Published in final edited form as:

Biochim Biophys Acta. 2017 April ; 1867(2): 69–83. doi:10.1016/j.bbcan.2016.11.005.

Evolutionary scalpels for dissecting tumor ecosystems

Daniel I. S. Rosenbloom^{1,2,*}, Pablo G. Camara^{1,2}, Tim Chu^{1,2}, and Raul Rabadan^{1,2}

¹Department of Systems Biology, Columbia University College of Physicians and Surgeons, 1130 St. Nicholas Avenue, New York, NY 10032, USA

²Department of Biomedical Informatics, Columbia University College of Physicians and Surgeons, 1130 St. Nicholas Avenue, New York, NY 10032, USA

Abstract

Amidst the growing literature on cancer genomics and intratumor heterogeneity, essential principles in evolutionary biology recur time and time again. Here we use these principles to guide the reader through major advances in cancer research, highlighting issues of “hit hard, hit early” treatment strategies, drug resistance, and metastasis. We distinguish between two frameworks for understanding heterogeneous tumors, both of which can inform treatment strategies: (1) The tumor as diverse *ecosystem*, a Darwinian population of sometimes-competing, sometimes-cooperating cells; (2) The tumor as tightly integrated, self-regulating *organ*, which may hijack developmental signals to restore functional heterogeneity after treatment. While the first framework dominates literature on cancer evolution, the second framework enjoys support as well. Throughout this review, we illustrate how mathematical models inform understanding of tumor progression and treatment outcomes. Connecting models to genomic data faces computational and technical hurdles, but high-throughput single-cell technologies show promise to clear these hurdles.

Keywords

cancer evolution; intratumor heterogeneity; combination therapy; drug resistance; genomics; mathematical modeling

1. Introduction

Medical science has long known tumors to consist of multiple different parts¹. Before the advent of modern genome sequencing, researchers routinely observed that different parts of a tumor could vary in cell morphology, karyotype, immunogenicity, and drug sensitivity², reported in frequent publications about tumor heterogeneity³. More recently, dramatic advances in genomics, transcriptomics, and technologies to isolate and characterize single

* dsr2131@columbia.edu.

Conflicts of interest:

The authors declare no conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

cells have enabled new quantitative understandings of heterogeneity. Large-scale studies have revealed both that each tumor bears a diverse complement of somatic mutations, some occurring at very low frequencies (intratumor heterogeneity), and that the genomic make-up of tumors can vary in clinically important ways across patients suffering from the same cancer type (interpatient heterogeneity)^{4,5}.

The field of precision medicine attempts to address interpatient heterogeneity by identifying drug targets specific to each patient, based on personalized molecular information. Yet technologies for precision medicine make abundantly clear that tumors are genetically and epigenetically complex entities that cannot be reduced to a mere list of mutations. Observations both *in vivo* and *in vitro* reveal that tumors are dynamic, heterogeneous populations of cells, a symbiotic biological community, which may respond in unexpected ways to therapy⁶⁻¹⁰. If therapy is to effectively take personalized genomic information into account, it must be informed by evolutionary principles. In this review, we will explore these principles and connect them to clinical findings.

Many previous reviews have discussed how heterogeneous tumors evolve^{2,3,11-16}, highlighting in particular the complexity of tumor ecosystems^{17,18}, the relationship between tumor heterogeneity and progression to metastasis¹⁹⁻²¹, and the importance of considering tumor evolution when designing therapy^{6-10,21}. Many computational methods have been designed to grapple with the onslaught of data produced by genomic²² and epigenomic²³ studies. New mathematical models are being invented to interpret the output of these computational methods, to understand the evolutionary processes underlying tumor progression and response to treatment^{24,25}. Amidst the growing literature, essential principles in evolutionary biology recur time and time again. This review offers an exposition of these principles, highlighting clinical scenarios. Our perspective is shaped deeply by mathematical and computational methods, our “scalpels” for dissecting the evolutionary history of heterogeneous tumors. In this review, we present the cutting edge of these scalpels without algebraic notation.

We begin by confronting the problem of drug resistance (Sections 2 and 3), reviewing the arguments, modeling frameworks, and clinical findings that have contributed to current understanding of evolution of tumors to evade targeted therapy and progress to metastasis. We will describe a predominant school of thought, based on population genetic modeling, that champions hard-hitting administration of multiple therapies early in disease progression to prevent the evolution of resistance^{24,26,27}. This strategy has seen outstanding success in some areas, particularly in transforming childhood acute lymphoblastic leukemia from an immediately fatal diagnosis to one in which most patients are cured²⁸. Some research, however, counsels caution in the enthusiastic extrapolation of this principle to all cancer types, suggesting that sometimes the best we may hope for is long-term restriction of a tumor to a small, nonthreatening volume^{7,29}.

