Exploring tumor heterogeneity
Fessler, E.

Citation for published version (APA):
Fessler, E. (2016). Exploring tumor heterogeneity
Chapter 1

Cancer heterogeneity – a multifaceted view.

Evelyn Fessler and Jan Paul Medema.

Adapted from EMBO reports 14 (8), 2013.
Felipe De Sousa E Melo, Louis Vermeulen, Evelyn Fessler, and Jan Paul Medema.
Abstract
Cancers of various organs have been categorized into distinct subtypes after increasingly sophisticated taxonomies. Additionally, within a seemingly homogeneous subclass, individual cancers contain diverse tumor cell populations that vary in important cancer-specific traits such as clonogenicity and invasive potential. Differences that exist between and within a given tumor type have hampered significantly both the proper selection of patients that might benefit from therapy, as well as the development of new targeted agents. In this review, we discuss the differences associated with organ-specific cancer subtypes and the factors that contribute to intra-tumor heterogeneity. It is of utmost importance to understand the biological causes that distinguish tumors as well as distinct tumor cell populations within malignancies, as these will ultimately point the way to more rational anti-cancer treatments.

Keywords: cancer stem cells, cancer subtype, clonal evolution, therapy resistance

Glossary:
ABL  abelson murine leukemia viral oncogene
APC  Adenomatous polyposis coli
BCR  B-cell receptor
BRAF  v-raf murine sarcoma viral oncogene homologue B1
c-KIT  v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homologue
c-MET  hepatocyte growth factor receptor
DLL4  delta-like ligand 4
EGFR  epidermal growth factor receptor
EMT  epithelial–mesenchymal transition
HER2  human epidermal growth factor receptor 2
HGF  hepatocyte growth factor
KRAS  kirsten rat sarcoma viral oncogene homologue
MLH1  MutL homologue 1
MMTV  mouse mammary tumor virus
NF-κB  nuclear factor kappa-light-chain-enhancer of activated B cells
PDGFR-α  platelet-derived growth factor receptor alpha
PyMT  polyoma middle T antigen
TGF  transforming growth factor
TGFB2  TGF beta receptor 2
TP53  tumor protein p53
Introduction

The first pathological reports made it clear that a single neoplasm presents with tremendous variability among individual tumor cells. This so-called intra-tumor heterogeneity poses an important challenge for predicting tumor behavior and clinical outcome. Differences exist not only within a single tumor, but also across individual patients presenting with cancers originating in the same organ. Increasing knowledge of this inter-tumor heterogeneity has led to an exhaustive categorization of tumor subsets according to staging, differentiation grade, cellular morphology, and marker expression.

Intra- and inter-tumor heterogeneity and how they relate to each other are the main focus of this review. The emergence of advanced molecular and biochemical technologies has led to a better understanding of the numerous mechanisms by which both faces of tumor heterogeneity are driven. Indeed, gene expression arrays and next-generation sequencing have paved the way for a more comprehensive indexing of a broad variety of cancer types, such as breast and brain tumors (1-3), allowing their separation into robust subtypes with markedly different molecular and clinical qualities. In addition, scrutiny within individual tumors has reached tremendous sensitivity, and many genetic and epigenetic modifications, as well as micro-environmental interactions, can be dissected at the single-cell level (4). Despite all these efforts, we are only starting to appreciate the complexity of tumors and our limited understanding of the molecular mechanisms by which individual cancers are driven has precluded the development of curative targeted therapies. In more advanced disease for instance, tumors relapse in most cases even after a seemingly successful therapy. This distressing fact has often been attributed to the presence of heterogeneity within tumors, and two main conceptual frameworks have been elaborated to conceptualize the link between intra-tumor heterogeneity and therapy resistance. The first, and most well-established idea, rests on classical evolutionary concepts and postulates that selective pressure acting on tumor cells ultimately leads to (epi-)genetically distinct, resistant clones (5, 6). The second and more recent theory is based on the realization that subsets of tumor cells inherently have a dissimilar ability to drive tumor growth and metastasis (7, 8). In this scenario, tumors are perceived as hierarchically organized tissues and differences between tumor cells are mainly explained by a gradient of differentiation that exists between cancer stem cells (CSCs) residing at the top of the hierarchy, and their more differentiated progeny. As we discuss below, these two theories coexist perfectly and are complementary rather than mutually exclusive (9). In addition, tumor heterogeneity is also influenced by the cell-of-origin and involves the presence of non-neoplastic cells that establish a tumor microenvironment.
Here, we provide an overview of several aspects that contribute crucially to tumor heterogeneity (Fig 1) and discuss their implications in a clinical context. Moreover, we highlight the connections that exist between intra- and inter-tumoral heterogeneity in an effort to draw parallels that might contribute to our basic understanding of these two sides of the same coin.

