identification of cells capable of metastasising and an assessment of their metastatic potential requires an accurate inference of metastatic migrations routes, robust conclusions about the drivers and steps of metastasis will necessitate comprehensive sampling.

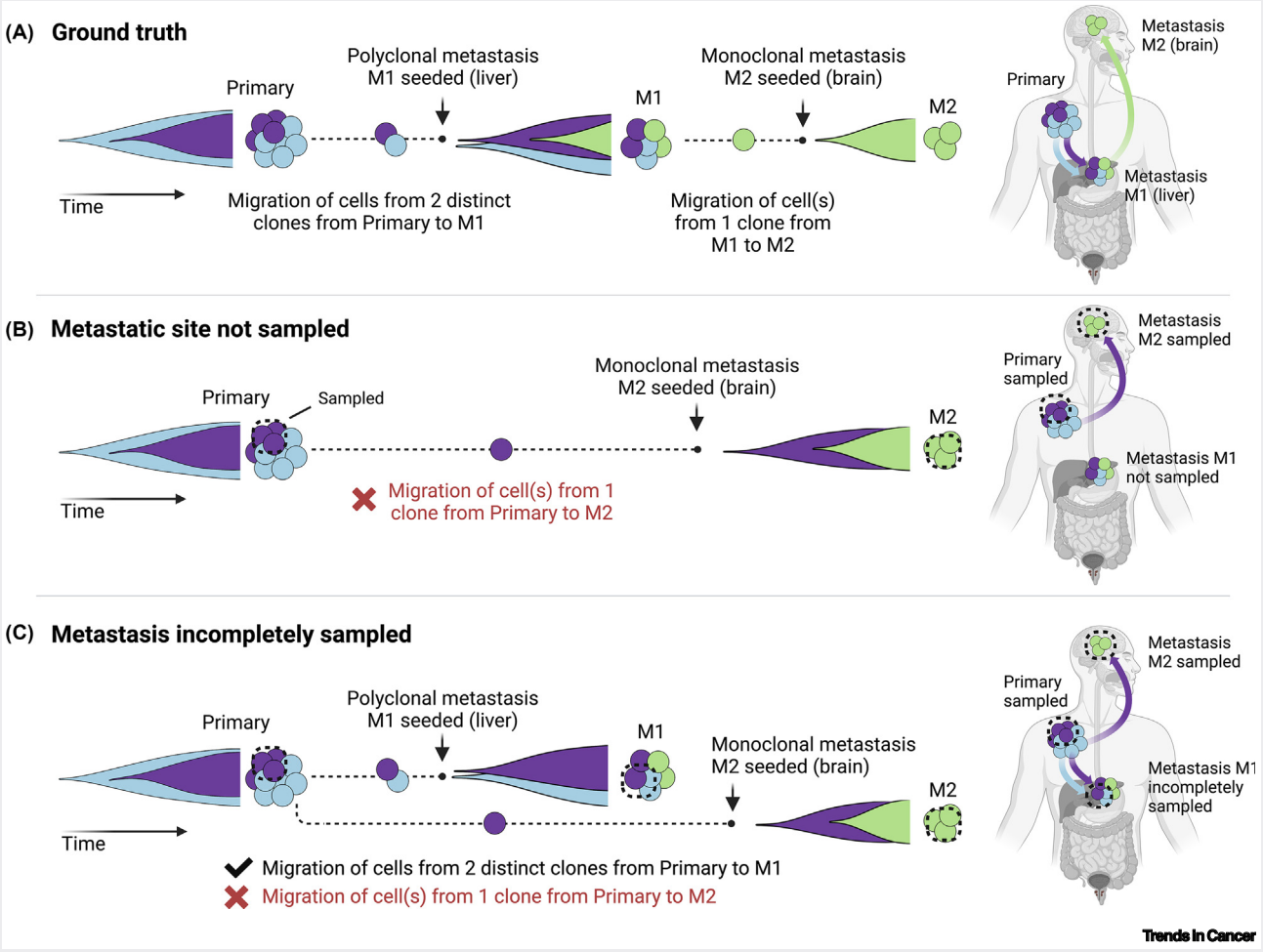

Figure I. Undersampling metastases impacts the inference of metastatic migrations. (A) A ground truth scenario is provided in which cells from two clones (blue and purple) migrate from the primary tumour to the liver, where they seed a polyclonal metastasis, M1. The blue and purple cell populations expand within metastasis M1 (indicated by the increasing width of the fishplot) and the purple clone gives rise to a descendant clone (green). One or more cells from the green clone migrate from M1 in the liver to the brain, where they seed a monoclonal metastasis, M2. The routes of dissemination (arrows coloured to match the clonal lineage of the migrating cells) involve a primary-to-metastasis migration and metastasis-to-metastasis migration. (B) If the primary tumour and metastasis M2 but not metastasis M1 were sampled (broken boxes), the ancestral relationship between the purple clone and the green clone would lead to the incorrect inference (indicated by red text) that cells from only one clone (purple) migrated from the primary to metastasis M2, thereby missing the presence of a polyclonal seeding event and a metastasis-to-metastasis migration. (C) If the primary tumour and both metastases were sampled, but a single region biopsy of metastasis M1 does not capture the presence of the green clone, the polyclonal seeding event from the primary tumour to metastasis M1 would be identified but the metastasis-to-metastasis migration between metastasis M1 and M2 would not. Instead, the route inferred would incorrectly include an additional migration from the primary tumour to metastasis M2.