As we proceed, we will distinguish between two different concepts of intratumor heterogeneity: the tumor as an *ecosystem* (Section 4), versus the tumor as an *organ* (Section 5). While both of these terms have been used before to describe aspects of tumor biology^{8,30}, they have not been contrasted explicitly. These two concepts involve different

assumptions about how portions of a tumor may regrow following reduction via effective therapy. Considering the tumor as an ecosystem, regrowth of a tumor is an instance of the evolutionary phenomenon of adaptive radiation^{31,32}, in which mutant cells manage to explore and exploit underused ecological niches, in which they can proliferate. Considering the tumor instead as an organ, regrowth of a tumor is akin to anatomic regeneration or morphallaxis^{33,34}, in which cells are directed by biochemical signals to restore lost function. This distinction corresponds to a key difference long understood in evolutionary biology – adaptation by selection of inherited alterations that arose randomly^{35,36}, versus selection of intrinsic biological programs capable of producing or restoring a desired phenotype³⁷.

Throughout this review, we will use quantitative concepts that are the bread-and-butter of evolutionary biology – population sizes, fitness, mutation frequencies, and measures of diversity. Sometimes, modeling approaches treat these concepts as given, as if they were easily readable from the book of nature. In truth they are complex, have definitions that change from author to author and over time, and are difficult to measure. These challenges have motivated many computational and experimental advances (Section 6 and Box 1). We close with three open questions provoked by the research reviewed (Section 7).

2. Intrinsic versus acquired drug resistance

Poor prognosis of many cancer types most often results from the failure of therapy to clear a treatment-resistant portion of the tumor. Resistance is often classified as intrinsic or acquired, depending on whether the cancer fails to respond initially to treatment (intrinsic resistance), or recurs as a resistant tumor following initial response to therapy (acquired resistance). Genomic mechanisms conferring resistance to targeted therapy include amplification of genes targeted for inhibition or acquisition of secondary (compensatory) mutations, and nongenomic mechanisms include activation of kinases restoring function of oncogenic pathways or transformation of cells to drug-tolerant stem cell phenotypes³⁸.

Intrinsic drug resistance is observed when a large portion of the initial tumor expresses a drug-resistant phenotype. In this case, the battle is over before it has begun, and a different treatment option must be pursued. Acquired drug resistance, on the other hand, proceeds by replacement of a population of drug-sensitive cells – which succumb to therapy – with a separate population of drug-resistant cells. The difference between intrinsic and acquired resistance can be understood in terms of the fitness costs and benefits of drug resistance³⁹⁻⁴². If a mutation that occurs within a growing untreated tumor confers drug resistance without incurring a fitness cost (Fig. 1A), it may or may not be present at high frequencies in the untreated tumor. Like any selectively neutral mutation occurring in a growing population, its frequency depends mainly on the time of its appearance: Cost-free mutations that occur early, when the tumor is small, are more likely to reach detectable levels before treatment, simply because they have more opportunity to grow, compared to equivalent mutations that occur later^{43,44}. The characterization of resistance as “intrinsic” or “acquired” may simply depend on this timing (Fig. 1C versus Fig 1D). This scenario may play out in the setting of relapsing childhood acute lymphoblastic leukemia (ALL), in which RAS-MAPK-activating mutations are associated to a high-risk of relapse. These mutations present a dual role, conferring resistance to the commonly used chemotherapeutic agent methotrexate and

sensitivity to vincristine, another common chemotherapeutic. Cell viability assays of methotrexate resistance, performed on two such mutations found in *KRAS*, revealed fitness diagrams similar to that in Fig. 1A, with no discernible fitness difference between mutated and control cells in the absence of drug⁴⁵.

If a mutation that confers drug resistance also incurs a fitness cost relative to the drug-sensitive clone in which it occurs (Fig. 1B), then it is less likely to reach substantial frequencies in the untreated tumor. As a result, such mutations would typically be classified as leading to acquired, but not intrinsic, resistance (Fig. 1E). In the ALL cohort described above, six different mutations in the gene *NT5C2* that may enhance cellular clearance of cytotoxic metabolites of chemotherapies were observed at relapse in 20% of patients with T-cell precursor ALL, but could not be detected by targeted deep-sequencing of the untreated samples⁴⁵. Other characteristically “late” driver mutations not often seen in untreated tumors – such as deletion of chromosome arm 17p and certain mutations in *TP53* in chronic lymphoid leukemia⁴⁶, and inactivating mutations of the mismatch repair pathway in glioblastoma⁴⁷ – may fit this fitness profile. In the latter case, inactivation of mismatch repair causes dramatic increases in somatic mutation rates, harming a cell’s ability to leave viable offspring⁴⁸. It is possible that small subclones with mismatch repair defects have developed other mechanisms to compensate for widespread DNA damage, granting them an advantage when subjected to treatment with the alkylating agent temozolomide.