The actors of tumor heterogeneity

The genetic contribution. Genetic variation is undoubtedly the most established foundation of both intra- and inter-tumor heterogeneity. The stepwise acquisition of mutations during cancer evolution was first advocated by Nowell in 1976 (10), and later on substantiated by a large body of experimental work. As tumors progress, multiple and changing environmental pressures operate on the cancer cell population and select for clones that are best endowed to survive and respond to these changes (Fig 2A). Adaptation can be enabled by genetic mutations that confer a selective advantage to a particular clone. This process is highly dynamic and usually selects for clones that will grow and dominate the tumor through so-called ‘selective sweeps’. This is

Figure 1. Schematic representation of the various factors (diagonal, red boxes) that affect intra- (x-axis, light blue) and inter- (y-axis, light grey) tumor heterogeneity. Several examples are depicted for each factor and for both forms of tumor heterogeneity. CIN – chromosomal instability; CSC – cancer stem cell; MSI – microsatellite instability; SHH – sonic hedgehog signaling.
reiterated as long as the tumor ecosystem evolves, leading to the formation of complex clonal patterns (6). In their landmark paper, Fearon and Vogelstein crystallized this concept by looking at different mutations in a series of patients comparing normal epithelium with cancer precursor lesions and cancer tissue. They reported increasing amounts of mutations with more advanced disease, but importantly, also proposed a preferential sequence of events by inferring the prevalence of certain mutations at each stage (11). Today the picture has expanded considerably and a full genomic landscape has been drawn for a large variety of malignancies (12-14), only to strengthen the notion that clonal evolution is crucial to tumorigenesis. Individual cancers can harbor from a handful (15, 16) to hundreds of mutations (17). This startling diversity has urged efforts to define patterns and sets of mutations that actively contribute to disease initiation and progression, also dubbed ‘driver’ mutations as compared with bystander or ‘passenger’ mutations, which are not causally implicated in oncogenesis but are carried along and accumulate as cancers develop. Identifying the biological nature of driver mutations in tumors, such as those conferring a growth advantage, is relatively straightforward, but the actual benefit they provide for tumor cells remains poorly investigated. Indeed, few studies have attempted to quantify the evolutionary advantage that mutations confer to a clone (18). Models in which clone size can be easily measured longitudinally (19) are needed for this purpose, such as the normal colonic crypt, which provides a unique architecture to study stem cell homeostasis. Colonic crypts become monoclonal over time, following a pattern of neutral drift dynamics in which functionally equivalent stem cells either expand or disappear stochastically until they either take over the crypt or are lost (20, 21). Therefore, a mutation specifically initiated in stem cells can affect both neutral drift and the speed at which an affected crypt reaches monoclonality. This can subsequently be used to infer the evolutionary benefit of that particular mutation (22). In contrast to the lack of quantitative measurements on genetic hits, it is more apparent that distinct mutations have different impacts on tumor heterogeneity. For instance, inactivation of the tumor suppressor MLH1 in colorectal cancer (CRC) impairs the DNA damage response and leads to the formation of tumors with a high mutation rate and a so-called ‘hypermutator’ phenotype (23). Finally, assessing clonal heterogeneity with existing techniques requires the evaluation of topologically distinct regions. This is beautifully illustrated in renal carcinoma in which whole-exome sequencing of multiple distinct tumor regions has revealed considerable genetic disparities (24). In keeping with this notion, Sottoriva and colleagues sampled single crypts from opposite sides of CRCs and note that the phylogenetic trees they construct based on genetic information would be different if only information from one side would have been considered (25). Additionally, the high resolution achieved in the genetic and epigenetic analyses allows the
authors to speculate that intra-tumor heterogeneity is installed early in tumor development and that the relative abundance of a selected feature is not a measure of its fitness but of its age (25).

In this scenario, the complexity of clonal patterns is not a result of Darwinian selection and ‘selective sweeps’ but results from a single expansion at an early stage. Surprisingly, the tumor microenvironment was assigned a passive role, as it did not affect tumor growth dynamics (25). Yet, microenvironmental components can act on transformed cells and alter their biological and clinical characteristics (26). Thus, how this so-called ‘Big Bang model’ can be reconciled with the evolutionary development of tumors and whether interconversion occurs upon selective pressure, such as therapy, warrants further investigation.

Genomic differences do not exist solely within an individual cancer but also, and arguably most prominently, among cancers. It is evident that the range of mutations between seemingly identical cancers can vary substantially between patients. Even the type of driver mutations found in each cancer can differ highly. This has often resulted in the identification of molecular subsets of cancers that are made up of diverse mutational spectra (Fig 2D). In the case of CRC, many previous reports have attempted to classify this disease in different categories according to its respective genetic defects (27, 28). For instance, activating mutations of the BRAF oncogene are associated with a particular type of CRC that displays instability in microsatellite (MSI) DNA repeats. Another category shows chromosomal instability and associates with mutations in the TP53 and KRAS genes. Although certain mutation patterns associate with a subset of cancers, the link between the genetic make-up of a tumor and its phenotype is fairly weak. For example, on the most abstract level, similar mutations can result in a vastly different phenotype, whilst the same phenotype can result from different mutations. CRC patients with BRAF mutations illustrate the former, and have markedly different biology and clinical outcome whether or not they present with MSI (27). Conversely, gastrointestinal stromal tumors (GISTs) that are induced by either c-KIT or PDGFR-a mutations, present with remarkably similar histology and therapy response (29). Furthermore, our group has shown that unbiased determination of the most biologically distinct CRCs by using gene expression arrays results in the identification of robust subsets that have crucial differences in prognosis, despite a significant overlap in their mutation spectra (30). Moreover, CRCs can show striking overlap in gene expression profiles and clinical behavior even though lacking unifying mutations (27, 31). This further demonstrates that determining mutations per se is insufficient to appreciate fully biological and, as a consequence, clinical behavior. Therefore, efforts have been focused on gene expression-based classifications for many cancer types which resulted in the detection of several subtypes within a given tumor.