[88]. Such alterations may be present in minor clones within the primary tumour and not appreciable in bulk sequencing data until the cells disseminate and colonise a distant metastatic site. Distinguishing these metastatic potential-conferring alterations from alterations that arise in metastases in response to therapy or tissue-specific pressures (described above) is critical for understanding the biological processes governing metastasis and, crucially, requires paired primary tumour and metastasis samples to be available for comparison [88]. In studies of primary–metastasis pairs alterations in known cancer drivers, typically acquired early in primary tumour evolution, were usually shared between the primary tumour and metastases [27,36,37,40,51,89]. However, the acquisition of additional alterations specific to clones that metastasise has also been described [35–37]. For example, the availability of multiregion sampling of primary renal cell carcinomas and paired metastases enabled Turajlic and colleagues to distinguish metastasising clones in the primary tumour from clones that did not metastasise [36]. Comparison of the genomic characteristics of these clones revealed that metastasising clones were characterised by chromosomal instability and were enriched for deletions in chromosomes 14q and 9p, encompassing *CDKN2A/B*, relative to those that did not metastasise [36].

Nevertheless, the number of potential metastatic drivers identified to date is limited [80,88]. A possible explanation for this, suggested by recent studies, is that metastasising cells only comprise a small fraction (<3%) of the total cells within primary tumours and metastases and, as such, their alterations are not readily appreciable in bulk DNA-sequencing data from single-region samples [36,90–95]. Beyond the improved ability to detect and characterise metastasising clones afforded by multiregion tissue sampling, single cell DNA and RNA-sequencing technologies are likely to further enhance the resolution with which metastatic clones can be studied by revealing the impact of genomic alterations on clone phenotypes [96–99]. Moreover, the application of these technologies to circulating tumour cells can provide insight into the genomic and phenotypic features that influence the intermediate steps of the metastatic cascade, such as survival in the circulation and tissue colonisation [21], which may not be fully captured by studying primary and metastatic tumour tissue.

The clonal origins of seeding cells and metastases

A further challenge in identifying features that confer metastatic potential is that the ability to metastasise is likely to be context specific. Through the application of a lineage-tracer system to a mouse model of metastasis, Quinn and colleagues revealed that tumour clones characterised by distinct transcriptional drivers have predilections for disseminating at different paces, to specific sites, and along particular routes, indicating that the metastatic phenotype is a continuum [20]. Therefore, knowledge not only of which cells metastasise, but also their routes, destinations, and propensity to do so will be important to reveal the full range of biological features that underpin metastasis.

The classical model of cancer progression suggests that metastases are founded or 'seeded' by a single progenitor cell that disseminates from the primary tumour [1,100]. However, several genomic sequencing studies have found metastases comprised of multiple clones from distinct evolutionary branches, indicating that cells belonging to more than one clone must have migrated to that site [6,16,17,27,34,36,91,101,102]. To distinguish these scenarios, metastases that have been seeded by a cell or cells from a single clone are termed 'monoclonal', while metastases that have been seeded by cells from multiple clone lineages are termed 'polyclonal'.

Whether metastases are monoclonal or polyclonal has implications for how we understand metastatic potential. Polyclonal metastases can arise from a single migration of multiple cells in a cluster or from multiple migrations of individual cells [103–105]. In the former scenario,

cooperation between cells from different clones may engender metastatic potential [92,93]. In the latter, multiple distinct clones may acquire metastatic potential through independent mechanisms [18,66,106,107]. Furthermore, the greater genetic diversity inherent to polyclonal metastases could provide a substrate for the development of therapy resistance [11].

The reported prevalence of monoclonal and polyclonal metastases varies across studies. Some of this variation may be accounted for by differences in definitions used, which we have endeavoured to harmonise here for comparison. In an analysis of 457 paired primary tumour and metastasis samples from 136 patients with breast, colorectal, or lung cancer, Hu and colleagues reported that monoclonal metastases were more common than were polyclonal metastases (76% vs. 24%) [91]. In this study, 80% of primary tumour and metastasis biopsies were single region and single site. In three studies with a greater number of multiregion and multisite metastasis samples, monoclonal metastases were also more common than were polyclonal metastases, but one or more polyclonal metastases were observed in 50% of patients with prostate cancer [6], 73% with breast cancer [102], and 100% with ovarian cancer [9]. Taken together, these studies indicate that metastases can result from the migration of a cell or cells from a single clonal lineage or from the migration of cells from multiple clonal lineages and both modes of spread can coexist in the same patient.

