

## REVIEW ARTICLE OPEN



## Subclonal heterogeneity and evolution in breast cancer

Ioanna Mavrommati<sup>1</sup>, Flora Johnson<sup>1</sup>, Gloria V. Echeverria<sup>2,3,4,5</sup> and Rachael Natrajan<sup>1</sup>✉

Subclonal heterogeneity and evolution are characteristics of breast cancer that play a fundamental role in tumour development, progression and resistance to current therapies. In this review, we focus on the recent advances in understanding the epigenetic and transcriptomic changes that occur within breast cancer and their importance in terms of cancer development, progression and therapy resistance with a particular focus on alterations at the single-cell level. Furthermore, we highlight the utility of using single-cell tracing and molecular barcoding methodologies in preclinical models to assess disease evolution and response to therapy. We discuss how the integration of single-cell profiling from patient samples can be used in conjunction with results from preclinical models to untangle the complexities of this disease and identify biomarkers of disease progression, including measures of intra-tumour heterogeneity themselves, and how enhancing this understanding has the potential to uncover new targetable vulnerabilities in breast cancer.

*npj Breast Cancer* (2021)7:155; <https://doi.org/10.1038/s41523-021-00363-0>

## INTRODUCTION

Breast cancer is a heterogeneous disease, driven by a myriad of genetic and non-genetic alterations that govern its clinical behaviour<sup>1</sup>. Tumour heterogeneity (TH) can exist at different levels (for example; genomic, transcriptomic and proteomic) and refers to the observation that cancer cells can have distinct molecular profiles from one another. This diversity of tumour cells' profiles, can be either observed between tumours, known as inter-tumour heterogeneity (inter-TH), or within the same tumours, known as intra-tumour heterogeneity (intra-TH)<sup>2</sup>.

Inter-TH in breast cancer has led to the classification based on histology and expression profiles of the molecular markers; oestrogen receptor (ER), progesterone receptor (PR) and the overexpression or gene amplification of human epidermal growth factor receptor 2 (HER2). The expression profiles of these markers have led to three main clinical breast cancer subgroups for which treatment regimes are based; ER+/HER2-, HER2+ or triple-negative breast cancer (TNBC) which lack ER, PR and HER2 expression<sup>3,4</sup>. The seminal studies using gene expression profiling, have further subdivided breast cancers into molecular and transcriptomic subtypes, which are of prognostic and predictive importance<sup>5–9</sup>. Together, these studies highlight the innate inter-TH amongst breast cancer patients.

Intra-TH escalates the complexity of the disease in many aspects including our understanding of cancer development, progression, metastasis, therapy response and acquired drug resistance. Two non-mutually exclusive models of intra-TH have been proposed, the cancer stem cell model and the clonal evolution model<sup>10</sup>, which are extensively reviewed elsewhere<sup>2,11–13</sup>. Despite the development of next-generation sequencing having enabled in depth analysis of intra-TH and tumour evolution, these two factors contribute to systemic treatment failure by initiating phenotypic diversity enabling drug resistance to emerge.

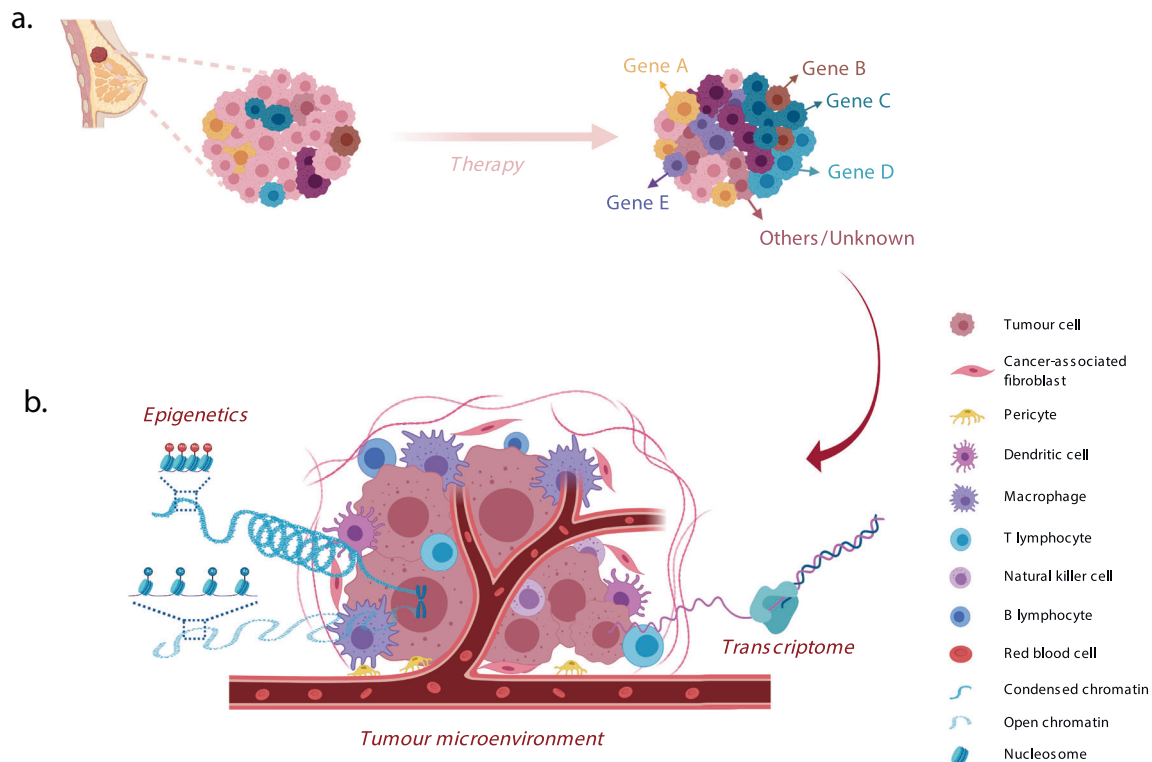
In this review, we summarise the importance of next-generation sequencing in identifying intra-TH in breast cancer at a genetic level, identify the clinical impacts and challenges of intra-TH to patients' treatment outcomes and how genetic intra-TH can be

assessed by single-cell approaches. Furthermore, we explore the non-genetic forms of intra-TH during breast cancer evolution, drug resistance and more importantly how single-cell approaches in conjunction with appropriate preclinical models can be used to untangle the complexity of the disease.

## Genetic intra-TH in breast cancer

In the past decade, through the development and implementation of next-generation sequencing, numerous studies have reported the existence of intra-TH in breast cancer at the genetic level. This has been attributed to chromosomal or genomic alterations affecting biological processes and functions<sup>3</sup>. For instance, through large-scale genome analysis of primary human breast cancers, the complex mutational landscape of breast cancer has been mapped, highlighting the existence of clonal tumour subpopulations in breast cancer<sup>14–22</sup>. These studies in general have highlighted that a large majority of mutations detected in primary breast cancers are subclonal<sup>17</sup> and not necessarily evenly distributed spatially among individual tumours highlighting the need for multiple sampling<sup>21</sup>. In primary untreated breast cancers, multi-region sequencing has highlighted that the extent of subclonal diversification varies considerably among individual tumours with no strict temporal order being evident, and genomic alterations such as point mutations and rearrangements affect the most common breast cancer genes, including *PIK3CA*, *TP53*, *PTEN*, *BRCA2* and *MYC*, occurring early in some tumours and late in others<sup>21</sup>. This is in contrast with other tumour types that show a specific temporal order of mutations<sup>23–26</sup>. This may be due to the fact breast cancers inherently are thought primarily to be driven by an array of copy number driver alterations, with most somatic mutations occurring later in tumour development<sup>19</sup> or may reflect a situation wherein some tumours that grow to a certain size, clonal sweeps occur in which an especially fit subclone replaces the majority of others in the tumour<sup>21</sup>. As the different subtypes of breast cancer are thought to arise from different cells-of-origin, this may additionally explain the differences in heterogeneity observed in the different subtypes of breast cancer<sup>5</sup>.

