Breast cancer is extremely complex, encompassing a wide variety of entities with respect to radiological appearance, histological and molecular subtypes, therapeutic options and responses to treatment and clinical outcomes. Pathologists have recognised the histological diversity for many years and in the recent World Health Organisation Classification of Tumours of the Breast, there are at least 19 different morphological subtypes of invasive breast cancer.1 Invasive carcinoma no special type (previously called invasive ductal carcinoma NST) is the most frequent type. This classification term is inherently wide ranging, as these tumours do not demonstrate specific morphological features to be classified as a ‘special’ subtype (e.g. invasive lobular carcinoma). The special subtypes are generally more homogeneous and some convey prognostic information. For example, tubular and mucinous carcinomas are associated with favourable outcomes, whereas metaplastic carcinomas have an aggressive clinical course.3 Tumours are further subdivided according to histological grade and phenotype (e.g. the status of oestrogen (ER) and progesterone (PR) receptor proteins and Human epidermal growth factor receptor 2 (HER2) amplification) to determine the most appropriate therapeutic options. The implementation of adjunct diagnostic tests including biomarker expression (e.g. EGFR, Ki67 and E-cadherin), gene expression profiles or DNA mutation analyses, will provide incremental improvements in patient outcomes by refining existing classification systems and treatment decision-making.2,8

One of the most exciting developments in this arena relates to the recent advances in massively parallel sequencing (next generation sequencing) technology that allow scientists and clinicians to delve deeper into the genome of a tumour in order to gain a greater understanding of the complex genetic mechanisms that drive tumour growth and progression. It is now possible to sequence multiple genes to entire genomes in clinically relevant time scales and at reasonably low cost (the Human Genome project cost $3 billion and took 13 years to complete). The diagnostic utility of this technology is now within reach, particularly with the development of assays for sequencing disease-specific panels of genes. In cancer, this enables important driver gene mutations to be rapidly identified, many of which are considered ‘actionable’ in that therapies targeting the mutated gene are already available. Sequencing the cancer genome to a high coverage (deep sequencing) also enables rare, low frequency variants to be discovered, which would not have been previously detected by traditional sequencing methods such as Sanger sequencing. These developments are revealing important insights into intratumour heterogeneity and the clonal progression of disease to metastasis.

Metastatic breast cancer

By comparison with the primary tumours, there is a far more limited understanding of the complexity of both metastatic progression and metastatic lesions, despite this being the final and often fatal stage of tumourigenesis. The analysis of historical autopsy series of patients who died of metastatic breast cancer and the collection and analysis of primary and matched metastatic samples from the same individuals has and will continue to make
important mechanistic and observational contributions to our knowledge. For example, as far back as 1889, Stephen Paget considered that dissemination was non-random and that the biological characteristics of certain tumours meant they were predisposed to colonise specific distant tissues in which the microenvironmental conditions were most appropriate (the famous ‘seed and soil’ theory). James Ewing countered this idea, suggesting that tumour cells simply seed distant sites according to the mechanical flow of the blood supply/lymphatics. Elements of both of these theories are accepted in breast cancer, where the most common sites of metastatic involvement are regional axillary lymph nodes, lungs, liver, bone and brain. The particular distribution and latency of dissemination to these and other organs is influenced by the histological type, grade and phenotype of the primary tumour. For instance, invasive ductal carcinoma and invasive lobular carcinoma colonise distant tissues with different frequencies. Invasive ductal carcinoma preferentially seed lung and brain metastases and invasive lobular carcinomas preferentially spread to the gastrointestinal tract, gynaecological organs and peritoneum. In general, low grade, ER positive breast carcinomas metastasise to bone and brain more frequently than ER negative tumours (e.g. bone versus lung, liver and brain). Furthermore, pioneering work by the Massague group using a combination of animal models and clinical samples, has demonstrated that specific gene expression programs in subpopulations of cells of the primary tumour mediate tissue specific metastatic spread to bone, lung or brain.

A number of conceptual models prevail in research regarding the complex evolution of metastatic disease. In general, histological, phenotypic or molecular (gene expression and genomic profiles) features of the primary tumour are reflected in the resulting metastases, supporting the clonal nature of progression. Thus for the most part, the molecular factors defining the salient characteristics of a primary tumour are stably maintained during progression, and this tends to support the ‘linear’ model of metastatic progression. The linear model implies that metastatic capability is acquired late in development after successive rounds of mutation and clonal selection and establishment of the primary tumour. This is in contrast to the ‘parallel’ model of progression, in which tumour cells can be shed from the primary tumour site early and at any time during its development. In this model, the primary tumour and the metastasis evolve in ‘parallel’, giving rise to increased biological variance. In simple terms, the distinction between these two pathways of progression has clinical implications since they infer that a metastasis should be treated either as if it has the same biological properties as the primary tumour (linear model), or it is different and thus should be biopsied to guide treatment (parallel model).