The mono- or polyclonality of metastases provides an indication of the number of clones that have successfully metastasised. As such, variation in the prevalence of polyclonal metastases across organ sites might reveal differences in how 'permissive' certain tissue types are to colonisation. Corroborating this notion, multiple studies have reported that lymph node metastases are more likely to be polyclonal compared with distant metastases [32,34,91,108]. Several properties of lymph nodes might support the survival of metastasising cells relative to distant sites, including the shorter distance between the site of origin and the nearest draining lymph node [109], the reduced mechanical stress during migration within the lymphatic system [110], and the less hostile chemical composition of lymph compared with blood [111]. These properties culminate in selective pressures distinct from those endured by cancer cells migrating to distant sites, reflected by differences in the genomic composition of lymph node and distant metastases [91]. Moreover, the observation that fewer clones successfully seed distant metastases implies that the clones that do may be particularly efficient at metastasising and, thus, may provide insight into the characteristics that confer metastatic potential.

Beyond anatomical boundaries, chemotherapy would be expected to impose an additional barrier to metastasis. Hu and colleagues found that metastases exposed to therapy were more likely to be monoclonal, in keeping with therapy reducing the number of successful seeding events. Moreover, in contrast to untreated metastases, they found evidence that treatment-exposed metastases were more likely to be seeded by a minor clone within the primary tumour equipped with subclonal driver alterations [91]. These findings have several implications of clinical importance. First, the clones that give rise to metastases in patients treated with adjuvant chemotherapy, even if fewer, are likely to be chemotherapy resistant and, thus, may contribute to treatment failure in patents with metastatic cancer. Second, exposure to therapy can remodel the clonal architecture of metastases and, therefore, may account for some of the variation in the prevalence of polyclonal metastases reported in the literature. Given that most metastases in studies of primary–metastasis pairs are sampled after exposure to therapy, the prevalence of polyclonal metastases in the natural disease course is likely to be greater than currently appreciated [91].

In silico simulations [91] and analyses of multiregion sampled metastases [38] have demonstrated that multiregion sequencing of primary tumours and metastases together increases the accuracy

of classifying monoclonal and polyclonal seeding compared with single-region sampling. Therefore, sufficient sampling of metastases such that polyclonality, if present, can be detected, is critical to understand the selective pressures that shape patterns of metastatic dissemination and, ultimately, the genetic characteristics of metastases.

Routes of cancer dissemination

Until recently, the cells that seed metastases were thought to originate from a single source: the primary tumour. However, studies in which multiple metastatic sites were sequenced have revealed the possibility that metastases can additionally be seeded by cells from other metastases [6,9,16–19,102]. This possibility supports the idea that several cells and several successive migrations contribute to the patterns of metastatic disease. One implication of this more complex model of metastasis is that the frequency and route of migrations are likely to determine the extent of clonal diversity found in metastases [107]. Given that a greater degree of tumour heterogeneity provides a substrate for therapy resistance [11], the specific steps that culminate in metastasis may influence therapy response rates. Furthermore, tracing migration routes of cells between organ sites could provide insight into cellular traits that facilitate survival within certain tissue microenvironments (or 'soils') [1] or, conversely, traits that enable cells to colonise a multitude of tissues, thus characterising the full spectrum of metastatic potential [20].

Studies that have traced the routes of metastatic migrations have described two dominant patterns: one in which multiple metastases are seeded by cells from the primary tumour, and another in which cells from a metastasis, originally seeded by the primary tumour, go on to seed another metastasis in a sequential manner. Noorani and colleagues found that, in ten patients with oesophageal cancer, metastases were seeded by cells from multiple clones from the primary tumour in parallel, a pattern they term the 'clonal diaspora', with no evidence of metastasis-to-metastasis migrations [66]. By contrast, in a cohort of ten patients with metastatic breast cancer who underwent research autopsies, sequential seeding of metastases by metastases was observed in all cases [16]. Gundem and colleagues also reported that metastasis-to-metastasis migrations were common in an autopsy cohort, occurring in eight out of ten patients with metastatic prostate cancer [6].

These studies interpreted the ancestral–descendant relationship between clones in two sites to represent the migration of cells. However, a cellular migration is not always a linear, unidirectional process and, therefore, it is not equivalent to a phylogenetic relationship; neither does the phylogenetic tree topology encode the anatomical location of ancestral clones [8,10,112]. Instead, resolving cell migrations between a primary tumour and multiple metastatic sites requires a method to determine the most likely combination of all possible migration routes. MACHINA is a computational algorithm that implements a theoretical framework based on the principle of maximum parsimony, that the most likely combination will involve the least number of migrations, to determine the migration pattern [10]. Applied to the cohort of patients with prostate cancer presented by Gundem and colleagues [6], MACHINA corroborated the presence of polyclonal metastases, but inferred alternative migration routes. In three patients, parallel seeding of all metastases from the primary was found to be more likely than the originally inferred metastasis-to-metastasis migrations [10]. Alternative more simple migration routes were also identified when MACHINA was applied to previously published series of patients with breast [17] or ovarian cancer [9], highlighting that uncertainty in bulk sequencing data and lack of an evolution-based statistical framework for determining migration routes can impede inferences regarding how cancer progresses.

An additional factor confounding the accurate inference of migration routes is the number of primary tumour regions and metastases sampled per patient. El Kebir and colleagues evaluated

the performance of MACHINA with varying sequencing coverage, sample purity, and number of sequenced samples, and found that the number of samples per patient had the largest impact on migration route precision and recall [10]. The authors did not distinguish between the number of multisite or multiregion samples, but either could conceivably alter inferred migration routes (Box 2).