As discussed in Chapter 4, an international subtyping consortium revealed the existence of four
Cancer heterogeneity – a multifaceted view.

The contribution of epigenetics to tumor heterogeneity. Besides genetic diversity, there is increasing recognition that epigenetic changes, such as DNA methylation and histone deacetylation, can occur throughout tumorigenesis (34). These changes, and the subsequent biological processes they affect, are inherited during cell division without altering the DNA sequence. DNA methylation is instrumental for crucial biological processes during embryonic development and cellular differentiation (35), and is mediated by DNA methyltransferases that transfer methyl groups to cytosine residues. This process occurs mostly at CpG dinucleotides, which are distributed with high frequency at the promoter region of genes (36). Aberrant DNA methylation is a characteristic detected in many cancer types that present often with global loss of DNA methylation (37). More specific alterations are also frequently seen, such as those that target p16 (38), which is a prime example of methylation-mediated inactivation of tumor suppressor genes. In addition to contributing to tumorigenesis, epigenetic alterations are also linked to tumor heterogeneity in numerous ways. Within an individual tumor, DNA methylation patterns are particularly polymorphic, even at single DNA promoter regions that are frequently methylated, which suggests the presence of heterogeneous cell populations with regards to DNA methylation (39). In breast cancers, subsets of tumor cells differentially methylate the TGFBR2 promoter, which results in differential TGF pathway activity (40). Additionally, cell type specific epigenetic modifications of DNA and of histone proteins shape the response to TGFβ, providing an explanation for the highly context-dependent nature of this signaling molecule (41). Furthermore, methylation-mediated silencing of key gatekeeper genes, such as MLH1, might fuel clonal evolution as they promote a hypermutation phenotype (23). Interestingly, this particular case is directly related to inter-tumor heterogeneity as MLH1 methylation is a defining feature of the MSI–CRC subtype (28). Moreover, methylation of a subset of Wnt target genes has been shown to identify colon cancers with poor prognosis (42). In addition, in Chapter 5 we show that promoter methylation of the miR-200 family displays high levels in CRCs belonging to the mesenchymal subtype and not only provides a means to identify these tumors, but also rationalizes their mesenchymal appearance (43). Importantly, data presented in Chapter 6 indicate that this epigenetic feature can already be detected at the premalignant stage. Finally, certain subtypes of CRCs and gliomas have even been defined by a high level of methylation at the CpG-rich promoter regions of a subset of genes (44, 45). The reversible nature of epigenetic
modifications has attracted much research efforts to develop agents that target the enzymes responsible for these changes. Several clinical examples illustrate the promise of epigenetic targeting, translated to the bedside by the US Food and Drug Administration approval of drugs that inhibit DNA methylation, and histone deacetylation for the treatment of myelodysplastic syndrome and cutaneous T-cell lymphoma, respectively (36).

The role of cancer stem cells in tumor heterogeneity. Normal tissue homeostasis is orchestrated by complex interactions between cells and their microenvironment that maintain a crucial balance between proliferation and differentiation. Importantly, all these processes occur in an ecosystem that is genetically identical and relies on a gradient of differentiation between stem cells and their differentiated progeny. An equivalent paradigm is not new in cancer (46), but only recent advances in technologies have made it possible to investigate critically the role of cancer cells endowed with stem cell properties. Initially described in leukemia, the concept of CSCs has been rapidly embraced by the scientific community and led to their identification in virtually all tumor types (7-9). The principle generally relies on the use of markers that enrich for cells able to propagate the malignancy in serial xenotransplantation assays. The exact identity of CSCs in many tumors is still disputed (47), especially due to the imperfect nature of the methodologies available to assess cancer stemness. Although more recent reports lend support for the existence of CSCs in endogenously growing neoplasia (48-50), they also have their own caveats. For instance, both intestinal adenoma and skin papillomas used in these studies are only benign precursor lesions that progress infrequently to carcinomas. The CSC concept nevertheless provides an additional source of heterogeneity in cancer; heterogeneous cell populations are defined by their differentiation state, non-differentiated (or CSCs) as compared with differentiated cells, as well as the various differentiation lineages that can be adopted (Fig 2B). Importantly, the genetic and CSC models are not mutually exclusive but rather complementary in fueling tumor heterogeneity (9). For example, CSCs are – by definition – the only cells with self-renewal capacity and as such would serve as a repository for clonal evolution to take place. Indeed, genetically diverse CSC populations within individual tumors have been demonstrated in both hematological and solid malignancies (40, 51, 52). Reciprocally, genetic mutations such as inactivation of TP53 can influence the CSC content (53). The molecular attributes that maintain the CSC state remain largely elusive but we have gained an increased insight into the molecular determinants of cancer stemness. For example, colon CSCs were shown to have relatively higher levels of Wnt pathway activity as compared with their differentiated progeny, despite the presence of Wnt-activating mutations (54). Although deregulated in all colon carcinomas, Wnt
does not necessarily define colon CSCs in every tumor. There might be tumors with CSCs that rely on alternative signal transduction cascades or that do not even follow a hierarchical organization (47, 55), as illustrated in CRC or breast cancer cell lines (56, 57).