<sup>1</sup>The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK. <sup>2</sup>Lester and Sue Smith Breast Center, Baylor College of Medicine, Houston, TX, USA. <sup>3</sup>Department of Medicine, Baylor College of Medicine, Houston, TX, USA. <sup>4</sup>Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX, USA. <sup>5</sup>Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, USA. ✉email: rachael.natrajan@icr.ac.uk



**Fig. 1** Pre-existing and acquired mechanisms of resistance to therapy in a model of clonal selection. **a** Tumours consist of genetically and phenotypically diverse clones (represented in different colours). Selective pressure such as therapy will eliminate a percentage of sensitive clones leaving behind some pre-existing resistant clones with particular genomic aberrations. These clones are able to survive and expand leading to drug resistance mediated through the selection of pre-existing clones (represented as gene A–D). New genomic aberrations may also arise due to the selective pressures of therapy leading to acquired drug resistance (represented as gene E). Remaining clones that are resistant to therapy may also be due to additional unknown resistance mechanisms and they may be pre-existing and/or acquired. **b** Pre-existing and/or acquired drug resistance may be due to transcriptional and epigenetic mechanisms affecting gene expression or the communication of tumour cells with the microenvironment. This figure was created with BioRender.com and is adapted in part from the template 'tumour microenvironment 2'.

### Tracking subclonal mutations in patients undergoing therapy

Subclonal changes can additionally occur upon the selective pressure of therapy (Fig. 1a)<sup>19,21,27–30</sup>. Drug resistance can be pre-existing (where therapy does not impact subclone frequency), selected (arising to the selection of rare pre-existing subclones that are able to expand), or acquired (where new genomic/transcriptomic/epigenetic aberrations are introduced contributing to the drug-resistant phenotype). Relapsed or metastatic breast cancers have been found to largely share the vast majority of their genomic alterations with the corresponding primary disease indicating pre-existing resistant clones, however many metastatic tumours also exhibit additional mutations that were not previously detected or are subclonal in the primary disease<sup>21,31–37</sup>. As an example, *ESR1* activating mutations are rarely present in primary ER+ breast cancer, even in those 15–20% patients that show intrinsic resistance to hormonal therapies<sup>22,38,39</sup>, whereas they are highly enriched in ER+ metastatic tumour samples from patients with acquired endocrine resistance (30–40% of ER+ patients) and in particular, are a known mechanism of resistance to aromatase inhibitors (AI)<sup>40–42</sup>. Recent large studies cataloguing the mutational repertoire of primary and relapsed metastatic breast cancers from patients have highlighted that multiple mechanisms of endocrine therapy resistance exist in ER+ BC cancers<sup>43</sup> including evidence that pre-existing low-frequency mutations can be responsible for endocrine resistance<sup>44,45</sup>.

Genetic mechanisms of resistance to chemotherapy are less well-documented in TNBC, and in general subclonal alterations are

not enriched upon chemotherapy treatment<sup>27,46</sup> highlighting that while a proportion of patient's tumours undergo clonal selection, others maintain a similar genomic architecture after chemotherapy<sup>46,47</sup>. It is worth mentioning that poly(ADP-ribose) polymerase (PARP1/2) inhibitor resistance is also observed in patients with homologous recombination defects including TNBC<sup>48–52</sup>. Although PARPi-based therapy is standard of care for patients with germline *BRCA1/2* mutations including breast and ovarian cancer, resistance has been observed in these patients, where reversion mutations of *BRCA1* or *BRCA2* are a well-documented resistance mechanism<sup>53,54</sup> and can be assessed through profiling of circulating free DNA (cfDNA)<sup>55</sup>. Although a plethora of potential resistance mechanisms to PARPi have been identified through high throughput functional genomic screens, many of these have yet to be validated in patients relapsing on PARPi treatment<sup>48–52,56,57</sup>. Overall, whether these reversion mutations are spontaneous, pre-existing and selected for under PARP1/2 inhibition, or induced by other therapeutic agents such as platinum previously administered to patients is still to be determined.

As well as bulk profiling of tumour biopsies, assessment of circulating tumour (ct) or cfDNA of cancer patients is a useful evolutionary methodology to track the evolution of intra-TH of cancer, the presence of treatment-resistant clones and to predict response and resistance. These tools have been shown to be useful in both metastatic patients<sup>58</sup> and in early breast cancer<sup>59</sup> to monitor response to therapy and are predictive of relapse<sup>42,60,61</sup>. Importantly, ctDNA analysis can provide important genomic

information that cannot be traced by a single biopsy at a specific cancer site and are becoming increasingly used in clinical studies as a non-invasive way of assessing the emergence of progressive disease and to direct therapy<sup>62–64</sup>. Although further research is necessary to determine whether ctDNA can fully represent tumour genomic profiles at any given time, recent data from a seminal clinical trial, plasmaMATCH, has assessed the validity and utility of ctDNA for targeted therapies without previous tissue testing in advance breast cancer patients<sup>65</sup>. Novel methylation profiling methods (so-called ‘molecular clocks’) can allow tracking of evolutionary changes of single CpG sites to track clonal haplotypes in ctDNA (‘epimutations’ that accumulate due to methyltransferase errors at cell division and occur 1000 to 10,000 times more frequently than DNA point mutations) at single-molecule resolution without a priori knowledge of the molecular make-up of the disease<sup>66</sup>. Implementation of such methodologies will likely overcome the limitations in terms of cost and prior knowledge of a patients’ tumour mutational repertoire for successful longitudinal patient tracking.

### Assessment of genetic intra-TH using single-cell approaches

Single-cell approaches to assess intra-TH have traditionally used fluorescence in situ hybridisation to measure copy number using probes against frequently amplified loci in breast cancer. Using diversity measures one study seeking to address whether copy number alterations drive neoadjuvant chemotherapy (NAC) resistance revealed that tumours that respond to NAC have a low diversity of copy number alterations before therapy, whereas those that had a poor response showed a higher diversity<sup>67</sup>. However, these patterns were unchanged between pre- and post-treatment biopsies, suggesting perhaps resistance to NAC is not driven by particular genomic alterations. Interestingly, this study did identify intra-TH based on the expression of cancer stem cell markers such as CD44 and CD24, that correlated with NAC response, suggesting phenotypic evolution plays a role in chemotherapy resistance<sup>67</sup>. Such studies have also highlighted that the spatial pattern of *PIK3CA* mutations together with *HER2* amplified cells play an important role in chemotherapy resistance<sup>68</sup>.

Recent advances in single-cell sequencing technologies now allow the assessment of intra-TH of tumours, as well as the non-genetic changes that may occur upon metastatic disease progression and upon selective pressures such as chemotherapy and targeted therapy in breast cancer. Single-cell sequencing overcomes the need for computational deconvolution of the theoretical contribution of individual clones but instead can profile the DNA, RNA, methylation and chromatin accessibility of each individual cell<sup>30,69–71</sup>.

Selective pressures such as targeted therapy can also alter the transcriptome of cancer cells, leading to drug resistance and patient relapse. In the first seminal single-cell sequencing study from Navin and colleagues, single-cell profiling was used to assess copy number heterogeneity in two breast cancer patients<sup>72</sup>. This study revealed that polygenomic tumours likely represent sequential clonal expansions and that seeding of metastasis from a monogenomic tumour and its liver metastasis indicated that a single clonal expansion formed the primary tumour and seeded the metastasis. Interestingly, the authors found that an unexpectedly abundant subpopulation of genetically diverse ‘pseudodiploid’ cells that did not travel to the metastatic site, suggesting that tumours grow by punctuated clonal expansions with few persistent intermediates in contrast to more gradual models of evolution<sup>72</sup>. A follow-up study looking at mutations at the single-cell level identified that copy number alterations occurred early in tumour evolution and remained highly stable suggesting clonal expansion, whereas, in contrast, point mutations evolved gradually, generating extensive clonal diversity, with many mutations

seen in less than 10% of the tumour cells<sup>30</sup>. Application of this technology to patients with ductal carcinoma in situ and concurrent invasive disease using topographical single-nucleus sequencing has revealed that the majority of genomic alterations developed in the ducts prior to the invasion, supporting a multiclonal invasion model, whereby more than one clone escapes to establish the invasive disease<sup>73</sup>. A similar study using single-cell profiling of formalin-fixed paraffin-embedded clinical tumour samples to assess copy number of the transition from ductal carcinoma in situ to invasive disease also found extensive intra-TH in pre-invasive lesions with evolution to invasive disease occurring via multiple mechanisms<sup>74</sup>. Together, these studies suggest genomic diversification of copy number alterations occur early during tumorigenesis due to clonal expansions and subclonal mutations occur later on, generating substantial mutational heterogeneity.