Biological rationale for research autopsies

As the complexities of the development of metastasis become apparent, so too do the steps of the process that need further investigation. These steps are often the most challenging to study and, as such, progressing the understanding of metastasis will require an approach that leverages the strengths of past studies and adopts new strategies to overcome their limitations.

The evidence summarised in this review indicates that comprehensive sampling of primary tumours and metastases at multiple timepoints throughout the disease course, combined with the use of established bulk and emerging single cell sequencing technologies that enable accurate mapping of tumour evolution, will be fundamental to understand the determinants of metastasis. This includes observations that minor clones often have a key role in metastasis [36,90,92–95], that multiple distinct clones can acquire metastatic potential either individually [18,66,106] or in cooperation with one another [92,93,103–105], and that the clonal composition of metastases and metastatic migration routes can be informative as to how selective pressures can influence the metastatic cascade [20,32,34,91]. The scarcity of studies that include multiregion sampling of primary tumours and multisite sampling of metastases (3% of patients in the studies surveyed had both) draws attention to an addressable gap in metastasis research. Establishing studies with longitudinal tissue acquisition strategies that enable resected primary tumour samples, circulating tumour cells or DNA, and relapse or progression tissue biopsies to be collected from patients who undergo a research autopsy at the time of death, is one way to bridge this gap (see Outstanding questions).

An example of such an approach is being trialled within the context of PEACE, a national research autopsy programme that recruits patients with metastatic cancer in the UK. With over 380 patients enrolled and over 220 research autopsies performed, PEACE has provided tissue to support the work of cancer researchers across a range of fields. The PEACE autopsy procedure involves a review of prior radiological imaging and pathologist-led examination to ensure samples (multiregion where possible) are taken from all identifiable metastatic sites. The pan-cancer, multicentre nature of PEACE enables the programme to dovetail with other clinical studies to create longitudinal datasets particularly suited to studying the progression of cancer from diagnosis to death. For example, patients with early-stage non-small cell lung cancer recruited in the Tracking Cancer Evolution through Therapy (Rx) (TRACERx; NCT01888601) study, who subsequently develop metastatic disease, are co-recruited into PEACE with the aim of studying tumour evolution from early- to late-stage disease and in response to therapy. For these patients, multiregion sampling of resected primary tumours and relapse or progression biopsy samples available through TRACERx, combined with metastases sampled during PEACE research autopsies, provides the opportunity to reconstruct the steps of tumour evolution, from initiation to metastasis, while minimising the risk of sampling bias.

Concluding remarks

The capacity to complete the metastatic cascade likely relies on the dynamic cooperation of cancer cells with one another [31,92], their microenvironment [26], and their host [113]. Disentangling these interactions requires an approach that takes into account both cancer cell-intrinsic and -extrinsic features, which can be enabled by the large volumes of tissue accessible through research autopsies. Furthermore, understanding these interactions in the context of

Outstanding questions

Does *de novo* metastatic disease follow the same pattern of progression as early-stage disease that subsequently relapses? It is unknown whether the patterns of metastatic dissemination observed in studies to date are representative of the pace and sequence of events that occur in patients who present with *de novo* metastatic disease. These patients are often excluded from translational research studies because they do not undergo resection of their primary tumour and rarely have multiregion biopsies taken during their clinical course. By contrast, patients who present with *de novo* metastatic disease comprise approximately half of the participants in research autopsy studies. A study recruiting patients with operable early-stage disease alongside patients with *de novo* metastatic disease that collects samples appropriate to the clinical setting during life and includes a research autopsy at the end of life would begin to address this important outstanding question.

How does therapy influence the pace and routes of metastasis? Research autopsy studies have consistently identified multiple, parallel alterations associated with therapy resistance across different metastatic sites. Due to the absence of samples taken from the same patients at earlier timepoints, the ability of these studies to map the order and timing of such alterations, which could be exploited to optimise therapeutic approaches, has been limited. The recruitment of patients who have participated in clinical drug trials, particularly those from whom pretherapy tissue samples were collected, to research autopsy studies would increase the availability of post-treatment samples and allow the emergence of therapy resistance to be tracked over both time and space.

disease setting from early to late stage, and in response to therapy, can be achieved by leveraging studies of tumour evolution. A combined strategy that incorporates longitudinal genomic studies, such as TRACERx, with research autopsies, such as PEACE, entailing comprehensive sampling of all sites of disease, has the real potential to further our current understanding of metastasis and potentially help identify therapeutic targets that may prevent or delay the onset of new metastases or progression of established metastases.

Acknowledgments

M.J-H. is a CRUK Career Establishment Awardee and has received funding from CRUK, IASLC International Lung Cancer Foundation, Lung Cancer Research Foundation, Rosetrees Trust, UKI NETs, NIHR, NIHR UCLH Biomedical Research Centre. C.S. acknowledges grant support from Pfizer, AstraZeneca, Bristol Myers Squibb, Roche-Ventana, Boehringer-Ingelheim, ArcherDx Inc. (collaboration in minimal residual disease sequencing technologies), and Ono Pharmaceutical.