2.1. Absolute versus relative fitness and the “mutant selection window”

We have been speaking of “fitness” in general terms so far, but there are many ways to define it, involving measures of expected offspring number, viability, net growth rate, or even complex lifecycles⁴⁹⁻⁵¹. Fitness is commonly described in *absolute* terms, as the actual survival probability or reproductive rate of an individual cell or clone, as well as in *relative* terms, measuring one particular clone against another. In the context of chemotherapy, it is imperative to keep these two concepts straight. Drug-resistant mutations are often defined in *relative* terms: does the mutation confer a fitness benefit, relative to an original clone that does not carry the mutation? Under treatment, the *fraction* of a tumor carrying a drug-resistant mutation is expected to grow, but this alone says almost nothing about the overall size of the tumor or its long-term prospects for metastasis. If the fitness cost incurred by the resistant clone was large, the resulting drug-resistant tumor may grow slowly or even fail to grow at all. In that case, chemotherapy would benefit the patient, even though it selects for an extremely (relatively) fit mutant (Fig. 1F). In the setting of antimicrobial drug resistance, this concept has been explored extensively, resulting in definition of a “mutant selection window,” or concentration of drug in which a drug-resistant clone both outperforms the sensitive wild type (is *relatively* fit) and still manages to grow (positive *absolute* fitness)³⁹⁻⁴². To optimize therapy, drug concentrations within this window should be avoided.

The schematics in Fig. 1 suggest that increasing the *relative* fitness of a resistant mutant compared to its drug-sensitive competitors, while not inherently desirable, need not lead to poor outcomes, if absolute fitness is below the threshold for tumor viability. Of course, tumors are far more diverse entities than suggested by simple diagrams of competition

between two subclones. The next section elaborates on this simple picture by considering resistance to multiple drugs.

3. Combination therapy: can evolutionary principles offer a new hope?

The extreme genetic diversity present in a large tumor means that it is nearly inevitable for some cells to harbor alterations conferring resistance to therapy^{52,53}. In glioblastoma (GBM), genomic amplifications commonly occur in multiple signaling pathways, suggesting that therapeutic targeting of only one pathway will merely allow growth of cellular populations that exploit an alternate pathway⁵⁴. In colorectal cancers initially sensitive to anti-EGFR antibody therapy, rapid appearance of drug-resistant disease following complete response to therapy implies that small drug-resistant subclones existed before treatment⁵⁵, and deep sequencing can detect rare resistance-conferring mutations in *KRAS* up to 10 months before treatment⁵⁶.

Combination drug therapy is one strategy for combatting inevitable resistance mutations. This strategy has been particularly successful in the treatment of infectious diseases such as HIV. If multiple therapeutic agents are used, requiring multiple separate mutations in order to confer resistance to the entire regimen, then the probability is low that all required mutations will occur in a single cancer cell (Figure 2A). A mathematical modeling analysis suggests that, even when failure of monotherapy is nearly certain, a dual therapy targeting two different genomic mechanisms (with no single mutation causing cross-resistance) may furnish patients with >70% probability of treatment success. The precise probability of success falls both with the somatic mutation rate and the size of the untreated tumor (Figure 2B)⁵⁷.

The model described above considers an idealized form of combination therapy – with effective doses of both drugs reaching all tumor cells. The design of actual therapy must account for pharmacokinetics and pharmacodynamics, as well as delivery to tumor sites. Pharmacokinetic gradients can hasten the evolution of drug resistance, as regions of intermediate drug penetration serve as “breeding grounds” for the acquisition of stronger drug-resistant mutants that may grow even at high concentrations⁵⁸⁻⁶⁰. Mathematical modeling suggests that when multiple drugs, with distinct pharmacokinetic and biopenetration profiles, are administered, this problem becomes more acute⁶¹: Regions subject to penetration by only one drug can serve to breed a large population of cancer cells with resistance to that drug. Subsequent accrual of genetic diversity in this singly-resistant compartment can serve as the raw material fueling acquisition of resistance to a second drug. In this manner, discordant drug penetration can trigger evolution of a multiple drug-resistant mutant that otherwise never would have evolved (Figure 2C).

Considering the relative ease with which resistance to single-drug regimens can emerge, researchers have often counseled a strategy of early treatment with large doses of multiple therapeutic agents. Sequential administration of therapy may “train” cancer cells to become resistant to each therapy in turn^{57,62}. If simultaneous treatment is infeasible due to toxicity, then, rapidly alternating treatment with non-cross-resistant agents is also proposed to prevent development of resistance^{26,27}. Evidence favoring alternating over sequential therapies is

generally mixed, however, perhaps owing to the many attempted permutations of therapy as well as the difficulty of managing side effects in an alternating therapy setting⁶³. In cases where the mechanism conferring resistance to one therapy confers sensitivity to co-administered therapy, a tumor may be stuck in an evolutionary “double-bind,” blocking the path to effective drug resistance^{6,45,64}. Understanding the tumor’s fitness costs and benefits of single- and multiple-drug resistance is therefore essential for design of effective combination therapies.