### Contribution of non-genetic forms of heterogeneity

Although the majority of the studies so far have focussed on assessing genetic alterations (mutations and copy number) to evaluate intra-TH, as discussed above, there is emerging evidence that non-genetic factors also play a major contributing role. Numerous studies have suggested the existence of differential transcriptomic make-up between primary tumour and the corresponding metastatic lesions<sup>75–78</sup>. Drug resistance can be through the selection of pre-existing alterations and/or acquired due to transcriptional and epigenetic mechanisms such as histone modifications and DNA methylation affecting gene expression or communication of tumour cells with the microenvironment (Fig. 1b)<sup>15,46,47,79–85</sup>. This has been addressed in a number of studies evaluating the single-cell transcriptomic make-up of the normal epithelium and of primary breast cancers.

Transcriptomic analyses from primary and metastatic cancers at both bulk and single-cell level has further highlighted the complexity of tumour evolution. Several reports have shown that metastatic breast cancer lesions show a higher transcriptomic heterogeneity than the corresponding primary disease<sup>75–78,86</sup>. Single-cell sequencing of circulating tumour cells (CTCs) of patients with metastatic breast cancer versus patients with primary breast cancer, highlighted increased transcriptomic heterogeneity in those patients<sup>76</sup>. Of note transcriptional signatures associated with epithelial to mesenchymal transition were enriched in CTCs derived from patients with metastatic disease<sup>76</sup>. Importantly, a single-cell study of patient-derived TNBC primary tumours in vivo and their corresponding lung and lymph-node micro-metastases exhibited a distinct transcriptome at a single-cell level. Interestingly, the mitochondrial oxidative phosphorylation (OXPHOS) pathway was upregulated in micro-metastases<sup>75</sup>. Notably, pharmacological inhibition of OXPHOS attenuated lung metastasis in vivo, highlighting the importance of this pathway in a metastatic spread as well as its potential in metastatic prevention in breast cancer patients<sup>75</sup>. These findings are in line with bulk RNA sequencing of primary tumours and their matched metastases of TNBC patients, highlighting consistent gene expression differences including enrichment of hypoxia, cellular metabolism, and fatty acid  $\beta$  oxidation pathways in the metastases, and were independent of the metastatic site<sup>86</sup>. Collectively, the above studies indicate the importance of transcriptomic heterogeneity in metastatic breast cancer as well as its potential role in the aggressiveness of this disease.

In a seminal study from Navin and colleagues, single-cell DNA and RNA sequencing was used to analyse the contribution of genetic and non-genetic heterogeneity in TNBC patients receiving NAC and treatment with the anti-VEGF therapy (bevacizumab)<sup>47</sup>. Although this study only looked longitudinally in a small number of tumours, some patients harboured enrichment of genomic alterations in residual disease post-NAC, however, overall, there

was no evidence that specific genetic differences contributed to chemotherapy resistance. Single-cell RNA sequencing of a subset of these tumour samples, however, highlighted that the transcriptomic profiles of pre-treatment and post-treatment samples were distinct, suggestive of a model of therapy resistance in which selective and acquired modes of evolution are responsible for the resistant population, in agreement with previous *in situ* based studies<sup>67</sup>.

A recent study performing single-nucleus RNA sequencing of six therapy naive primary TNBC fresh tumours highlighted that in the absence of chemotherapy therapeutic pressure, TNBCs show extensive intra-TH at the transcriptomic level<sup>87</sup>. This appeared to be driven by copy number alterations within the cells, suggestive that at least in the primary disease setting, the genotype can influence transcriptomic heterogeneity of individual subpopulations. Perhaps most interesting from this study was the identification of distinct subclonal populations of malignant cells that were shared between the different patients characterised by multiple signatures of treatment resistance and metastasis, including glycosphingolipid metabolism and associated innate immunity pathways. Moreover, these signatures identified from the subclonal cellular populations were able to predict long-term prognosis in independent cohorts of primary TNBC from bulk RNA-profiling. This suggests that minority populations of cells have already acquired the transcriptomic features to allow them to become metastatic. Moreover, if this signal can be elucidated from bulk tissue once identified, this heralds a paradigm shift for identification and specific targeting of these subpopulations a priori to circumvent metastatic spread.

A number of recent studies have attempted to address whether drug resistance including endocrine therapy or chemotherapy in breast cancer is selected for and/or acquired. Upon endocrine therapy in ER + breast cancer cell line, MCF7, a rare subpopulation of pre-adapted cells with high plasticity have been found to have distinct transcriptomic signatures (upregulation of genes involved in p53 signalling pathway, epithelial-mesenchymal transition (EMT), hypoxia and cell polarity) with features of dormancy as well as mixed epithelial and mesenchymal characteristics. Interestingly, single-cell sequencing highlighted that these pre-adapted cells that display features of dormancy and mixed epithelial and mesenchymal traits are able to survive upon short-term endocrine therapy, however, these pathways are not seen in fully resistant cells, suggesting that the cells undergo further transcriptomic reprogramming to become fully resistant to endocrine therapy and to generate subsequent metastasis<sup>79</sup>. Additionally, single-cell sequencing of ER + cell lines has identified genetically and transcriptionally distinct subpopulations including rare populations of MCF7 cells highly expressing an apoptosis-related signature, positively correlated with a pre-adaptation signature to oestrogen deprivation<sup>88</sup>. These results suggest a multi-step mechanism of drug resistance to endocrine therapy where both genetic and nongenetic alterations contribute to this phenotype. In a similar study, authors used molecular barcoding coupled with single-cell transcriptomic sequencing to track the ER + cell line, MCF7, upon long-term endocrine therapy *in vitro*. Their findings suggested that the majority of fulvestrant or tamoxifen resistance clones were pre-existing and highly selected during treatment<sup>80</sup>.

Single-cell RNA profiling of CTCs from liquid biopsies have further revealed the extent of intra-TH, with rare CTCs displaying divergent gene expression signatures of EMT, cancer stemness, DNA repair, kinase signalling and immune-tumour cross-talk<sup>89</sup>. Moreover, single-cell profiling of CTC clusters identified that these clusters are oligoclonal, and although rare have up to a 50-fold increase in metastatic potential and are characterised by high expression of the cell junction component plakoglobin, suggesting this could be a useful biomarker to detect cells with high metastatic potential in the primary tumour<sup>90</sup>. Profiling of CTCs

from metastatic breast cancer patients has also highlighted the intra-TH of *ESR1* mutations in patients with endocrine resistance suggesting independent mechanisms of resistance to these agents, and highlighting the clinical applicability of single-cell approaches for detecting genomic alterations mediating targeted therapy resistance<sup>91</sup>.

### The role of epigenetic heterogeneity in therapy resistance

It is now well established that epigenetic mechanisms play a vital role in driving phenotypic heterogeneity and evolution and contribute to treatment resistance in breast cancer patients. In endocrine resistance in ER + MCF7 breast cancer cells, for example, drug-induced epigenetic states have been shown to involve large topologically associating domains and the activation of super-enhancers. In particular, MCF7 cells resistant to AI have been shown to upregulate cholesterol biosynthesis through stable epigenetic activation, which leads to constitutive ER-activation in AI-resistant cells<sup>82</sup>. More recently, enhancer mapping of H3K27 acetylation states in 47 matched primary and metastatic endocrine-resistant tumour biopsies has allowed the inference of tumour evolution and highlighted that endocrine therapies select for phenotypic clones under-represented at diagnosis<sup>92</sup>. These studies suggest phenotypic heterogeneity governed by inheritable cell-type-specific transcription is maintained through the activation of epigenetically defined regulatory regions including promoters and enhancers<sup>92</sup>.