Declaration of interests

S.H., P.F., and S.Z. have no conflicts of interests to declare. M.J-H. has consulted for, and is a member of, the Achilles Therapeutics Scientific Advisory Board and Steering Committee, has received speaker honoraria from Pfizer, Astex Pharmaceuticals, Oslo Cancer Cluster, and holds a patent PCT/US2017/028013 relating to methods for lung cancer detection. C.S. is an AstraZeneca Advisory Board Member and Chief Investigator for the MeRmaiD1 clinical trial and has consulted for Amgen, Pfizer, Novartis, GlaxoSmithKline, MSD, Bristol Myers Squibb, AstraZeneca, Illumina, Genentech, Roche-Ventana, GRAIL, Medicxi, Bicycle Therapeutics, Metabomed, and the Sarah Cannon Research Institute; has stock options in ApoGen Biotechnologies, Epic Bioscience, and GRAIL; and has stock options and is co-founder of Achilles Therapeutics. C.S. holds patents relating to assay technology to detect tumour recurrence (PCT/GB2017/053289); to targeting neoantigens (PCT/EP2016/059401), identifying patent response to immune checkpoint blockade (PCT/EP2016/071471), determining HLA LOH (PCT/GB2018/052004), predicting survival rates of patients with cancer (PCT/GB2020/050221), to treating cancer by targeting insertion/deletion mutations (PCT/GB2018/051893), identifying insertion/deletion mutation targets (PCT/GB2018/051892); methods for lung cancer detection (PCT/US2017/028013), identifying responders to cancer treatment (PCT/GB2018/051912); and a patent application to determine methods and systems for tumour monitoring (PCT/EP2022/077987).

References

1. Fidler, I.J. (2003) The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. *Nat. Rev. Cancer* 3, 453–458
2. Massague, J. and Obenauf, A.C. (2016) Metastatic colonization by circulating tumour cells. *Nature* 529, 298–306
3. Lambert, A.W. *et al.* (2017) Emerging biological principles of metastasis. *Cell* 168, 670–691
4. Valastyan, S. and Weinberg, R.A. (2011) Tumor metastasis: molecular insights and evolving paradigms. *Cell* 147, 275–292
5. Turajlic, S. and Swanton, C. (2016) Metastasis as an evolutionary process. *Science* 352, 169–175
6. Gundem, G. *et al.* (2015) The evolutionary history of lethal metastatic prostate cancer. *Nature* 520, 353–357
7. Comen, E. *et al.* (2011) Clinical implications of cancer self-seeding. *Nat. Rev. Clin. Oncol.* 8, 369–377
8. Reiter, J.G. *et al.* (2017) Reconstructing metastatic seeding patterns of human cancers. *Nat. Commun.* 8, 14114
9. McPherson, A. *et al.* (2016) Divergent modes of clonal spread and intraperitoneal mixing in high-grade serous ovarian cancer. *Nat. Genet.* 48, 758–767
10. El-Kebir, M. *et al.* (2018) Inferring parsimonious migration histories for metastatic cancers. *Nat. Genet.* 50, 718–726
11. McGranahan, N. and Swanton, C. (2017) Clonal heterogeneity and tumor evolution: past, present, and the future. *Cell* 168, 613–628
12. Gerlinger, M. *et al.* (2014) Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat. Genet.* 46, 225–233
13. Jamal-Hanjani, M. *et al.* (2017) Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* 376, 2109–2121
14. Turajlic, S. *et al.* (2018) Deterministic evolutionary trajectories influence primary tumor growth: TRACERx renal. *Cell* 173, 595–610
15. Zaccaria, S. and Raphael, B.J. (2020) Accurate quantification of copy-number aberrations and whole-genome duplications in multi-sample tumor sequencing data. *Nat. Commun.* 11, 4301
16. Brown, D. *et al.* (2017) Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. *Nat. Commun.* 8, 14944
17. Hoadley, K.A. *et al.* (2016) Tumor evolution in two patients with basal-like breast cancer: a retrospective genomics study of multiple metastases. *PLoS Med.* 13, e1002174
18. Sanborn, J.Z. *et al.* (2015) Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10995–11000
19. Choi, Y.J. *et al.* (2017) Intraindividual genomic heterogeneity of high-grade serous carcinoma of the ovary and clinical utility of ascitic cancer cells for mutation profiling. *J. Pathol.* 241, 57–66
20. Quinn, J.J. *et al.* (2021) Single-cell lineages reveal the rates, routes, and drivers of metastasis in cancer xenografts. *Science* 371, eabc1944
21. Keller, L. and Pantel, K. (2019) Unravelling tumour heterogeneity by single-cell profiling of circulating tumour cells. *Nat. Rev. Cancer* 19, 553–567
22. Cresswell, G.D. *et al.* (2020) Mapping the breast cancer metastatic cascade onto ctDNA using genetic and epigenetic clonal tracking. *Nat. Commun.* 11, 1446
23. Abbosh, C. *et al.* (2023) Tracking early lung cancer metastatic dissemination in TRACERx using ctDNA. *Nature* 616, 553–562
24. LaFave, L.M. *et al.* (2020) Epigenomic state transitions characterize tumor progression in mouse lung adenocarcinoma. *Cancer Cell* 38, 212–228
25. Garcia-Recio, S. *et al.* (2023) Multiomics in primary and metastatic breast tumors from the AURORA US network finds microenvironment and epigenetic drivers of metastasis. *Nat. Cancer* 4, 128–147