In the endometrium, the metastasis (C) was ER and PR negative (positive staining for the epithelial marker CK8/18 is shown in the inset to demonstrate the carcinoma cells).

**Figure 1:** A clinical case to illustrate the complexity of metastatic progression. This case illustrates not only the long latency of metastatic progression for an ER positive primary tumour, but also the diversity of disease during the evolution of the metastatic clone, involving a change in the morphological growth pattern and down regulation of hormone receptors, possibly in response to exposure to endocrine-based therapy. The key episodes in the clinical history of a breast cancer patient are outlined (A). The patient was initially diagnosed with a primary breast carcinoma in 1986 and died 23 years later in 2009 of metastatic disease. The primary tumour was an invasive ductal carcinoma NST that was positive for oestrogen (ER) and progesterone (PR) receptors (B) and was HER2 negative (not shown). The first evidence of metastatic progression was in 1991 with the diagnosis of a bone metastasis. The patient underwent prolonged chemotherapy and radiation therapy at various stages throughout the disease course, and was given Tamoxifen before switching to an aromatase inhibitor prior to the development of an endometrial metastasis, which was diagnosed in 2005. The endometrial metastasis (C) was ER and PR negative (positive staining for the epithelial marker CK8/18 is shown in the inset to demonstrate the carcinoma cells).

**Morphological and phenotypic complexity**

Intratumour heterogeneity within a primary tumour, in the context of cellular or phenotypic plasticity or clonal heterogeneity, has important implications for disease progression and patient outcomes. Breast carcinomas, exhibiting mixed growth patterns reflecting different histological subtypes, are perhaps the most obvious examples of primary tumour heterogeneity. Mixed tumours evolve either as independent tumours in collision, or from a common clonal population that diverges following ongoing genetic instability or cellular plasticity. Mixed invasive carcinoma of no special type (NST) and invasive lobular carcinoma account for 3-5% of all breast cancers and are composed of both ductal and lobular histological components. The morphology of resulting lymph node metastases is more likely to reflect the predominant histological component of the primary tumour, but either...
or both components may disseminate. The diffuse growth pattern of lobular cells may be more efficient at colonising some anatomical tissues, such as the peritoneum or ovary.

Metastatic carcinomas account for up to 5% of all breast cancers and are characterised by metaplastic transformation to squamous and/or mesenchymal components. Consistent with this, tumours show striking phenotypic plasticity with heterogeneous staining patterns of markers for: i) luminal and basal/myoepithelial cell differentiation (CK19 and basal markers such as CK5/6, CK14, p63, CD10); ii) stem cell-like characteristics (CD44+/CD24−/low); and iii) mixed epithelial to mesenchymal plasticity (including down regulation of E-cadherin). This inherent ‘plastic’ nature likely contributes to their very aggressive clinical course, providing cellular capabilities to adapt, evade treatment and disseminate.

Phenotypic and genomic analysis of primary breast carcinomas and their matched metastases highlight how heterogeneous progression can be. Studies limited to analysing isolated metastases are complemented by the analysis of autopsy series of patients who have died of metastatic disease, where a more comprehensive analysis of multiple metastases can be undertaken. Key findings related to therapeutic targets in breast cancer demonstrate that ER and PR are frequently discordant and findings related to therapeutic targets in breast cancer analysis of multiple metastases can be undertaken. Key findings related to therapeutic targets in breast cancer demonstrate that ER and PR are frequently discordant and most notably down-regulated during metastatic progression (figure 1). Such down-regulation may vary between different metastases within the same patient and be non-randomly associated with colonisation of specific tissues, such as lung, bone and liver. HER2 overexpression, determined by the amplification of the ERBB2/HER2 gene, is generally more stable during progression, yet discordance has been reported in approximately 10% of cases. There are a number of explanations for this phenomenon, including the fact that expression of hormone receptors are dynamically controlled and hence may be readily down regulated during the emergence of tumour clones that are resistant to endocrine-related therapy (figure 1). Metastases may also arise from primary tumours that are themselves heterogeneous. For instance, they may exhibit non-uniform expression (e.g. ER positivity in 30% of cells) or be clonally diverse (e.g. ERBB2/HER2 amplification occurs in a tumour subclone). An integrative analysis of genomic (patterns of common amplification e.g. 8q24, 11q13) and phenotypic (status of CD24 and CD44 expression) heterogeneity at a single cell resolution reveals significant heterogeneity between primary tumours and distant metastases, and even between neighbouring metastatic cells within the same metastatic foci.