Evidence also suggests epigenetic intra-TH governs resistance to novel therapeutic agents such as Bromodomain and Extra-Terminal motif inhibitors<sup>93</sup>. Indeed, alterations in epigenetic mechanisms through the induction of several transcriptional states, contribute to intra-TH and thus to the aggressiveness of the disease and drug resistance<sup>94</sup>. In a recent pivotal study by Hinohara and colleagues, the epigenetic enzyme, KDM5B, was found to be a regulator of transcriptomic heterogeneity in a panel of ER + breast cancer cell lines, promoting resistance to endocrine therapy. Its high expression in ER + breast cancer cell lines was found to be associated with higher transcriptomic heterogeneity<sup>80</sup> and was reversed upon treatment with KDM5 small molecule inhibitors, providing proof of concept that modulating the activity of epigenetic enzymes, can lead to improved responses to the standard of care treatment. Interestingly in cellular models, resistance to KDM5 inhibitors was acquired rather than selected for as in endocrine therapy resistance. Additional studies have also identified epigenetic reprogramming of the transcriptional repressive mark, H3K27me3, in cell populations resistant to chemotherapy in TNBC and to tamoxifen in ER + breast cancer patient-derived xenografts (PDXs)<sup>83</sup>.

### Single-cell approaches to dissect the contribution of the tumour microenvironment

Increasing evidence suggests that epigenetic and subsequent transcriptomic alterations in cells within the tumour microenvironment (e.g. stromal, immune cells and non-cellular components that make up the extra-cellular matrix) contribute to tumour initiation, progression and metastasis mediated through paracrine signalling with the tumour cells<sup>95,96</sup>. Gene expression signatures derived from the tumour stroma can be predictive of patients' clinical outcome<sup>97</sup>. Thus, there is increasing interest in addressing the contribution of the microenvironment in terms of tumour development, progression, drug resistance as well as its potential as a prognostic factor and a therapeutic target. The use of single-cell RNA sequencing has enabled deeper insights into intra-TH as well as the contribution of tumour microenvironment which are not detectable by other methods<sup>98-100</sup>. Recently, the immune atlas in breast cancer has been mapped. Using unbiased single-cell RNA sequencing analysis, the immune cells from eight primary breast cancers patients, as well as their matched normal breast

tissue, blood and lymph nodes were profiled in order to investigate the role of immune cells in tumour microenvironment and its contribution to cancer progression and immunotherapy response<sup>98</sup>. The authors identified significantly increased phenotypic expansion of intra-tumoral immune cells in comparison to normal breast tissue. Similarly, the tumour microenvironment of primary TNBC tumours has been profiled by high throughput single-cell RNA sequencing<sup>100</sup>. The authors identified two cancer-associated fibroblast and two perivascular-like subpopulations, where each population clustered into different states. These signatures were found to be strongly associated with cytotoxic T-cell dysfunction, leading to immune evasion<sup>100</sup>.

It is becoming clear that the spatial organisation of tumour microenvironments plays an important role in therapy response and patient outcome in breast cancer<sup>101,102</sup>. Tissue micro-dissection paired with gene expression analysis has provided valuable insights into tumour–microenvironment relationships. For example, it was found that in TNBC T-cell infiltration within a tumour mass, but not peripheral localisation of T cells, was associated with good patient outcome<sup>103</sup>. Digital pathology coupled with computational assessment allows a unique insight into how the tumours are organised and is an exciting emerging field. For example, the spatial heterogeneity of immune cells, stromal cells and epithelial cancer cells assessed by haematoxylin and eosin routine histopathology sections has been shown to be associated with a poor response to standard therapies in both ER+ and ER- primary breast cancers, suggesting that having a mixture of these cells within the tumour microenvironment allows adaptation as the cancer evolves<sup>102</sup>. Moreover, increased immune cell spatial clustering has been shown to predict both early (0–5 years) and late (5–10 years) recurrence in ER+ breast cancer patients after endocrine therapy, whereas total immune cell abundance showed no such an association<sup>101</sup>. A recent study of 20 primary breast cancers subjected to single-cell RNA coupled with epitope sequencing (CITE-seq), has revealed new PD-L1/PD-L2+ macrophages associated with clinical outcome<sup>100</sup>. Using spatial transcriptomic sequencing on the corresponding tissue sections of these tumours further revealed that stroma-immune niches were spatially organised. Moreover, single-cell signatures were able to stratify large breast cancer cohorts at the bulk RNA level into nine “ecotype” clusters associated with unique clinical outcomes which were distinct from the classical intrinsic molecular subtypes of breast cancer<sup>100</sup>. Recent technological advances in proteomics have additionally allowed assessment of the epithelial and stromal cell make up of primary breast cancers whilst maintaining the spatial architecture. Using imaging cytometry time of flight (CyTOF) to measure multiple proteins in a single experiment now allows unprecedented assessment of the tumour microenvironment. These studies have revealed the extensive tumour and immune ecosystems in primary breast cancers<sup>104,105</sup> and highlighted associations between genomic alterations such as *TP53* mutations and distinct epithelial and stromal phenotypes, that hypoxic environments may induce immune tolerance, and identified that single-cell pathology subgroups have distinct clinical outcomes<sup>104,105</sup>. Further assessment of these during tumour evolution will be important going forward to validate these as biomarkers of disease progression.

### Potential clinical impacts of intra-TH in breast cancer

Whilst genetic intra-TH has been documented to be associated with poor patient outcomes in tumour types such as leukaemia and lung cancer<sup>106,107</sup>, its prognostic association in breast cancer is less clear. Multi-regional sequencing analysis of 75 TNBC primary tumours suggested genomic intra-TH of copy number status in several driver oncogenes is correlated with metastasis<sup>108</sup>. In the context of evolution from primary to metastatic disease, high depth bulk profiling of primary and matched metastatic tumours

from the same patient have revealed that most distant metastases acquire genomic driver mutations not seen in the primary tumour, drawing from a wider repertoire of cancer genes than early drivers, including a number of clinically actionable mutations inactivating SWI-SNF and JAK2-STAT3 pathways<sup>36</sup>. In the context of treatment response, analyses of relatively small cohorts using multi-regional profiling have revealed contrasting results associated with therapy responses, with one study suggesting correlation<sup>67</sup> and another study finding lack thereof<sup>21</sup> between the extent of genomic intra-TH with chemotherapy response. In a study assessing the mutational repertoire in primary HER2 amplified breast cancers with heterogeneous HER2 protein expression, the clinical impact of intra-TH is perhaps clearer. Heterogeneous HER2 amplification is seen in around 5% of HER2 amplified breast cancers. Through bulk dissection and genomic profiling of HER2-amplified and non-HER2 amplified components of clinically classified HER2-positive breast cancers, a recent study identified that the HER2-negative component of these tumours harboured additional driver mutations, including one case with an activating HER2 mutation, highlighting convergent evolution in this patient. Overall, however, the HER2-negative component would not be predicted to respond to anti-HER2 therapies, this highlights the need for close monitoring of these patients undergoing HER2 targeted therapeutic interventions<sup>109</sup>. Further assessment of the role of phenotypic heterogeneity driving the prognosis of breast cancers from single-cell epigenomic and transcriptomic studies is starting to reveal that transcriptomic features are seen in minority subpopulations and novel immune subpopulations in primary disease correlates with patient outcome<sup>87,100</sup>. There is a major need in the field to conduct large-scale survival analysis with genome-wide indicators of intra-TH to thoroughly address the question of whether primary tumour intra-TH is truly associated with patient outcomes in breast cancer.