Treatment of childhood acute lymphoblastic leukemia (ALL) by combination therapy is one of the major breakthroughs in pediatric medicine in the late 20th century²⁸. After a period in which single chemotherapies and palliation were dominant, a landmark study in the 1960s demonstrated the superiority of simultaneous high doses of several chemotherapeutic and immunosuppressive agents (including methotrexate and vincristine, discussed above), administered as different combinations during induction, intensification, and maintenance treatment phases^{65,66}. Further studies supported the “hit hard, hit early” treatment concept⁶⁷ and introduced immunophenotypic and genomic risk subtyping to customize treatment. These advances have transformed pediatric ALL from an immediately fatal diagnosis to one in which over 80% of patients are cured^{68,69}.

Looking ahead, a promising line of research is putting gene-targeted combination therapy to the test in gastrointestinal stromal tumors (GIST) by examining the performance of the tyrosine kinase inhibitor imatinib together with an agent inhibiting the PI3K/AKT signal cascade. In both three-dimensional culture and xenografts of GIST cells, combination therapy manages to control tumor growth and improve survival beyond what was achieved by either drug alone⁷⁰. This work built on previous studies showing efficacy of imatinib when combined with one of several PI3K/AKT inhibitors, in xenografted tumors. In those cases, sensitivity to each drug combination depended on the tumor’s oncogenic genotype, suggesting a way forward for personalizing regimens^{71,72}.

Some care must be taken in extrapolating from these promising preclinical results to conclusions about the evolution of resistance to dual therapy in GIST. A majority of patients treated with tyrosine kinase inhibitors alone develop mutations in the tyrosine kinase *KIT* conferring drug resistance⁷³; in fact, multiple different mutations often occur in a single patient, each within a separate lesion⁷⁴. This observation suggests that the population of cells with drug-resistant *KIT* may be large and diverse in these patients prior to therapy. Where a great many cells already have one resistance mutation, the tumor may seize the opportunity to evolve resistance to a second drug, initiating a clone that is resistant to both. It is unclear whether the treatment periods of two weeks to four months in the above experiments would have been sufficient time for the appearance of resistance to dual therapy. Furthermore, while inhibition of KIT blocks activation of the downstream PI3K/AKT signaling pathway in imatinib-sensitive tumors, overexpression⁷⁵ or KIT-independent activation⁷⁶ of components in this pathway can contribute to imatinib resistance. Understanding interactions between inhibitors of KIT and of the PI3K/AKT pathway, as well as the (possibly overlapping!) mechanisms conferring resistance to each, is critical to designing drug combinations that will not elicit cross-resistance.

3.1. Phenotypic heterogeneity and combination therapy

Intratumor heterogeneity of protein signaling networks can also contribute to drug resistance, even where no genomic cause is evident. Using a patient-derived xenograft mouse model of glioblastoma multiforme (GBM), Wei et al. examined tumor response to an mTOR kinase inhibitor^{77,78}. Using single-cell characterization of phosphoprotein levels, they found that response to treatment – which delayed tumor growth – diminished the role of mTOR complexes in protein signaling, while increasing the tumor's reliance on ERK and Src signaling to compensate for inhibition. The authors suggest that this compensation occurs by rapid “rewiring” of protein networks in individual cells, a particular case of the more general phenomenon of robustness of cellular functions to genetic variation or environmental insult⁷⁹⁻⁸¹. Another possible mechanism for the rapid population-level change is a “functional bottleneck” that kills off cells failing to express the compensatory phenotype, analogous to a genetic bottleneck. Whatever the underlying mechanism – adaptation by individual cells, or Darwinian selection at the population-level – this work underscores the importance of targeting multiple pathways simultaneously, to minimize the chance that a cell takes advantage of an un-targeted pathway.

3.2. The tumor strikes back: Can cancer cells “cooperate” to evade combination therapy?

The above framework conceives of resistance to therapy as a property of individual cancer cells – the presence or absence of particular mutations, genetic pathways, or protein functions enabling a cell and its descendants to survive therapy. This framework meshes well with common theoretical models in evolutionary biology and population genetics, and it informs interventions to prevent emergence of multiple drug-resistant cells. This framework has been extremely useful in the combatting of microbial infections such as tuberculosis and HIV⁸². It may, however, discount the role of tumor cells as participants in a complex ecosystem, in which heterogeneity of function and cooperation between cells is essential. The “societal” relationships between cancer subpopulations has long been recognized as contributing to tumor viability, metastasis, and drug resistance^{3,17}.