26. Vazquez-Garcia, I. *et al.* (2022) Ovarian cancer mutational processes drive site-specific immune evasion. *Nature* 612, 778–786
27. De Mattos-Arruda, L. *et al.* (2019) The genomic and immune landscapes of lethal metastatic breast cancer. *Cell Rep.* 27, 2690–2708
28. Er, E.E. *et al.* (2018) Pericyte-like spreading by disseminated cancer cells activates YAP and MRTF for metastatic colonization. *Nat. Cell Biol.* 20, 966–978
29. Esposito, M. *et al.* (2019) Bone vascular niche E-selectin induces mesenchymal-epithelial transition and Wnt activation in cancer cells to promote bone metastasis. *Nat. Cell Biol.* 21, 627–639
30. Cleary, A.S. *et al.* (2014) Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* 508, 113–117
31. Marusyk, A. *et al.* (2014) Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* 514, 54–58
32. Reiter, J.G. *et al.* (2020) Lymph node metastases develop through a wider evolutionary bottleneck than distant metastases. *Nat. Genet.* 52, 692–700
33. Hu, Z. *et al.* (2019) Quantitative evidence for early metastatic seeding in colorectal cancer. *Nat. Genet.* 51, 1113–1122
34. Naxerova, K. *et al.* (2017) Origins of lymphatic and distant metastases in human colorectal cancer. *Science* 357, 55–60
35. Gibson, W.J. *et al.* (2016) The genomic landscape and evolution of endometrial carcinoma progression and abdominopelvic metastasis. *Nat. Genet.* 48, 848–855
36. Turajlic, S. *et al.* (2018) Tracking cancer evolution reveals constrained routes to metastases: TRACERx renal. *Cell* 173, 581–594
37. Yates, L.R. *et al.* (2017) Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* 32, 169–184
38. Al Bakir, M. *et al.* (2023) The evolution of non-small cell lung cancer metastases in TRACERx. *Nature.* 616, 534–542
39. Pereira, B. *et al.* (2021) Cell-free DNA captures tumor heterogeneity and driver alterations in rapid autopsies with pre-treated metastatic cancer. *Nat. Commun.* 12, 3199
40. Siegel, M.B. *et al.* (2018) Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J. Clin. Invest.* 128, 1371–1383
41. Avigdor, B.E. *et al.* (2017) Mutational profiles of breast cancer metastases from a rapid autopsy series reveal multiple evolutionary trajectories. *JCI Insight* 2, e96896
42. Savas, P. *et al.* (2016) The subclonal architecture of metastatic breast cancer: results from a prospective community-based rapid autopsy program "CASCADE". *PLoS Med.* 13, e1002204
43. Da Silva, L. *et al.* (2010) HER3 and downstream pathways are involved in colonization of brain metastases from breast cancer. *Breast Cancer Res.* 12, R46
44. Corella, A.N. *et al.* (2020) Identification of therapeutic vulnerabilities in small-cell neuroendocrine prostate cancer. *Clin. Cancer Res.* 26, 1667–1677
45. Kumar, A. *et al.* (2016) Substantial interindividual and limited intraindividual genomic diversity among tumors from men with metastatic prostate cancer. *Nat. Med.* 22, 369–378
46. Friedlander, T.W. *et al.* (2012) Common structural and epigenetic changes in the genome of castration-resistant prostate cancer. *Cancer Res.* 72, 616–625
47. Grasso, C.S. *et al.* (2012) The mutational landscape of lethal castration-resistant prostate cancer. *Nature* 487, 239–243
48. Vergara, I.A. *et al.* (2021) Evolution of late-stage metastatic melanoma is dominated by aneuploidy and whole genome doubling. *Nat. Commun.* 12, 1434
49. Makohon-Moore, A.P. *et al.* (2021) The genetic evolution of treatment-resistant cutaneous, acral, and uveal melanomas. *Clin. Cancer Res.* 27, 1516–1525
50. Rodriguez, D.A. *et al.* (2021) Multiregional genetic evolution of metastatic uveal melanoma. *NPJ Genom. Med.* 6, 70
51. Makohon-Moore, A.P. *et al.* (2017) Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. *Nat. Genet.* 49, 358–366
52. Yachida, S. *et al.* (2010) Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* 467, 1114–1117
53. Yachida, S. *et al.* (2012) Clinical significance of the genetic landscape of pancreatic cancer and implications for identification of potential long-term survivors. *Clin. Cancer Res.* 18, 6339–6347