### Modelling intra-TH and subclonal evolution

The observed intra-TH over the past years in breast cancer has raised the need for the generation and use of relevant preclinical models mirroring the full spectrum of the disease. Although established cancer cell lines have served as a fundamental in vitro model in breast cancer research over the past decades, they do not fully recapitulate all aspects of breast cancer including the natural structure of tumours, altered cell morphology, microenvironment, molecular and genomic resemblance to breast cancer tumours<sup>110,111</sup>. The generation and use of PDXs, PDX derived organoids (PDXOs)<sup>110–114</sup>, and patient-derived organoids (PDOs)<sup>115</sup> from surgical or biopsy patients' samples have proven to be a very powerful, preclinical in vitro and in vivo tool to assess drug response and resistance, modelling metastasis, partially elucidate tumour microenvironment and for biomarker development as they more accurately represent the patients' tumours. Importantly, PDXs<sup>33,78,116–118</sup> and PDXOs<sup>115</sup> have both been shown to preserve much of the intra-TH found in originating patients' biopsies, making them attractive models with which to investigate intra-TH dynamics in the laboratory. PDXOs and PDOs are currently attractive models with which to readily probe the activities of hundreds of experimental compounds in patient-derived breast tumour cells<sup>115,119</sup>. The primary limitation of PDX models is the lack of a fully intact immune system. Thus, future studies implementing pharmacologic and genetic screening technologies in models with 'humanised' immune systems will be of great value to the field.

Many scientists now use PDXs and PDOs as preclinical models to address intra-TH and evolution in cancer (see ref. <sup>120</sup> for an extensive review). One methodology in achieving this is lineage tracing by using highly complex barcode libraries, enabling the tracking of individual cells or subclones during both the metastatic cascade and through therapy resistance, using

next-generation sequencing as a readout<sup>121</sup>. These technologies label up to millions of cells with unique molecular barcode identifiers that allow transmission during cell division and hence single-cell tracing over time. This approach allows the investigation of the presence of pre-existing resistant clones in the tumour mass, molecular changes that can occur due to therapy and resistance, or therapeutic approaches to be taken into consideration to successfully target these pre-existing resistant clones through single-cell DNA or RNA sequencing<sup>46,79,88,94,122</sup>. This method in particular has been used to study clonal dynamics and therapy response in primary TNBC<sup>46,78,123</sup>. In these elegant studies, patient-derived models of primary TNBC were used and cells were barcoded with a complex lentiviral library. Residual tumours after chemotherapy exposure were found to maintain their barcode complexity rather than show a pattern of enrichment, suggestive of cellular plasticity, rather than pre-existing genomic alterations that emerged. OXPHOS was one of the most significantly upregulated pathways in residual tumours and treatment with an inhibitor of the mitochondrial electron transport chain complex I delayed the regrowth of chemotherapy-treated residual tumours in vivo. Interestingly, residual PDX tumours had differentially activated epigenetic regulators including HDAC7, HDAC10 and SIRT7, suggesting that these epigenetic regulators can act as promising candidates for TNBC tumours with residual disease in combination with metabolic inhibitors<sup>46</sup>. In vivo barcoding in primary untreated TNBC PDX models has also allowed investigation of the subclonal evolution of metastatic spread. These studies harness the utility of the fact that many TNBC PDX's spontaneously metastasise to distant organs in the absence of selective pressure such as chemotherapy. Using the same barcoding approach, Echeverria and colleagues showed that metastatic spread occurs through a selective bottleneck revealing an enrichment of shared clonal lineages across multiple metastatic sites, indicative of polyclonal seeding. Interestingly, rather than only a few clones leading to seeding at distant sites, the authors also observed extremely rare populations of cells that were able to seed and were maintained at low levels at these metastatic sites<sup>78</sup>. Similarly, studies using barcoded syngeneic 4T1 mammary tumour derived cells have demonstrated a bottleneck in metastatic seeding from the primary tumour in CTC's analysed in the blood and cells at distant sites. Interestingly in this model, these lineages were largely non-overlapping suggestive of independent subclones colonising distant sites and showed tropism for entering the lymphatic or vasculature system. Vascular tropism in particular was correlated with increased expression of *Serpine2* and *Slpi*, which were necessary and sufficient to programme these cells for vascular mimicry<sup>124</sup>. In addition, molecular barcoding has been used coupled with CRISPR technologies to trace a cell phenotype immediately after the introduction of DNA mutations in cancer<sup>125,126</sup>.

Given the polyclonal nature of metastatic spread and the fact CTC clusters have been shown to have a substantial increase in metastatic seeding<sup>90</sup>, this highlights the possibility that subclones interact with one another to mediate therapy resistance and spread. This is in where one subclone may maintain the fitness advantage of others through distinct paracrine contributions by maintaining the fitness of the whole or subpopulations resulting in a clonally diverse tumour phenotype. This concept has been evaluated using a *Trp53*-null mouse model of basal-like breast cancer, where a unique subpopulation of tumour cells expressing mesenchymal-like markers was identified to cross-talk with tumour initiating cells promoting tumorigenicity<sup>127</sup>. Cleary et al. show that, in some *Wnt1* driven mammary tumours, neither luminal nor basal cancer cells alone could form a new tumour when injected into the mammary fat pads of immunocompromised mice, whereas mixtures of the two cell populations induced biclonal tumour formation with high efficiency, suggestive of

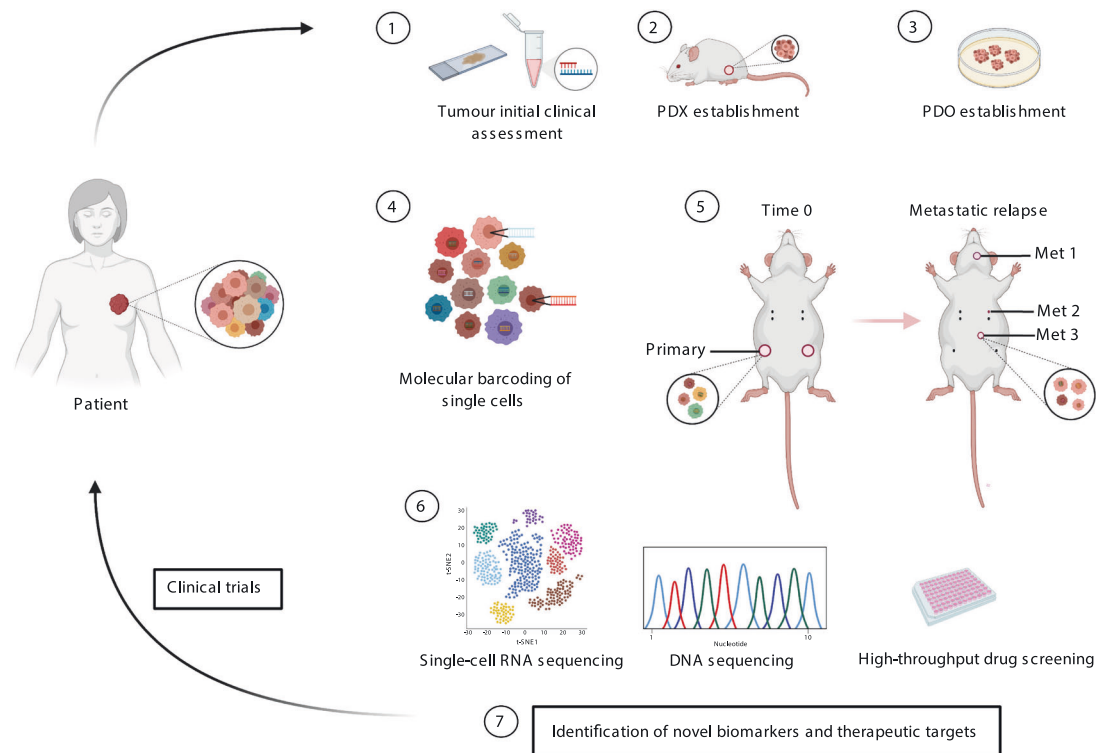
cooperativity between the two cellular populations<sup>128</sup>. More recently, in vivo assessment of isolated secreted factor-expressing subclones of the TNBC cell line, MDA-MB-468, revealed non-cell-autonomous maintenance of intra-TH<sup>129</sup> in which minor subclones expressing IL11 and FIGF (VEGFD) cooperate to promote metastatic progression and generate polyclonal metastases composed of a driver and neutral subclones. This cooperation was indirect and mediated through a specific neutrophil population that was stimulated by the IL11-expressing minor subclone<sup>130</sup>. Together, these studies demonstrate that non-cell-autonomous cooperative interactions exist between different cellular populations within tumours and highlight potential intervention strategies to circumvent these. In the context of NAC resistance in TNBC in particular, subclonal cooperativity may explain the reason for the lack of recurrent molecular alterations identified in NAC-resistant tumours.