Particularly worrisome, cooperation between distinct tumor subclones may boost chemoresistance. In a colorectal cancer cell model, exposure to oxaliplatin selected for resistance via production of soluble cell factors. Introduction of these resistant cells into chemosensitive cultures enhanced growth of the chemosensitive population in the presence of treatment, suggesting that production of resistance-mediating factors can benefit the entire tumor⁸³. Other experiments have demonstrated that factors transported between cells can contribute to resistance in models of chronic myeloid leukemia⁸⁴, myeloma^{85,86}, pancreatic cancer⁸⁷, and other cancers (reviewed in⁸⁸⁻⁹⁰). Recall that the evolutionary argument for combination therapy hinges on the low probability that multiple required drug-resistance mutations would occur *in the same cell*, in the limited time before therapy eliminates that cell. If resistance mutations can occur in multiple different cells and yet confer multiple drug-resistance upon the entire tumor, this optimistic view of combination therapy is on weaker ground (Figure 2D).

4. Adaptive radiation and ecological niches: Understanding the cancer ecosystem

Interactions between cancer cells extend far beyond the mechanisms of drug resistance, and investigation has revealed a rich cancer ecosystem. A beautiful study of evolution of heterogeneous tumors 35 years ago demonstrates how the concept of an adaptive radiation, while originally used to describe the origin of animal species specializing in different niches, applies to cancer ecosystems as well⁹¹. In melanoma tumor cells, metastatic capability was measured by number of lung metastases formed following intravenous injection of dissociated cells in mice. In heterogeneous populations of cancer cells, the distribution of metastatic capability remained stable through both *in vitro* passage and *in vivo* transfer of xenografts. However, when homogeneous stock derived from single-cell dilution was passaged, the distribution in this trait drifted substantially. Stability of the phenotype distribution appeared to be regulated by ecological interactions between the different subpopulations present in heterogeneous culture. When separate homogeneous cultures, distinguishable by drug sensitivity, were mixed, each maintained its own trait distribution over 20 passages, suggesting that each subpopulation fills a separate “niche” in the tumor ecosystem. Only after a treatment was applied to eradicate some of the subpopulations did the remaining cells expand their range of metastatic abilities to fill the “gap” left by the extinct subpopulation (Figure 3A,B). This behavior strongly suggests a phenomenon analogous to adaptive radiations following mass extinction events, during which phenotypic evolution accelerates to generate diversity and explore newly available niches^{3,32}. In many species, convergent adaptive radiations are observed, with a similar suite of traits recurring in ecologically similar habitats, whether in the field⁹² or in the lab⁹³. This pattern suggests that, despite the underlying randomness of mutation, and despite distinct genetic mechanisms that may be at play in each habitat, natural selection can act deterministically (Figure 3C)³¹.

In the decades since these observations were made, the genomic and other molecular aspects underlying metastatic capability have been explored. The mechanisms underlying epithelial-mesenchymal transition (EMT), a hallmark of metastatic transition, are hotly pursued, and this research has led to understanding of the roles that the Ras signaling pathway, receptor tyrosine kinases, and cadherin downregulation have in destabilizing cellular junctions and enabling cell motility⁹⁴. Within this vast literature, there is a growing appreciation for how the molecular and biophysical mechanisms promoting metastasis can be understood in an ecological setting. In a number of model systems, interactions between subclones, which separately do not exhibit metastatic potential, can induce EMT and metastases when introduced together to the same animal^{20,21}. These observations indicate a kind of “division of labor” involved in development of metastasis, with some cells destabilizing extracellular matrix, others recruiting blood vessels, and others taking advantage of these conditions by detaching from the tumor mass and entering the circulation¹⁸.

Detailed understanding of these pro-metastatic interactions is possible at the level of cell morphology and gene expression. In a mouse model of small cell lung cancer, a minority of cells were consistently found in which oncogenic Ras activation induced dedifferentiation to

a mesenchymal phenotype, although the majority of each tumor still expressed markers of neuroendocrine differentiation, a characteristic of this type of cancer. It was observed, furthermore, that the balance of neuroendocrine to mesenchymal cells in an *in vitro* system determined the gross morphology of the culture, with neuroendocrine cells attaching to a mesenchymal monolayer (if such cells are available) or floating free in the medium (if no mesenchymal cells are provided). Further experiments in the mouse model suggested that presence of both cell types within the same tumor site is required for establishment of distant metastases⁹⁵. A subsequent study implicated fibroblast growth factor 2 secretion by mesenchymal cells in enabling metastatic activity of the neuroendocrine population⁹⁶.

Although the metastatic capability of a single cell may be measured in isolation – and indeed was, as described above in Figure 3(A,B)⁹¹ – the presence of interaction effects may mean that metastatic potential is better described as a “group-level” or “community” trait, akin to human language, cooperative hunting behavior, or reproductive efficiency of insect colonies^{97,98}. Where future growth and expansion of a population depends on interactions, it is not always straightforward to define the evolutionary fitness of a single clone or genotype. Mathematical models of tumor evolution have begun to contend with this issue and may be able to predict long-term tumor dynamics from short-term measurements of growth and interaction. As a case in point, a study of breast cancer-derived xenografts found that heterogeneous tumors containing a subclone engineered to overexpress the cytokine IL11 outgrew others. This faster overall growth was not caused by expansion of the IL11-expressing subclone itself, but rather was due to the pro-growth effect of IL11 secretion into the microenvironment. Further investigation revealed that IL11 promoted remodeling of vasculature and extra-cellular matrix surrounding the tumor, allowing the otherwise indolent tumor to expand⁹⁹. This growth-promoting behavior can be described using mathematical models of frequency-dependent fitness, in which the growth rate of one subclone depends on the frequency of other subclones. By fitting such a model to observed growth rates, one can predict whether tumor heterogeneity will be preserved over the long term, beyond experimentally feasible timespans^{99,100} (Figure 4).