The role of CSCs in regulating inter-tumor heterogeneity is more elusive, although subtype-specific regulation of CSC traits has also been noted. First, experimental mouse models of lung cancer with distinct genetic backgrounds were shown to have a crucial impact on the identity of the CSC population (58). Second, an elegant study by Caldas et al. reported that TGFβ potently induces a stem cell phenotype in basal and low-claudin subtypes, but has a markedly different and even opposing role in other breast subtypes, such as luminal (Fig 2E; (59)). Finally, besides xenotransplantation assays, analyses of phylogenic trees of tumor populations have also resolved that CSC fractions vary among cancers (60). These subtype-specific features of CSCs might explain some long-lasting inconsistencies in the field, such as the discrepancies associated with markers commonly used to identify CSC populations (47). Regardless, these results point towards the CSC phenotype as an important contributor to heterogeneity across cancer subtypes. The implications of the CSC model for tumor growth are further illustrated by mathematical modelling of tumors that follow this model, which has revealed that hierarchical organization promotes the acquisition of new genetic and epigenetic traits and, thereafter, increased heterogeneity (61). It should be noted however that differentiation hierarchy is not the only non-genetic explanation of tumor heterogeneity. In some instances, more stochastic processes seem to dominate the scene (56, 62).

The impact of the tumor microenvironment on heterogeneity. Environmental cues have a major role in tumor development and progression, but also determine the heterogeneity observed within and across tumors. The tumor microenvironment (TME) comprises a wide variety of non-neoplastic cells, coopted by tumor cells that influence the various steps of cancer development. The TME includes cancer-associated fibroblasts, endothelial cells, and a variety of infiltrating immune cells, and promotes tumor heterogeneity in various ways. The mere presence of TME components is, by itself, a source of heterogeneity, and these cell types can be distributed in different tumor microenvironments, such as the invasive compared with the non-invasive edge (63). Extrinsic factors produced by stromal cells also affect tumor heterogeneity by mediating changes in clonal evolution rate or stem cell content. For example, chronic inflammation can lead to oxidative stress and increased production of reactive oxygen species, which are potent inducers of DNA damage (64). Consequently, this might influence intra-tumor heterogeneity in
tumor regions with localized inflammation, but also affect cancer subtypes that associate strongly with chronic inflammation.

In the context of CSCs, the TME regulates tumor hierarchy and stem cell traits. Normal stem cell homeostasis is ensured by the integration of signals that fine-tune the balance between self-renewal and lineage commitment. Part of these signals emanate from the microenvironment, commonly referred to as the ‘stem cell niche’ (65). An equivalent entity has been shown to have similar albeit aberrant functions in cancer. For instance, tumor-associated myofibroblasts produce HGF that binds to c-MET and activates Wnt activity to support colon CSCs (54). The Notch ligands DLL4 and more recently Jagged1 were shown to have similar effects (66, 67). In

![Figure 2. Cartoon representation of several factors affecting tumor heterogeneity. (A–C) Intra-tumor heterogeneity. (A) From a common ancestor (blue founder cell), different subclones (represented by different colors) emerge due to selection and distinct mutations. (B) In the CSC model, clonal evolution still takes place but only acts on the CSCs. These cells can self-renew and give rise to the various cell lineages present in a tumor (differentiated cells in respective colors). (C) The TME affects intra-tumor heterogeneity by its composition (myofibroblasts in pink and immune cells in grey) as well as their derived factors that can induce the reversion of differentiated cells to CSCs. (D–F) Inter-tumor heterogeneity. (D) The MSI or CIN colon cancer subtypes are distinct subtypes that associate with BRAF or APC and KRAS mutations, respectively. (E) TGFβ pathway activation results in an increase or decrease of CSC phenotype when triggered in the basal or luminal breast cancer subtype, respectively. (F) SDPPs can be used to identify breast cancer patients that differ in their survival probability. APC – adenomatous polyposis coli; BRAF – v-raf murine sarcoma viral oncogene homologue B1; CIN – chromosomal instability; CSC – cancer stem cell; KRAS – kirsten rat sarcoma viral oncogene homologue; MSI – microsatellite instability; SDPP – stroma-derived prognostic profile; TGFβ – transforming growth factor beta; TME – tumor microenvironment.](image)
glioma, endothelial cells co-localize to and are required for glioma stem cell maintenance (68).