Taken together, these studies highlight the polyclonal and complex nature of metastatic spread in breast cancer. Further assessment of these in humanised patient-derived models to encompass competent immune systems will be necessary to further elucidate mechanisms of metastatic spread and resistance to particular therapies. One can envision a strategy whereby this could be intersected with an extensive characterisation of clinical tumour sampling at single-cell resolution in matched models to fully assess the intra-TH at a genetic and genomic level in an individual patient. Only then, we will be in a better place to tackle the complexity of intra-TH and evolution in cancer (Fig. 2).

## CONCLUSION

Intra-TH remains a challenge in terms of appropriate cancer therapeutic strategies. Current biomarker analysis does not fully represent the complexity of TH; the existence of rare subclones or accurately predicting treatment outcomes of cancer patients. Deconvoluting the genomic, epigenomic and functional heterogeneity in breast cancer can improve our understanding of the drivers of the disease, mechanisms of resistance and metastasis, identify more accurate biomarkers and novel therapeutic strategies. The use of advanced methodologies such as molecular barcoding, single-cell RNA profiling together with mathematic modelling have proven to be of high importance in terms of addressing these issues. However, further efforts are necessary in order to comprehensively determine and control intra-TH in breast cancer. Furthermore, the development of novel cfDNA molecular tracking such as the use of 'molecular epigenetic clocks', can provide unbiased tools for longitudinal patient monitoring in the clinic.

The continuous transcriptional and epigenetic contribution during cancer progression or upon selective pressures have implications on the gene expression of tumour cells, making the tumour more heterogeneous and thus difficult to treat, leading to drug resistance and metastasis. Although the development of advanced methodologies such as single-cell RNA sequencing and single-cell tracing molecular barcoding methodologies in relevant preclinical models have significantly improved our understanding of TH and evolution, further studies in immune-competent models are necessary in order to map the genomic and epigenetic landscape of these rare drug-resistant clones. Moreover, additional studies are needed to fully appreciate the potential inter-clonal cooperativity mediated by both non-tumour cells and between different tumour subclones. This will then allow insights into the identification of novel drugs that can disrupt these interactions and shift these clones from a drug-resistant state to a drug-responsive state. Additionally, the transcriptomic profiles from the studies described above can be applied as tools to predict which patients will respond to specific therapies, provide prognostic signatures for patients' outcomes besides the histopathological techniques already in use, as well as promising candidates for new



**Fig. 2 Models tackling intra-TH in breast cancer.** Once the primary tumour is removed from a patient, the tumour sample is used for (1) initial clinical assessment and characterisation, followed by (2) PDX (patient material grown in mice) and (3) PDO (patient material directly grown ex vivo) establishment for research purposes and compared to the primary tumour sample. The remaining tumour is (4) dissociated to single cells, barcoded and (5) injected into the mammary glands of mice. Primary and matched metastatic tumours are removed from mice and subjected to (6) single-cell RNA sequencing, DNA sequencing and high throughput drug screenings thus (7) identifying the genetic, non-genetic make-up and evolution of these tumours at a single-cell level as well as identifying novel therapeutic strategies and biomarkers. This is then linked back into the analysis of patient samples in clinical trials. This figure was created with BioRender.com.

therapeutic strategies in breast cancer. These could be used as novel combinatorial approaches together or using intermittent scheduling approaches with existing therapies including immunomodulatory agents targeting the tumour microenvironment.

Received: 8 March 2021; Accepted: 26 November 2021;

Published online: 21 December 2021

## REFERENCES

- Ellsworth, R. E., Blackburn, H. L., Shriver, C. D., Soon-Shiong, P. & Ellsworth, D. L. Molecular heterogeneity in breast cancer: State of the science and implications for patient care. *Semin Cell Dev. Biol.* **64**, 65–72 (2017).
- Koren, S. & Bentires-Alj, M. Breast tumor heterogeneity: source of fitness, hurdle for therapy. *Mol. Cell* **60**, 537–546 (2015).
- Turashvili, G. & Brogi, E. Tumor heterogeneity in breast cancer. *Front Med.* **4**, 227 (2017).
- Harbeck, N. et al. Breast cancer. *Nat. Rev. Dis. Prim.* **5**, 66 (2019).
- Perou, C. M. et al. Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
- Sorlie, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl Acad. Sci. USA* **98**, 10869–10874 (2001).
- Prat, A., Ellis, M. J. & Perou, C. M. Practical implications of gene-expression-based assays for breast oncologists. *Nat. Rev. Clin. Oncol.* **9**, 48–57 (2011).
- Prat, A. et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast* **24**(Suppl 2), S26–S35 (2015).
- Lehmann, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
- Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976).
- Marusyk, A. & Polyak, K. Tumor heterogeneity: causes and consequences. *Biochim Biophys. Acta* **1805**, 105–117 (2010).
- Davis, A., Gao, R. & Navin, N. Tumor evolution: linear, branching, neutral or punctuated? *Biochim. Biophys. Acta Rev. Cancer* **1867**, 151–161 (2017).
- Battle, E. & Clevers, H. Cancer stem cells revisited. *Nat. Med.* **23**, 1124–1134 (2017).
- Banerji, S. et al. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**, 405–409 (2012).
- Ciriello, G. et al. Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* **163**, 506–519 (2015).
- Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
- Nik-Zainal, S. et al. The life history of 21 breast cancers. *Cell* **149**, 994–1007 (2012).
- Stephens, P. J. et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature* **486**, 400–404 (2012).
- Shah, S. P. et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395–399 (2012).
- Navin, N. et al. Inferring tumor progression from genomic heterogeneity. *Genome Res.* **20**, 68–80 (2010).
- Yates, L. R. et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat. Med.* **21**, 751–759 (2015).
- Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
- Green, M. R. et al. Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood* **121**, 1604–1611 (2013).
- Moskaluk, C. A., Hruban, R. H. & Kern, S. E. p16 and K-ras gene mutations in the intraductal precursors of human pancreatic adenocarcinoma. *Cancer Res* **57**, 2140–2143 (1997).
- Powell, S. M. et al. APC mutations occur early during colorectal tumorigenesis. *Nature* **359**, 235–237 (1992).
- Yachida, S. & Iacobuzio-Donahue, C. A. Evolution and dynamics of pancreatic cancer progression. *Oncogene* **32**, 5253–5260 (2013).