As the above experiments make clear, the simple fitness profiles in Figures 1A,B elide important mechanisms of tumor evolution. When different parts of a tumor interact, it is not possible to assign a single “fitness value” or “metastatic capability” to all cells of the same genotype. Treatment which shrinks the tumor can alter the tumor microenvironment or intratumor interactions, in ways not captured by these fitness profiles.

5. Morphallaxis and self-regulation: Is the tumor an organ?

The studies discussed above explore how different populations of cells interact, together determining the fate of the tumor. The fitness – or growth potential – of each type of cell depends on which other cell types are present in their local environment, a scenario that fits comfortably within frameworks established by theoretical evolutionary biologists to describe frequency-dependent fitness and spatially heterogeneous environments. There is a deeper sense, however, in which the fates of subclones may be intertwined, one which does not fit as readily into existing theory: Different parts of a tumor may be integrated tightly in a pattern of mutual control, such as in an organ.

In nearly all scenarios in evolutionary biology, mutations occur randomly, without regard to their consequences on organismal fitness. As the saying goes, “Mutation proposes, selection disposes”¹⁰¹. Error-prone mechanisms of DNA replication fit this description of randomness. The mechanisms proposed above to maintain tumor heterogeneity – adaptive radiation and cooperative interactions between subclones – are mechanisms of *selection*. IL11-secreting cells in the breast cancer model described above⁹⁹ did not “ask” neighboring cells to express other phenotypes, but they did alter the fitness landscape, enabling these other phenotypes to thrive.

Heritable epigenetic traits, however, may be regulated in a more directed manner. Of course, this is what happens in normal organism development, with signaling between differentiating cells generating stable heterogeneous phenotypes¹⁰². These stable development pathways evolved over the history of multicellular life, refined over 600 million years for functional organism coordination. But cancer is a kind of evolution gone awry, which does not persist past the life of the host. Without this multigenerational selection for functional coordination, is it possible that stable heterogeneity can be maintained by signals between subclones that enforce mutual control? If so, re-establishment of functional heterogeneity in a tumor following treatment (Fig. 3A,B) may not be a case of adaptive radiation at all, but rather the consequence of a paracrine signaling program hijacked to restore tumor function.

One manner in which functional heterogeneity can be generated is by proximity of different parts of the tumor to different environments. Numerous regulatory pathways involved in organism development can provide tumor environments that promote or restrict self-renewal, pluripotency, epithelial-mesenchymal plasticity, or other aspects of cancer cell fate. Pathway activity can depend on proximity to oxygen-rich regions or the invasive front of tumors³⁰ and can be manipulated experimentally, by introduction of transcription factors or tumor growth factors¹⁰³. In particular, the roles of transforming growth factor- β (TGF- β) signaling in controlling cell proliferation and apoptosis, and, when dysregulated, promoting tumor initiation and mesenchymal transition, has been an intense focus of study^{104,105}. Environmental gradients of TGF- β may establish niches that act as “sources” and “sinks” for mesenchymal cancer cells, enabling distant metastases¹⁰⁵. Even if a therapy smote all stemlike cancer cells, or all capable of metastatic outgrowth, the tumor environment might subsequently coax these cells to regenerate.

The situation may be similar to regeneration of amputated organs by morphallaxis in animals such as the hydra or certain oligochaete worms cells^{33,34}. Position-dependent signals in hydra cause neurons directly below an amputated head to convert to ganglion cells, restoring lost function¹⁰⁶ (Fig. 3D). Differentiated cells in animals capable of morphallaxis are highly plastic, evoking the plasticity observed in cancer cell function.