Altogether, these and other studies (63) have demonstrated the impact of the TME on the heterogeneity of tumors. Perhaps even more importantly, as discussed in Chapter 2, factors provided by tumor-associated cells not only maintain stem cell attributes but also induce the phenotype in more differentiated cells (Fig 2C). For instance, HGF was shown to induce differentiated colon cancer cells to revert to a CSC state (54), and similarly TGFβ triggers an EMT–CSC program in mammary epithelial cells (69). Moreover, in Chapter 3 we provide evidence that basic fibroblast growth factor derived from tumor-associated endothelial cells can direct differentiated glioma cells to a CSC-like state (70). This extrinsic control of tumor cell flexibility poses an obvious challenge in the design of therapies, especially those aimed at targeting the clonogenic core of tumors (71), and suggests that alternative therapeutical options have to be considered (72). The composition of the TME is not only important in the regulation of tumor cell properties, but it also can be used as a defining peculiarity among different cancer subtypes and might provide opportunities for targeted therapy. For example, activation of CD8+ T cells in specific tumor regions predicts tumor recurrence in colon cancer (63). Furthermore, stromal-derived profiles from breast cancer have been associated with cancer subtypes suffering with a worse outcome (Fig 2F; (73)). Interestingly, the impact of the TME might also be specific to certain tumor subtypes. Mammary tumor metastasis in the MMTV-PyMT breast cancer model depends crucially on the recruitment of CD4+ T cells (74) unlike MMTV-NeuT-derived tumors (75). A recent study delineated the microenvironmental make-up of the consensus molecular subtypes of CRC (76). Whereas two of the subtypes did not display distinctive microenvironmental signatures, both the subtype associated with good and with poor prognosis were characterized by an inflammatory signature (76). Integrating information on the phenotype and behavior of malignant cells with the composition and activation of stromal components might provide a rationale for distinct therapeutic approaches, such as immunotherapies using checkpoint inhibitors (76). The recognition of the TME as a defining feature of cancer subtypes is crucial when differences in mutations or stem cell content cannot explain the apparent diversity. For example, similar mutations might accumulate in a different sequential order, leading to a distinct requirement towards the environment. Alternatively, analogous mutations following the same sequence of events might give rise to divergent molecular subtypes depending on the inflammatory or environmental response. Indeed, data presented in Chapter 7 indicate that sessile serrated adenomas – precursor lesions to CRC characterized by the \(BRAF^{V600E}\) mutation and DNA hypermethylation (77, 78) – can either develop to the MSI or mesenchymal CRC subtype and that this divergence can be orchestrated by differential TGFβ
signaling activity \( (79) \). It is therefore important to view distinct cancer subtypes as complex systems in which both the presence of malignant and non-transformed cells as well as their crosstalk shape the resulting phenotype.

**The effect of the cell-of-origin on tumor heterogeneity.** Cancer develops from a single founding cell, the identity of which is still highly debated \( (80, 81) \). In most adult tissues, self-renewal capability is restricted to stem cells and therefore these are probably the prime target for cellular transformation. Adult somatic stem cells are usually long-lived, which is a prerequisite to accumulate sufficient transforming events. Mouse models of intestinal cancer have initially confirmed the latter hypothesis as adenomas were shown to develop preferentially when loss of \( APC \) is induced in the stem cell compartment \( (82) \). However, more recent work has challenged this view and has demonstrated that dedifferentiation of intestinal enterocytes can occur and lead to intestinal tumors \( (81, 83) \), provided they have the right conditions. The Verma group reported another example in which they used a model of glioma and showed that terminally differentiated cells are also potent inducers of cancer, given that the right combination of genetic aberrations is induced. More specifically, the authors \( (80) \) deleted the tumor suppressor \( p53 \) together with the activation of the \( KRAS \) oncogene in differentiated neurons. They observed the induction of tumors that progressed rapidly towards invasive gliomas. Although these examples suggest that there is no restriction on the cell type that is initially targeted, the presence or acquisition of stem cell traits could still represent a limiting step to transformation. In the intestine, enforced activity of both Wnt signaling and NF-\( \kappa B \) in enterocytes is required to initiate adenomas from these cells and is accompanied by the re-expression of genes present in the intestinal stem cell signature. Similarly, acute myeloid leukemia can develop either from hematopoietic stem cells (HSCs) or more committed progenitor cells on translocation of the mixed lineage leukemia \( (MLL) \) gene. When occurring in committed granulocyte–macrophage progenitors, \( MLL-AF9 \) activates a self-renewal signature normally restricted to HSCs and generates tumors with high leukemic stem cell content \( (84) \).

Moreover, as discussed in Chapter 8, the cell-of-origin can be a source for inter-tumor heterogeneity, and mapping the susceptible cells of different cancer subtypes is intensely investigated. A striking example was reported in medulloblastomas, one of the most frequent pediatric brain tumors originally thought to arise predominantly in the cerebellum. At least two main subtypes have been described – one driven by aberrant sonic hedgehog signaling (SHH) and the other seemingly dependent on Wnt activation – which have a dismal and good prognosis, respectively. By comparing the transcriptomes of these subtypes to those of distinct precursor
cells, Gibson et al. revealed that tumors from the SHH subtype were closely associated with committed cerebellar granule neuron precursor cells and developed from the cerebellum (85). By contrast, the Wnt-driven subtype is associated with a gene expression profile of neural precursor cells located in the embryonic dorsal brainstem. Several additional examples have been described (86), and it is expected that many molecularly distinct subtypes from various tumors will be quickly trailed back to their origin. Importantly, in some cases, transcriptional programs alone might be insufficient or even misleading in inferring the cell-of-origin of cancer subtypes. Basal cell carcinomas (BCCs) develop through activation of the hedgehog pathway, and during progression these tumors often show molecular markers normally expressed by hair follicles. Although this would suggest these tumors originate from hair follicle stem cells, conditional activation of the hedgehog pathway in various skin lineages has demonstrated that most BCCs develop preferentially from stem or progenitor cells found in the interfollicular epidermis (87). In line with these observations are studies in breast cancer which confirm that the differentiation state of the tumor does not always allow drawing conclusions about the cell-of-origin (88-91).