27. Balko, J. M. et al. Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. *Cancer Discov.* **4**, 232–245 (2014).
28. Gellert, P. et al. Impact of mutational profiles on response of primary oestrogen receptor-positive breast cancers to oestrogen deprivation. *Nat. Commun.* **7**, 13294 (2016).
29. Miller, C. A. et al. Aromatase inhibition remodels the clonal architecture of estrogen-receptor-positive breast cancers. *Nat. Commun.* **7**, 12498 (2016).
30. Wang, Y. et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
31. Brastianos, P. K. et al. Genomic characterization of brain metastases reveals branched evolution and potential therapeutic targets. *Cancer Discov.* **5**, 1164–1177 (2015).
32. De Mattos-Arruda, L. et al. Capturing intra-tumor genetic heterogeneity by de novo mutation profiling of circulating cell-free tumor DNA: a proof-of-principle. *Ann. Oncol.* **25**, 1729–1735 (2014).
33. Ding, L. et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999–1005 (2010).
34. Hoadley, K. A. et al. Tumor evolution in two patients with basal-like breast cancer: a retrospective genomics study of multiple metastases. *PLoS Med.* **13**, e1002174 (2016).
35. Savas, P. et al. The subclonal architecture of metastatic breast cancer: results from a prospective community-based rapid autopsy program “CASCADE”. *PLoS Med.* **13**, e1002204 (2016).
36. Yates, L. R. et al. Genomic evolution of breast cancer metastasis and relapse. *Cancer Cell* **32**, 169–184.e167 (2017).
37. Shah, S. P. et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809–813 (2009).
38. Jeselsohn, R. et al. Emergence of constitutively active estrogen receptor- $\alpha$  mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin. Cancer Res.* **20**, 1757–1767 (2014).
39. Lei, J. T., Anurag, M., Haricharan, S., Gou, X. & Ellis, M. J. Endocrine therapy resistance: new insights. *Breast* **48**, S26–S30 (2019).
40. Toy, W. et al. ESR1 ligand-binding domain mutations in hormone-resistant breast cancer. *Nat. Genet.* **45**, 1439–1445 (2013).
41. Robinson, D. R. et al. Activating ESR1 mutations in hormone-resistant metastatic breast cancer. *Nat. Genet.* **45**, 1446–1451 (2013).
42. Schiavon, G. et al. Analysis of ESR1 mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. *Sci. Transl. Med.* **7**, 313ra182 (2015).
43. Razavi, P. et al. The genomic landscape of endocrine-resistant advanced breast cancers. *Cancer Cell* **34**, 427–438.e426 (2018).
44. Nagarajan, S. et al. ARID1A influences HDAC1/BRD4 activity, intrinsic proliferative capacity and breast cancer treatment response. *Nat. Genet.* **52**, 187–197 (2020).
45. Xu, G. et al. ARID1A determines luminal identity and therapeutic response in estrogen-receptor-positive breast cancer. *Nat. Genet.* **52**, 198–207 (2020).
46. Echeverria, G. V. et al. Resistance to neoadjuvant chemotherapy in triple-negative breast cancer mediated by a reversible drug-tolerant state. *Sci. Transl. Med.* **11**, eaav0936 (2019).
47. Kim, C. et al. Chemoresistance evolution in triple-negative breast cancer delineated by single-cell sequencing. *Cell* **173**, 879–893.e813 (2018).
48. Henneman, L. et al. Selective resistance to the PARP inhibitor olaparib in a mouse model for BRCA1-deficient metaplastic breast cancer. *Proc. Natl Acad. Sci. USA* **112**, 8409–8414 (2015).
49. Jaspers, J. E. et al. Loss of 53BP1 causes PARP inhibitor resistance in Brca1-mutated mouse mammary tumors. *Cancer Discov.* **3**, 68–81 (2013).
50. Liu, Y. et al. RAD51 mediates resistance of cancer stem cells to PARP inhibition in triple-negative breast cancer. *Clin. Cancer Res.* **23**, 514–522 (2017).
51. Pettitt, S. J. et al. Genome-wide and high-density CRISPR-Cas9 screens identify point mutations in PARP1 causing PARP inhibitor resistance. *Nat. Commun.* **9**, 1849 (2018).
52. Schoonen, P. M. et al. Progression through mitosis promotes PARP inhibitor-induced cytotoxicity in homologous recombination-deficient cancer cells. *Nat. Commun.* **8**, 15981 (2017).
53. Edwards, S. L. et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature* **451**, 1111–1115 (2008).
54. Sakai, W. et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature* **451**, 1116–1120 (2008).
55. Khaliq, S. et al. Longitudinal analysis of a secondary BRCA2 mutation using digital droplet PCR. *J. Pathol. Clin. Res.* **6**, 3–11 (2020).
56. Pettitt, S. J. & Lord, C. J. Dissecting PARP inhibitor resistance with functional genomics. *Curr. Opin. Genet. Dev.* **54**, 55–63 (2019).
57. Li, H. et al. PARP inhibitor resistance: the underlying mechanisms and clinical implications. *Mol. Cancer* **19**, 107 (2020).
58. Dawson, S. J. et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N. Engl. J. Med.* **368**, 1199–1209 (2013).
59. Garcia-Murillas, I. et al. Assessment of molecular relapse detection in early-stage breast cancer. *JAMA Oncol.* **5**, 1473–1478 (2019).
60. Alix-Panabieres, C. & Pantel, K. Circulating tumor cells: liquid biopsy of cancer. *Clin. Chem.* **59**, 110–118 (2013).
61. Gasch, C., Pantel, K. & Riethdorf, S. Whole genome amplification in genomic analysis of single circulating tumor cells. *Methods Mol. Biol.* **1347**, 221–232 (2015).
62. Pearson, A. et al. Inactivating NF1 mutations are enriched in advanced breast cancer and contribute to endocrine therapy resistance. *Clin. Cancer Res.* **26**, 608–622 (2020).
63. O’Leary, B. et al. The genetic landscape and clonal evolution of breast cancer resistance to palbociclib plus fulvestrant in the PALOMA-3 trial. *Cancer Discov.* **8**, 1390–1403 (2018).
64. Razavi, P. et al. Alterations in PTEN and ESR1 promote clinical resistance to alpelisib plus aromatase inhibitors. *Nat. Cancer* **1**, 382–393 (2020).
65. Turner, N. C. et al. Circulating tumour DNA analysis to direct therapy in advanced breast cancer (plasmaMATCH): a multicentre, multicohort, phase 2a, platform trial. *Lancet Oncol.* **21**, 1296–1308 (2020).
66. Cresswell, G. D. et al. Mapping the breast cancer metastatic cascade onto ctDNA using genetic and epigenetic clonal tracking. *Nat. Commun.* **11**, 1446 (2020).
67. Almendro, V. et al. Inference of tumor evolution during chemotherapy by computational modeling and in situ analysis of genetic and phenotypic cellular diversity. *Cell Rep.* **6**, 514–527 (2014).
68. Janiszewska, M. et al. In situ single-cell analysis identifies heterogeneity for PIK3CA mutation and HER2 amplification in HER2-positive breast cancer. *Nat. Genet.* **47**, 1212–1219 (2015).
69. Hedlund, E. & Deng, Q. Single-cell RNA sequencing: technical advancements and biological applications. *Mol. Asp. Med.* **59**, 36–46 (2018).
70. Clark, S. J. et al. Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat. Protoc.* **12**, 534–547 (2017).
71. Cusanovich, D. A. et al. Multiplex single cell profiling of chromatin accessibility by combinatorial cellular indexing. *Science* **348**, 910–914 (2015).
72. Navin, N. et al. Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
73. Casasent, A. K. et al. Multiclonal invasion in breast tumors identified by topographic single cell sequencing. *Cell* **172**, 205–217.e212 (2018).
74. Martelotto, L. G. et al. Whole-genome single-cell copy number profiling from formalin-fixed paraffin-embedded samples. *Nat. Med.* **23**, 376–385 (2017).
75. Davis, R. T. et al. Transcriptional diversity and bioenergetic shift in human breast cancer metastasis revealed by single-cell RNA sequencing. *Nat. Cell Biol.* **22**, 310–320 (2020).
76. Powell, A. A. et al. Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. *PLoS ONE* **7**, e33788 (2012).
77. Nguyen, A., Yoshida, M., Goodarzi, H. & Tavazoie, S. F. Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. *Nat. Commun.* **7**, 11246 (2016).
78. Echeverria, G. V. et al. High-resolution clonal mapping of multi-organ metastasis in triple negative breast cancer. *Nat. Commun.* **9**, 5079 (2018).
79. Hong, S. P. et al. Single-cell transcriptomics reveals multi-step adaptations to endocrine therapy. *Nat. Commun.* **10**, 3840 (2019).
80. Hinojara, K. et al. KDM5 histone demethylase activity links cellular transcriptomic heterogeneity to therapeutic resistance. *Cancer Cell* **35**, 330–332 (2019).
81. Magnani, L. et al. Genome-wide reprogramming of the chromatin landscape underlies endocrine therapy resistance in breast cancer. *Proc. Natl Acad. Sci. USA* **110**, E1490–E1499 (2013).
82. Nguyen, V. T. et al. Differential epigenetic reprogramming in response to specific endocrine therapies promotes cholesterol biosynthesis and cellular invasion. *Nat. Commun.* **6**, 10044 (2015).
83. Gosselin, K. et al. High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat. Genet.* **51**, 1060–1066 (2019).
84. Caiado, F., Silva-Santos, B. & Norell, H. Intra-tumour heterogeneity - going beyond genetics. *FEBS J.* **283**, 2245–2258 (2016).
85. Huntsman, D. G. & Ladanyi, M. The molecular pathology of cancer: from pan-genomics to post-genomics. *J. Pathol.* **244**, 509–511 (2018).
86. Siegel, M. B. et al. Integrated RNA and DNA sequencing reveals early drivers of metastatic breast cancer. *J. Clin. Invest.* **128**, 1371–1383 (2018).
87. Karaayvaz, M. et al. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. *Nat. Commun.* **9**, 3588 (2018).
88. Chen, F. et al. Single-cell transcriptomic heterogeneity in invasive ductal and lobular breast cancer cells. *Cancer Res.* **81**, 268–281 (2021).