Rich microenvironmental cues drive heterogeneity and establish a complex interface between tumor and host³⁰. Little is known about the potential for tumors to self-regulate by manipulating these cues directly. If a tumor indeed behaves as a tightly integrated organ, it may hijack developmental pathways to maintain balance between its functional components. Brain tumors are particularly heterogeneous at an epigenetic level¹⁰³ and so may be a

natural place to pursue this question. Recent studies suggest mechanisms of endogenous control in glioblastoma. Brain tumor stem cells (BTSCs) isolated from patient-derived tumors may express proneural or mesenchymal phenotypes, with the latter subset promoting tumor regrowth after treatment^{107,108}. Transition from the proneural to the mesenchymal phenotype is stimulated by secretion of tumor necrosis factor- α (TNF- α) by macrophages¹⁰⁷, which themselves are recruited to the perivascular tumor environment by BTSCs¹⁰⁹. Thus an initially proneural BTSC population may manipulate the perivascular environment to trigger a portion of its own cells to transition to the mesenchymal type. In mixed culture of both BTSC types, spheroid colonies regularly arranged themselves as mesenchymal cores surrounded by proneural periphery, with unique sets of pro-oncogenic factors contained in extracellular vesicles produced by each cell type¹¹⁰. Once a mixed BTSC population is established, then, it potentially could be maintained by vesicular exchange of proteins altering tumor growth and spatial arrangement.

Taking a step back, this scenario of two BTSC types cooperating to promote tumor growth is similar to the experiment described in the previous section, in which populations of breast cancer-derived cells engineered to express different secreted factors could collaborate to enhance malignancy⁹⁹. Analysis of that experiment fit neatly into established theoretical frameworks of population genetics and frequency-dependent fitness. Yet the concept of a tumor as a tightly regulated organ, in which differentiated groups of cells work in concert to determine each other's structure and function, includes a layer of complexity not captured in these frameworks. "Hybrid" or "multiscale" mathematical models of tumor growth that capture both intracellular factors (e.g., nutrient level, cell age) and extracellular conditions (e.g., microenvironmental cues, neighboring cancer cells)^{25,111} may be particularly useful for integrating this concept with experimental data.

6. How can we measure intratumor heterogeneity "in the wild"?

So far, we have discussed experimental and theoretical approaches for understanding tumor evolution, progression, and drug resistance. We have taken as given researchers' ability to identify and quantify specific tumor subclones. In clinical settings, identifying genomic subclones involves statistical methods and substantial guesswork. It is therefore difficult to apply many population genetic approaches directly, as they presuppose the ability to genetically sequence individual organisms and identify subpopulations. Here we highlight some of the challenges and methods associated with identifying and measuring subclonal diversity.

6.1. Lessons from population genetic theory

When a population geneticist studies an evolving population, one of the first questions raised is, "How much genetic diversity is maintained by this population?" Two main components contribute to genetic diversity. First, the mutation rate – expected number of genetic mutations per individual replication event – generates diversity. This rate may be expressed per base pair, per entire sequenced region, or at any other convenient scale. If some mutations result in cell death or very low relative fitness, they may be excluded from the rate. Second, the "effective size" of a population determines its capacity to maintain the

diversity that was generated by mutation. Effective size is related to the total number of tumor cells, but may be thought of as excluding “evolutionary dead-ends”. For example, in a tumor where only a small number of very fit cells – or perhaps cancer stem cells – leave many generations of descendants, the effective size may be orders of magnitude smaller than the actual number of cells in the tumor. In populations of small effective size, mutations may quickly go extinct or quickly reach fixation (100% of the population), meaning that little diversity is maintained. Populations of large effective size, on the other hand, are able to maintain a large number of mutations at intermediate frequencies for long periods of time¹¹².

The product of these two components – mutation rate times effective population size – directly relates to the level of genetic diversity, or polymorphism, predicted to appear in a population sample¹¹³. Usually, population genetic samples are of individual organisms; common diversity measures include the number of polymorphic genomic sites or the average pairwise genetic diversity between sampled sequences¹¹³. To use these widespread metrics in the cancer context would require single-cell DNA sequencing. Yet this type of measurement is typically infeasible, and has been performed in only a handful of studies at a whole-exome level¹¹⁴⁻¹¹⁷. Even where single cells can be isolated, PCR error, allele drop-out, and poor amplification of some genomic regions can make both false positives (spurious mutations found in a minority of cells) and false negatives (undetected mutations) frequent and difficult to diagnose¹¹⁴, though methodological advances continue to push error rates down^{118,119}. The impact of this technological barrier is that a surprising amount of population genetic theory developed over the past century is not directly applicable to tumor genomic data. The above lesson, however, holds in general: levels of genetic diversity depend on a product of mutational processes (rate of non-lethal mutations) and demographic factors (population size, fitness, and stem cell structure)^{43,120}.

6.2. Ecological diversity measures

With single-cell technology in its infancy, the field makes do with bulk sequencing analysis for most purposes. As a result, methods have been developed to “computationally dissect” tumor populations and infer the presence of individual subclones, each with a unique genotype²². These methods assume a particular picture of tumor evolution: Each adaptive driver mutation in a tumor may lead to proliferation of the cell in which the mutation occurred. This cell is likely to have other mutations of minimal effect – “passengers” – distinguishing it from the surrounding tumor. As a result, when the cell proliferates, it initiates a clonal lineage containing several founder mutations not found elsewhere. If we assume that none of these mutations is reversed in this lineage nor occurs again in a different lineage (a fair assumption given the small fraction of genomic sites typically mutated), then all founder mutations share the same fate during throughout the life of the tumor^{22,121}. The key logical step in computational inference of subclones is this: If two or more somatic mutations appear in equal fractions of cells in every sample of the tumor, then these mutations are assumed to occur together in the same exact cells. In other words, it would be a strange coincidence for two mutations occurring in different subclones to have matching cellular frequencies in every sample. Clearly, the strength of this argument grows with the number of samples obtained.