54. Campbell, P.J. *et al.* (2010) The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467, 1109–1113
55. Xie, T. *et al.* (2015) Whole exome sequencing of rapid autopsy tumors and xenograft models reveals possible driver mutations underlying tumor progression. *PLoS One* 10, e0142631
56. Martinez-Marti, A. *et al.* (2021) Genetic evolution to tyrosine kinase inhibitory therapy in patients with EGFR-mutated non-small-cell lung cancer. *Br. J. Cancer* 125, 1561–1569
57. Suda, K. *et al.* (2021) Inter- and intratumor heterogeneity of EGFR compound mutations in non-small cell lung cancers: analysis of five cases. *Clin. Lung Cancer* 22, e141–e145
58. Chen, H.Z. *et al.* (2021) Genomic and transcriptomic characterization of relapsed SCLC through rapid research autopsy. *JTO Clin. Res. Rep.* 2, 100164
59. Roper, N. *et al.* (2020) Clonal evolution and heterogeneity of osimertinib acquired resistance mechanisms in EGFR mutant lung cancer. *Cell Rep. Med.* 1, 100007
60. Roper, N. *et al.* (2019) APOBEC mutagenesis and copy-number alterations are drivers of proteogenomic tumor evolution and heterogeneity in metastatic thoracic tumors. *Cell Rep.* 26, 2651–2666
61. Suda, K. *et al.* (2016) Heterogeneity in resistance mechanisms causes shorter duration of epidermal growth factor receptor kinase inhibitor treatment in lung cancer. *Lung Cancer* 91, 36–40
62. Boyle, T.A. *et al.* (2020) A community-based lung cancer rapid tissue donation protocol provides high-quality drug-resistant specimens for proteogenomic analyses. *Cancer Med.* 9, 225–237
63. Kunimasa, K. *et al.* (2020) Rapid progressive lung cancers harbouring multiple clonal driver mutations with big bang evolution model. *Cancer Genet.* 241, 51–56
64. Patch, A.M. *et al.* (2015) Whole-genome characterization of chemoresistant ovarian cancer. *Nature* 521, 489–494
65. Brastianos, P.K. *et al.* (2017) Resolving the phylogenetic origin of glioblastoma via multifocal genomic analysis of pre-treatment and treatment-resistant autopsy specimens. *NPJ Precis. Oncol.* 1, 33
66. Noorani, A. *et al.* (2020) Genomic evidence supports a clonal diaspora model for metastases of esophageal adenocarcinoma. *Nat. Genet.* 52, 74–83
67. Parikh, A.R. *et al.* (2019) Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. *Nat. Med.* 25, 1415–1421
68. Sanchez-Vega, F. *et al.* (2019) EGFR and MET amplifications determine response to HER2 inhibition in ERBB2-amplified esophagogastric cancer. *Cancer Discov.* 9, 199–209
69. Siravegna, G. *et al.* (2018) Radiologic and genomic evolution of individual metastases during hER2 blockade in colorectal cancer. *Cancer Cell* 34, 148–162
70. Kang, M. *et al.* (2022) Gallbladder adenocarcinomas undergo subclonal diversification and selection from precancerous lesions to metastatic tumors. *eLife* 11, e78636
71. Au, L. *et al.* (2021) Determinants of anti-PD-1 response and resistance in clear cell renal cell carcinoma. *Cancer Cell* 39, 1497–1518
72. Winters, B.R. *et al.* (2019) Genomic distinctions between metastatic lower and upper tract urothelial carcinoma revealed through rapid autopsy. *JCI Insight* 5, e128728
73. Faltas, B.M. *et al.* (2016) Clonal evolution of chemotherapy-resistant urothelial carcinoma. *Nat. Genet.* 48, 1490–1499
74. Iacobuzio-Donahue, C.A. *et al.* (2019) Cancer biology as revealed by the research autopsy. *Nat. Rev. Cancer* 19, 686–697
75. Pisapia, D.J. *et al.* (2017) Next-generation rapid autopsies enable tumor evolution tracking and generation of preclinical models. *JCO Precis. Oncol.* 2017, PO.16.00038
76. Slamon, D.J. *et al.* (2001) Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N. Engl. J. Med.* 344, 783–792
77. Shaw, A.T. *et al.* (2014) Ceritinib in ALK-rearranged non-small-cell lung cancer. *N. Engl. J. Med.* 370, 1189–1197
78. Mok, T.S. *et al.* (2009) Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N. Engl. J. Med.* 361, 947–957