As cells metastasise, they must also adapt to survive and meet the physiological requirements of the new local microenvironment, where growth signals may be quite different to the breast tissue of origin. The brain is a foreign environment for breast cancer cells, yet some breast cancers, and cancers from other sites readily colonise brain tissue. Some breast and lung carcinomas have been shown to up-regulate HER3 signalling during colonisation as a possible mechanism of growth and survival in the brain. This adaptive response is likely due to the abundance of the HER3 ligand neuregulin, which is expressed by multiple cell types in the brain. It has also been shown that some breast cancer metastases in the brain adopt a GABAergic phenotype similar to that displayed by neurons, as a potential adaptive response to enhance survival in this new microenvironment. Understanding mechanisms of colonisation and adaptive responses of tumour cells will be pivotal in trying to both prevent colonisation of specific distant organs and develop new metastatic site-specific therapeutics.

Genomic diversity in metastatic progression

All cancers are characterised by the acquisition of somatic mutations that accumulate over the lifetime of the tumour. The pattern of mutations is extremely diverse among individual tumours, prompting large-scale initiatives to catalogue this diversity using molecular profiling (of gene expression and DNA copy number) and deep sequencing to better understand the evolution of the cancer genome during tumour development and progression, and to identify key ‘driver’ alterations for therapeutic targeting. Driver mutations confer selective growth and survival advantages to the tumour cell lineage and make a significant contribution to clonal diversity, treatment resistance and metastatic progression. Typically, critical driver mutations (mutation of TP53 and PIK3CA, amplification of MYC, CCND1, ERBB2/HER2) occur within the early phases of tumour growth and are thus present in the majority of cells of the primary tumour. The late acquisition of new driver mutations or amplifications may further drive clonal diversity in a primary tumour or a metastasis, and these alterations may facilitate dissemination and/or treatment resistance.

Elegant massively parallel sequencing and copy number profiling studies of breast, renal cell, prostate, colorectal, and pancreatic carcinomas, as well as leukaemias, exemplify these findings at nucleotide resolution and demonstrate clonal evolution and metastatic progression can be very complex in some instances, but monoclonal in others. Multiple sampling from spatially defined tumour regions or multiple metastatic biopsies, and phylogenetic relationship modelling based on mutation clonal frequencies, clearly delineates significant subclonal diversity. The trunk of the phylogenetic tree represents founder mutations that are responsible for the initiation and establishment of the tumour, and these mutations are present within all resulting subclones. In fact, many of the mutations occur in the primary tumour. The branches of the tree represent the evolution of distinct subclonal populations that arise due to the acquisition of ‘progressor’ somatic mutations in specific lineages, some of which progress to metastases in different organs. The data imply that different subclones within a primary tumour have developed metastatic capability. As a consequence, this yields genomic diversity in metastatic samples from the same patient. Further clonal evolution can occur at the metastatic sites involving the alteration of key driver genes such as KRAS, MYC and CCNE1. Interestingly, distinct and spatially separated inactivating mutations within the same tumour suppressor genes (e.g. SETD2 and PTEN) were observed in the same patient, suggesting the evolution and selection of separate clones with convergent phenotypic characteristics.
Conclusion and clinical challenges

A considerable challenge with treating breast cancer patients lies in the diversity of disease, with respect to subtypes, treatment responses and outcomes. The inherent nature of some individual tumours to exhibit diversity, compounds this complexity and plays a significant role in the development of resistance and metastatic progression (figure 1). An understanding of both the basis and the extent of genomic and phenotypic diversity would therefore be extremely valuable in helping to determine guidelines for the clinical management of metastatic disease, including predicting sites of relapse, identifying mechanisms of treatment resistance and defining the most appropriate treatment options.

When tumours progress after treatment, taking a biopsy of the metastatic deposit to guide treatment decisions is not routine, and so clinical decisions are made based on predictive biomarkers performed on the primary tumour specimen. From the clinical point of view, it is not always practical or even possible to biopsy metastases, in which case there is little choice but to be guided by the biological features of the primary tumour. Taking evidence from the detailed genomic studies illustrating the clonal heterogeneity of tumour development and progression, it might be most fruitful to target the early driver mutations found in the ‘trunk’ of the phylogenetic evolutionary tree, since these alterations are likely to be consistent in all resulting subclones that may develop. Where biopsy of metastatic deposits is possible, it may help delineate treatment choices. In a recent multicentre clinical trial, biopsies of accessible breast cancer metastases were obtained from 407 patients, and of these, 283 were analysed to identify potentially targetable genomic alterations using DNA copy number profiling and Sanger sequencing for hot spot mutations in PIK3CA and AKT1. A potentially targetable alteration was identified in 46% of patients and 13% of the cohort received targeted therapy based on their molecular results. This study has demonstrated the feasibility of performing molecular testing on metastatic biopsy samples for the development of personalised medicine.

References