In summary, we have reviewed two forms of tumor heterogeneity, namely intra- and inter-tumor heterogeneity, and discussed various sources that can promote them (Fig 1). First, we described the genetic contribution to heterogeneity resulting from an accumulation of genetic mutations in tumor cells and subsequent clonal selection; the process of clonal selection acts mostly on the CSCs as they are uniquely endowed with the capacity to self-renew. Moreover, we have discussed how CSCs add another layer of heterogeneity as these cells can spin off different tumor lineages of differentiated tumor cells. Finally, we concluded this first part by presenting the role of the tumor microenvironment and the cell-of-origin.

Implications and crosstalk of both forms of heterogeneity

Clinical implications. The mere presence of heterogeneity, regardless of the underlying foundation that promotes it, has a profound clinical impact. In the case of inter-tumor heterogeneity, the presence of oncogenic mutations is often used to guide treatment decisions. Two main situations can be described: on the one hand, the presence of a mutation could directly predict a lack of response to a particular treatment. This is the case in metastatic CRC, for which anti-EGFR therapy is relatively effective in KRAS wild-type cancer subtypes, but is not effective in KRAS-mutant tumors (92). On the other hand, certain tumors are ‘addicted’ to oncogenic aberrations and targeting these mutations has proven clinically useful in those tumors. For instance, melanomas with a BRAF mutation are eligible and sensitive to treatment with the BRAF inhibitor vemurafenib as opposed to BRAF wild-type tumors (93). As mentioned
previously, mutation-based categorization incompletely recapitulates the complexity and
diversity of cancer subtypes, and sub-classification of tumors based on gene expression profiles
potentially improves clinical decisions. In this respect, unbiased identification of CRCs on the
basis of gene expression has revealed the presence of a resistant subset to anti-EGFR treatment,
independent of *KRAS* mutation status (30). Similarly, different pancreatic adenocarcinoma
subtypes show differential response to anti-EGFR treatment despite the presence of *KRAS*
mutations (94). The use of genomic technologies has resulted in an extensive characterization of
breast cancers into distinct molecular subtypes and has yielded a major clinical benefit for some
of these subsets. For example, the use of tamoxifen (an estrogen receptor antagonist; (95)) and
trastuzumab (an anti-HER2 monoclonal antibody; (96)) has crucially improved the survival of
patients belonging to the estrogen receptor or HER2 subtype, respectively.

Intra-tumor heterogeneity also has an impact on clinical outcome. In diagnostic terms, the choice
of a therapeutic intervention is constrained by topological heterogeneity, as sampling procedures
do not yield a fully representative picture of a tumor (24, 97). Furthermore, diagnosis is
generally derived from the primary tumor whereas treatment often aims at eradicating metastatic
disease presenting with phenotypic attributes that have progressed beyond the primary tumor.
Finally, much of our knowledge on tumor heterogeneity is derived from ensemble
measurements, which reflect changes present in most cells. Although emerging, the genetic
make-up of single cells within a tumor (4) remains a formidable challenge, but will result in a
better interpretation of clonal genotypes. A recent study, for instance, made use of single cell
RNA sequencing of individual glioma cells and revealed that not only pieces from different areas
of a given tumor, but also single cells within a tumor correspond to different glioma subtypes
(98). The question remains, however, of whether these diverse issues are rate-limiting steps that
need to be resolved to improve clinical outcome for patients? In other words, is it essential to
obtain multi-regional sampling to have a correct diagnosis and successful therapy? For instance,
some approaches might be available to filter through tumor complexity. One of them relies on
the use of xenotransplantation as xenografts of human breast primary tumors display an
enrichment of mutations present in the metastatic lesion counterpart. This suggests that this
method could be used as an extra step to enrich for the most aggressive clones (99), from which
diagnosis and therapy design should be made. Furthermore, the clinical benefit of obtaining
genetic information on single tumor cells remains uncertain as it will only provide a complex
architectural view of the clonal genotype without further revealing the dependence of tumor cells
on specific growth factors or signaling pathways. Intriguingly, it was recently shown that cells
within one glioma tumor display differential expression of receptor tyrosine kinases (RTKs)
Cancer heterogeneity – a multifaceted view.

Inhibition of RTKs is a mainstay of targeted therapy and the observed cell-to-cell variability has thus profound implications for therapeutic strategies. Yet, the relevance of single-cell-derived diagnosis is unclear, as data derived from these cells does not provide any direct information on the remnant tumor cell population, and other more accessible methods can be used to infer clonal relations (5, 100).