89. Gorges, T. M. et al. Accession of tumor heterogeneity by multiplex transcriptome profiling of single circulating tumor cells. *Clin. Chem.* **62**, 1504–1515 (2016).
90. Aceto, N. et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110–1122 (2014).
91. Paoletti, C. et al. Comprehensive mutation and copy number profiling in archived circulating breast cancer tumor cells documents heterogeneous resistance mechanisms. *Cancer Res.* **78**, 1110–1122 (2018).
92. Patten, D. K. et al. Enhancer mapping uncovers phenotypic heterogeneity and evolution in patients with luminal breast cancer. *Nat. Med.* **24**, 1469–1480 (2018).
93. Shu, S. et al. Response and resistance to BET bromodomain inhibitors in triple-negative breast cancer. *Nature* **529**, 413–417 (2016).
94. Hinohara, K. & Polyak, K. Intratumoral heterogeneity: more than just mutations. *Trends Cell Biol.* **29**, 569–579 (2019).
95. Hu, M. et al. Distinct epigenetic changes in the stromal cells of breast cancers. *Nat. Genet.* **37**, 899–905 (2005).
96. Hu, M. & Polyak, K. Microenvironmental regulation of cancer development. *Curr. Opin. Genet. Dev.* **18**, 27–34 (2008).
97. Finak, G. et al. Stromal gene expression predicts clinical outcome in breast cancer. *Nat. Med.* **14**, 518–527 (2008).
98. Azizi, E. et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* **174**, 1293–1308.e1236 (2018).
99. Costa, A. et al. Fibroblast heterogeneity and immunosuppressive environment in human breast cancer. *Cancer Cell* **33**, 463–479.e410 (2018).
100. Wu, S. Z. et al. Stromal cell diversity associated with immune evasion in human triple-negative breast cancer. *EMBO J.* **39**, e104063 (2020).
101. Heindl, A. et al. Relevance of spatial heterogeneity of immune infiltration for predicting risk of recurrence after endocrine therapy of ER+ breast cancer. *J. Natl Cancer Inst.* <https://doi.org/10.1093/jnci/djx137> (2018).
102. Natrajan, R. et al. Microenvironmental heterogeneity parallels breast cancer progression: a histology-genomic integration analysis. *PLoS Med.* **13**, e1001961 (2016).
103. Gruosso, T. et al. Spatially distinct tumor immune microenvironments stratify triple-negative breast cancers. *J. Clin. Invest.* **129**, 1785–1800 (2019).
104. Jackson, H. W. et al. The single-cell pathology landscape of breast cancer. *Nature* **578**, 615–620 (2020).
105. Wagner, J. et al. A single-cell atlas of the tumor and immune ecosystem of human breast cancer. *Cell* **177**, 1330–1345.e1318 (2019).
106. Landau, D. A. et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714–726 (2013).
107. Zhang, J. et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* **346**, 256–259 (2014).
108. Yang, F. et al. Intratumor heterogeneity predicts metastasis of triple-negative breast cancer. *Carcinogenesis* **38**, 900–909 (2017).
109. Ng, C. K. et al. Intra-tumor genetic heterogeneity and alternative driver genetic alterations in breast cancers with heterogeneous HER2 gene amplification. *Genome Biol.* **16**, 107 (2015).
110. Tentler, J. J. et al. Patient-derived tumour xenografts as models for oncology drug development. *Nat. Rev. Clin. Oncol.* **9**, 338–350 (2012).
111. Pompili, L., Porru, M., Caruso, C., Biroccio, A. & Leonetti, C. Patient-derived xenografts: a relevant preclinical model for drug development. *J. Exp. Clin. Cancer Res.* **35**, 189 (2016).
112. Invrea, F. et al. Patient-derived xenografts (PDXs) as model systems for human cancer. *Curr. Opin. Biotechnol.* **63**, 151–156 (2020).
113. Murayama, T. & Gotoh, N. Patient-derived xenograft models of breast cancer and their application. *Cells* <https://doi.org/10.3390/cells8060621> (2019).
114. Dobrolecki, L. E. et al. Patient-derived xenograft (PDX) models in basic and translational breast cancer research. *Cancer Metastasis Rev.* **35**, 547–573 (2016).
115. Bruna, A. et al. A biobank of breast cancer explants with preserved intra-tumor heterogeneity to screen anticancer compounds. *Cell* **167**, 260–274.e222 (2016).
116. Eirew, P. et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **518**, 422–426 (2015).
117. Li, S. et al. Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. *Cell Rep.* **4**, 1116–1130 (2013).
118. Woo, X. Y. et al. Conservation of copy number profiles during engraftment and passaging of patient-derived cancer xenografts. *Nat. Genet.* **53**, 86–99 (2021).
119. Powell, R. T. et al. Pharmacologic profiling of patient-derived xenograft models of primary treatment-naïve triple-negative breast cancer. *Sci. Rep.* **10**, 17899 (2020).
120. Baek, M., Chang, J. T. & Echeverria, G. V. Methodological advancements for investigating intra-tumoral heterogeneity in breast cancer at the bench and bedside. *J. Mammary Gland Biol. Neoplasia* <https://doi.org/10.1007/s10911-020-09470-3> (2020).
121. Bhang, H. E. et al. Studying clonal dynamics in response to cancer therapy using high-complexity barcoding. *Nat. Med.* **21**, 440–448 (2015).
122. Ge, J. Y. et al. Acquired resistance to combined BET and CDK4/6 inhibition in triple-negative breast cancer. *Nat. Commun.* **11**, 2350 (2020).
123. Merino, D. et al. Barcoding reveals complex clonal behavior in patient-derived xenografts of metastatic triple negative breast cancer. *Nat. Commun.* **10**, 766 (2019).
124. Wagenblast, E. et al. A model of breast cancer heterogeneity reveals vascular mimicry as a driver of metastasis. *Nature* **520**, 358–362 (2015).
125. Guernet, A. et al. CRISPR-barcoding for intratumor genetic heterogeneity modeling and functional analysis of oncogenic driver mutations. *Mol. Cell* **63**, 526–538 (2016).
126. Stewart, J., Banerjee, S., Pettitt, S. J. & Lord, C. J. Modelling the cancer phenotype in the era of CRISPR-Cas9 gene editing. *Clin. Oncol.* **32**, 69–74 (2020).
127. Zhang, M. et al. Intratumoral heterogeneity in a Trp53-null mouse model of human breast cancer. *Cancer Disco.* **5**, 520–533 (2015).
128. Cleary, A. S., Leonard, T. L., Gestl, S. A. & Gunther, E. J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113–117 (2014).
129. Marusyk, A. et al. Non-cell-autonomous driving of tumour growth supports subclonal heterogeneity. *Nature* **514**, 54–58 (2014).
130. Janiszewska, M. et al. Subclonal cooperation drives metastasis by modulating local and systemic immune microenvironments. *Nat. Cell Biol.* **21**, 879–888 (2019).

## ACKNOWLEDGEMENTS

This work was funded by Programme grants from Breast Cancer Now as part of Programme Funding to the Breast Cancer Now Toby Robins Research Centre to R.N. G.V.E. is a Cancer Prevention Research Institute of Texas (CPRIT) Scholar in Cancer Research. G.V.E. is supported by a CPRIT faculty recruitment award RR200009 and National Institutes of Health (NIH) 1K22CA241113-01. The funders have no input into the research. The content is solely the responsibility of the authors and does not necessarily represent the official views of Breast Cancer Now NIH or CPRIT.

## AUTHOR CONTRIBUTIONS

I.M., G.V.E. and R.N. wrote the manuscript. I.M. and F.J. generated the figures. All authors approved the final version of the manuscript.

## COMPETING INTERESTS

R.N. has grants and nonfinancial support from Pfizer and is an associate editor of The Journal of Pathology and Cancer Research Communications. The remaining authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Rachael Natrajan.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021