79. Priestley, P. *et al.* (2019) Pan-cancer whole-genome analyses of metastatic solid tumours. *Nature* 575, 210–216
80. Robinson, D.R. *et al.* (2017) Integrative clinical genomics of metastatic cancer. *Nature* 548, 297–303
81. Pleasance, E. *et al.* (2020) Pan-cancer analysis of advanced patient tumors reveals interactions between therapy and genomic landscapes. *Nat. Cancer* 1, 452–468
82. Nguyen, B. *et al.* (2022) Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *Cell* 185, 563–575
83. Zehir, A. *et al.* (2017) Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* 23, 703–713
84. Pradat, Y. *et al.* (2023) Integrative pan-cancer genomic and transcriptomic analyses of refractory metastatic cancer. *Cancer Discov.* 13, 1116–1143
85. Bailey, M.H. *et al.* (2018) Comprehensive characterization of cancer driver genes and mutations. *Cell* 173, 371–385
86. Jeselsohn, R. *et al.* (2018) Allele-specific chromatin recruitment and therapeutic vulnerabilities of ESR1 activating mutations. *Cancer Cell* 33, 173–186
87. Armenia, J. *et al.* (2018) The long tail of oncogenic drivers in prostate cancer. *Nat. Genet.* 50, 645–651
88. Birkbak, N.J. and McGranahan, N. (2020) Cancer genome evolutionary trajectories in metastasis. *Cancer Cell* 37, 8–19
89. Reiter, J.G. *et al.* (2018) Minimal functional driver gene heterogeneity among untreated metastases. *Science* 361, 1033–1037
90. Haffner, M.C. *et al.* (2013) Tracking the clonal origin of lethal prostate cancer. *J. Clin. Invest.* 123, 4918–4922
91. Hu, Z. *et al.* (2020) Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat. Genet.* 52, 701–708
92. Janiszewska, M. *et al.* (2019) Subclonal cooperation drives metastasis by modulating local and systemic immune microenvironments. *Nat. Cell Biol.* 21, 879–888
93. Kok, S.Y. *et al.* (2021) Malignant subclone drives metastasis of genetically and phenotypically heterogenous cell clusters through fibrotic niche generation. *Nat. Commun.* 12, 863
94. Angelova, M. *et al.* (2018) Evolution of metastases in space and time under immune selection. *Cell* 175, 751–765
95. Hong, M.K. *et al.* (2015) Tracking the origins and drivers of subclonal metastatic expansion in prostate cancer. *Nat. Commun.* 6, 6605
96. Zaccaria, S. and Raphael, B.J. (2021) Characterizing allele- and haplotype-specific copy numbers in single cells with CHISEL. *Nat. Biotechnol.* 39, 207–214
97. Laks, E. *et al.* (2019) Clonal decomposition and DNA replication states defined by scaled single-cell genome sequencing. *Cell* 179, 1207–1221
98. Funnell, T. *et al.* (2022) Single-cell genomic variation induced by mutational processes in cancer. *Nature* 612, 106–115
99. Minussi, D.C. *et al.* (2021) Breast tumours maintain a reservoir of subclonal diversity during expansion. *Nature* 592, 302–308
100. Talmadge, J.E. *et al.* (1982) Evidence for the clonal origin of spontaneous metastases. *Science* 217, 361–363
101. Brastianos, P.K. *et al.* (2015) Genomic characterization of brain metastases reveals branched evolution and potential therapeutic targets. *Cancer Discov.* 5, 1164–1177
102. Ullah, I. *et al.* (2018) Evolutionary history of metastatic breast cancer reveals minimal seeding from axillary lymph nodes. *J. Clin. Invest.* 128, 1355–1370
103. Aceto, N. *et al.* (2014) Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 158, 1110–1122
104. Cheung, K.J. *et al.* (2016) Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc. Natl. Acad. Sci. U. S. A.* 113, E854–E863
105. Maddipati, R. and Stanger, B.Z. (2015) Pancreatic cancer metastases harbor evidence of polyclonality. *Cancer Discov.* 5, 1086–1097
106. Leung, M.L. *et al.* (2017) Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res.* 27, 1287–1299
107. Heyde, A. *et al.* (2019) Consecutive seeding and transfer of genetic diversity in metastasis. *Proc. Natl. Acad. Sci. U. S. A.* 116, 14129–14137
108. Ulintz, P.J. *et al.* (2018) Lymph node metastases in colon cancer are polyclonal. *Clin. Cancer Res.* 24, 2214–2224
109. Leong, S.P. *et al.* (2022) Molecular mechanisms of cancer metastasis via the lymphatic versus the blood vessels. *Clin. Exp. Metastasis* 39, 159–179
110. Sleeman, J.P. (2000) The lymph node as a bridgehead in the metastatic dissemination of tumors. *Recent Results Cancer Res.* 157, 55–81
111. Ubellacker, J.M. *et al.* (2020) Lymph protects metastasizing melanoma cells from ferroptosis. *Nature* 585, 113–118
112. Hong, W.S. *et al.* (2015) Inferring the origin of metastases from cancer phylogenies. *Cancer Res.* 75, 4021–4025
113. Hursting, S.D. and Berger, N.A. (2010) Energy balance, host-related factors, and cancer progression. *J. Clin. Oncol.* 28, 4058–4065
114. Koba, H. *et al.* (2021) Molecular features of tumor-derived genetic alterations in circulating cell-free DNA in virtue of autopsy analysis. *Sci. Rep.* 11, 8398
115. Zhao, Z.M. *et al.* (2016) Early and multiple origins of metastatic lineages within primary tumors. *Proc. Natl. Acad. Sci. U. S. A.* 113, 2140–2145
116. Nowell, P.C. (1976) The clonal evolution of tumor cell populations. *Science* 194, 23–28
117. Burrell, R.A. *et al.* (2013) The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501, 338–345
118. Schwartz, R. and Schaffer, A.A. (2017) The evolution of tumour phylogenetics: principles and practice. *Nat. Rev. Genet.* 18, 213–229
119. Tarabichi, M. *et al.* (2021) A practical guide to cancer subclonal reconstruction from DNA sequencing. *Nat. Methods* 18, 144–155
120. Gerlinger, M. *et al.* (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366, 883–892
121. Spain, L. *et al.* (2023) Late-stage metastatic melanoma emerges through a diversity of evolutionary pathways. *Cancer Discov.* Published online March 28, 2023. https://doi.org/10.1158/2159-8290.cd-22-1427