Therapy response is also affected by heterogeneity as targeted drugs often suffer from their highly selective nature towards specific gene alterations. When a selective pressure is applied, such as therapy, the fittest – most resistant – subclones are invariably selected to survive and expand to manifest clinically as a tumor relapse (5, 6). Mutations can directly affect the target itself, as in the case of chronic myelogenous leukemia (CML) when treated with imatinib (101). Alternatively, mutations can act synergistically to reactivate signaling pathways that are targeted. As mentioned above, relapses of metastatic CRC treated with anti-EGFR are frequently accompanied with mutations in KRAS, a downstream effector of EGFR signaling (92). Alternatively, feedback mechanisms are often activated to confer resistance. One such example has been described in CRC in which resistance to BRAF inhibition is bypassed by EGFR overexpression (102). In most cases, therapy resistance is accompanied by the outgrowth – selective sweep – of genetic variants that existed before the introduction of the treatment. In CML, a variant of the ABL kinase, a crucial target of the drug imatinib can be detected before therapy in patients that will develop resistance to that drug (103). Similarly, the emergence of KRAS-mutant clones might be noticed months before radiographic progression in metastatic CRC patients treated with anti-EGFR therapy (104). Likewise, around 5% of all gliomas are composed of coexisting subclones of which each is characterized by the amplification of a distinct RTK (105). This mosaicism might hold important implications for therapy: if these clones form an interconnected network and depend on each other, a single RTK inhibitor might be sufficient for effective treatment. However, subclones could also represent distinct self-sustaining entities, which might explain the low efficacy of anti-EGFR therapy in glioma and might point to the requirement of a multiple RTK-targeting strategy (105). More recently, longitudinal assessment of genomic alterations before and after treatment in a panel of chronic lymphocytic leukemias has revealed that an adverse clinical outcome was related to the expansion of subclones that were already present before treatment (100). Moreover, a recent study analyzed the dynamics of resistance development and showed that the EGFR T790M mutation, which is responsible for resistance in 50-60% of lung cancer cases treated with EGFR inhibitors, was detectable in lung cancer cell lines prior to treatment with the EGFR inhibitor gefitinib (106). Upon EGFR inhibition the EGFR T790M-mutant cells are selected for and lead to
rapid outgrowth of drug-resistant colonies. In addition, the authors also observed outgrowth of colonies only after prolonged EGFR inhibition, which suggests that these colonies were initiated by the formation of *de novo* mutations and not by pre-existing mutant cells. Mutations associated with this late resistance are either the $\text{EGFR}^{T790M}$ mutation itself or mutations in downstream signaling components (106). This study therefore offers an explanation for both heterogeneity observed in the time to recurrence development as well as heterogeneity in sensitivity to second line therapeutics (106).

Importantly, in the context of CSCs, therapy seems to select for pre-existing, highly resistant tumor cells. It is generally assumed that these cells are more resistant to a variety of treatments (71, 107, 108). Evidently this has a major clinical impact, as eradication of this subpopulation should be the priority when designing new therapeutic interventions. For example, leukemic stem cells (LSCs) identified in CML are reportedly more resistant to the drug imatinib than their differentiated progeny. This is often attributed to the quiescent nature of LSCs (109, 110) as imatinib preferentially targets dividing cells. Similarly, the quiescent state of a fraction of colon CSCs has also been associated with chemotherapy resistance (62), and altogether these studies highlight the potential therapeutic benefit of reverting quiescence (110). Additional experimental evidence comes from glioma in which CSC marker-expressing cells repair DNA damage more efficiently (111), and the ability of breast CSCs to maintain low levels of reactive oxygen species, which protects them against radiation (112). More clinical examples have also been reported, for instance in GISTs, whereby patients showing complete disease remission under imatinib treatment relapsed quickly after imatinib withdrawal (113). This suggests that a fraction of cancer cells remain untouched during the treatment and cause a relapse.

The number of CSCs might relate to metastatic potential and, therefore, it is not too surprising that many efforts have been made to better identify tumors at a higher risk of relapse, with the assumption that they would present with a higher CSC content. The identification of patients with poor prognosis in breast cancers (114), colon cancers (42, 115), and leukemia (116), based on their association with CSC signatures, has indeed supported the relevance of CSCs in a clinical setting. Frequently, the underlying biological association of the tumor transcriptome to a CSC signature has been thought to reflect the fraction of CSCs present in a tumor. Although these associations might exist in malignancies that have a high CSC content, it is unlikely to be a general feature, as most cancers are believed to harbor only a minute number of CSCs. Consequently, several important conclusions can be drawn from the prognostic power of CSC signatures. First, CSC-associated expression profiles are clinically relevant as elements of stemness influence the clinical outcome of various cancers. Second, the association of CSC-
derived profiles with prognosis should be interpreted with caution. This is well-illustrated for CRC in which a mouse intestinal stem cell (ISC)-derived signature was shown to associate strongly with patients at a high risk of relapse (115), suggesting that stem cell content is important. However, our data show that the presence of ISC-related genes in both pure ISC or colon CSC-derived gene signatures is not a crucial determinant of the prognostic power of these signatures (42). Importantly, in the illustrated case, ISC and CSC signatures pointed to a distinct, more immature and poorly differentiated subset, and thereby reflected a clonal trait of the malignant tissue, rather than CSC number.

Connecting inter- and intra-tumor heterogeneity. As described above, both forms of heterogeneity influence clinical outcome in many ways. The mere presence of intra-tumoral heterogeneity impinges on adequate diagnosis and results frequently in therapy resistance. Indeed, it has been shown that high levels of intra-tumor heterogeneity correlate with poor prognosis in glioma (98). Moreover, the realization that distinct molecular subsets exist requires a shift in cancer drug development, from a ‘one-size fits all’ chemotherapy to a more personalized and group-based drug design, or even an approach directed towards specific tumor clones (52). Unfortunately, our crude understanding of the molecular mechanisms by which intra-tumor heterogeneity is driven, coupled with uncertainty about the diversity of existing molecular subtypes, might preclude the development of significantly improved therapies, particularly for more advanced metastatic disease.

Can one find analogy between these two forms of heterogeneity that could be exploited to improve rational therapy design? In other words, could a better appreciation of the biological foundation of distinct subtypes be used to infer the behavior of intra-tumor heterogeneity? A study of acquired resistance to EGFR inhibitors in lung cancer revealed that recurrent tumors in a fraction of patients undergo a subtype shift from non-small cell to small cell carcinoma (117). The acquisition of the small cell lineage is marked by the acquisition of traits specific for this particular lineage, including sensitivity to conventional chemotherapy. Although the authors conclude that a lineage switch occurred by an unknown mechanism, it is equally feasible that robust subsets of lung cancer were pre-existing within these tumors and contracted or expanded during the therapy. Although this observation remains anecdotal, the presence of a heterogeneous tumor cell population in a tumor could be a reflection of the diverse molecular subtypes that belong to a malignancy (Fig 3).
Conclusion and outlook

Tumors are recognized as heterogeneous entities but the crucial question remains of how to use that knowledge and turn it into a clinical benefit. For instance, heterogeneity in Barrett esophagus, measured by several indexes initially developed for ecological studies, is an important predictor of progression to esophageal carcinoma; a similar study performed in breast tissue suggests that the degree of heterogeneity could be a general predictor for disease progression (118). Whether the presence of more clones is crucial for progression or simply the reflection of genomically unstable tumors remains to be established. Ingeniously designed, targeted therapeutic agents are tested in preclinical models at tremendous speed but at a high cost as well. Tumor heterogeneity reflects the biggest challenge for drug development – single agents that target known aberrations to which cancer cells might be addicted – such as BCR–ABL in CML – might simply fail due to the presence of distinct non-targeted clones. Moreover, a better appreciation of inter-tumor heterogeneity and how this might affect tumor response is needed when drug responses are being evaluated. For example, ‘drug x’ might be discarded because it

Figure 3. Connecting intra- and inter-tumor heterogeneity. (Top) Multiple subclones exist in an individual lung tumor each being represented by a different subtype of lung cancer. (Bottom) Some subtypes of lung cancer are more sensitive to therapy than others. This is reflected in an individual tumor for which a therapy selects the most aggressive subclone, such as the subclone resembling the most aggressive cancer subtype.
Cancer heterogeneity – a multifaceted view.

does not produce clinical benefit in the overall population, although it might be potent in a small subset of the population given that the latter is represented adequately. This is not straightforward as the identification of subtypes is generally achieved on primary tumor-derived material, whereas phase II clinical trials are usually performed on late-stage cancers that might not necessarily encompass the same diversity of subtypes as those present in early stages. Assuming the latter is true, how do we progress towards more effective treatments? One of the limiting steps here is having appropriate models to use to tackle these questions. Mouse models have been cleverly engineered to recapitulate crucial aspects of human tumor progression but, in particular, fail to reflect the heterogeneity that is present in human cancers, especially inter-tumorally. Human xenograft models are superior in that respect as they are derived directly from patients and can be maintained almost indefinitely in immunocompromised mice. The main criticism here rests on the lack of interaction with the microenvironment and the selection of tumor cells and clones that are more apt to survive in the mouse environment. Moreover, drug screening in this setting is cumbersome. A potential alternative is to fall back on established and well-characterized cancer cell lines. Studies have begun to elucidate drug response in large panels of cell lines (119), but these lack sufficient insight into the extent to which these represent the cancer population. Based on the unbiased identification of distinct subtypes of primary colon cancers, we have identified subtypes of colon cancer cell lines that have important biological attributes, such as invasive properties and therapy resistance, corresponding with primary cancers (30). Research in this area will be facilitated by a wealth of data such as genotype, drug response, and expression profiles that are available for a wide range of cell lines from various lineages. Although clearly imperfect, these simple in vitro models might be the first crucial step to understanding the different biological backgrounds that are hard-wired into distinct cancer entities.

Acknowledgements
We thank Maarten Bijlsma for critically reading the manuscript and for useful comments.

Footnotes
The authors declare that they have no conflict of interest.
Chapter 1

References

36. I. Verbrugge, R. W. Johnstone, M. Bots, Promises and challenges of anticancer drugs that


76. E. Becht et al., Immune and stromal classification of colorectal cancer is associated with molecular subtypes and relevant for precision immunotherapy. Clinical cancer research, (2016).
79. E. Fessler et al., TGFβ signaling directs serrated adenomas to the mesenchymal colorectal cancer subtype. EMBO Mol Med 8, 745-760 (2016).
85. P. Gibson et al., Subtypes of medulloblastoma have distinct developmental origins. Nature 468, 1095-1099 (2010).
89. G. Molynieux et al., BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. Cell stem cell 7, 403-417 (2010).
102. A. Prabhallad et al., Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. Nature 483, 100-103 (2012).
103. C. Roche-Lestienne et al., Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. Blood 100, 1014-1018 (2002).