Accurate computational inference of subclones asks much of noisy genomic data. Two mutations in the same subclone – present in precisely the same cells – may actually be sequenced in two very different fractions of reads, as PCR amplification of reads alters their observed frequency²². High depth of sequencing, of a large and highly pure tumor sample, with minimal amplification needed, is therefore ideal for minimizing error¹²². Even if we could rule out the noise contributed by sampling of amplified reads, each mutation may be either heterozygous or homozygous, meaning that the proportion of mutant reads coming from a cell that contains the mutation may be near 50% or 100%. Chromosomal amplifications and deletions can alter these proportions, as well. Trustworthy estimates of heterozygosity, copy number alterations, and sample purity are therefore important resources for subclonal inference, supplementing the raw somatic mutation data. A variety of computational tools have been developed to help researchers contend with these issues¹²¹⁻¹³⁰; many more are reviewed in ref. ²².

If a researcher is confident in the quality of subclonal inference performed on a tumor sample, this information can be used to characterize overall diversity in the tumor. A tumor containing many different subclones, of roughly equal size, carries more diversity than does a tumor with a single dominant clone. To quantify this notion of diversity, cancer biologists have adapted measurements commonly used by ecologists, such as the Shannon diversity index, Simpson's index, and others¹³¹. In one cohort, a high diversity index was found to predict progression of esophageal premalignancy to carcinoma^{132,133}. These measures may serve well alongside other more typical molecular biomarkers in stratifying patients by disease risk^{134,135}.

6.3. Single-cell genomics and transcriptomics

Moving beyond bulk sequencing, immunofluorescence in situ hybridization (iFISH) can be used to measure amplification or deletion of chromosomal regions within single cells¹³⁶ and thus identify genomic tumor subclones and the evolutionary relationships between them¹³⁷. This technique sidesteps the modeling assumptions involved in the subclonal inference discussed above, but of course cannot capture the magnitude of mutation data produced by sequencing of entire genomes or exomes. When iFISH was used to analyze 15 breast carcinomas, associations were found between high diversity and the extent of tumor spreading and necrosis. This pattern suggests that tumor size and oxygen restriction may prevent a single fast-growing subclone from “sweeping” through the entire tumor, wiping away diversity¹³⁸. Subsequent investigation found that diversity values calculated from iFISH were similar regardless of the chromosomal region probed, supporting more general use of this technique to interrogate subclonal architecture of breast cancer¹³⁹. More recently, high diversity of copy number alterations was observed to relate to poor prognosis in squamous cell carcinoma of the tongue¹⁴⁰.

iFISH provides a partial and low-throughput view of individual cancer cells' genomes. Single-cell sequencing has recently emerged as a high-throughput and more comprehensive tool for accessing the genome and transcriptome of hundreds or thousands of individual cells, offering deeper understanding of tumor heterogeneity and progression (Figure 5A)¹¹⁴. Methods for single-cell DNA sequencing^{117,141-143} dissect tumor clonal structure by

identifying both copy number alterations and single nucleotide variants, overcoming the limitations of bulk sequencing. In a recent study¹⁴⁴, copy-number profiling and exome sequencing of individual cells from two breast cancer patients allowed reconstruction of phylogenetic relationships between individual cells, revealing each tumor's unique evolutionary history (Figure 5B). This study found that aneuploidy rearrangements occurred early in tumor development, whereas point mutations occurred gradually.

Single-cell RNA sequencing methods¹⁴⁵⁻¹⁵⁰ provide a new perspective on tumor heterogeneity that considers epigenetic alterations of individual cells. These methods can identify groups of cancer cells with similar transcriptional profiles and can characterize cellular composition and transcriptional state of tumor-infiltrated cells. Notably, these methods have revealed that established prognostic markers based on bulk tumor expression can vary greatly among cells constituting the tumor. For instance, single-cell RNA sequencing of five primary glioblastoma tumors¹⁵¹ showed that each tumor consisted of cells expressing multiple glioblastoma subtypes¹⁵². In metastatic melanoma, two distinct transcriptional cell states were discovered¹⁵³ (Figure 5C), one of which is linked to intrinsic resistance to RAF and MEK inhibition¹⁵⁴⁻¹⁵⁶. The same study also investigated cellular composition of the melanoma tumor microenvironment (Figure 5D). Moreover, using this data, inter-patient differences in the level of T-cell exhaustion could be identified, highlighting the potential of these methods to inform future development of immunotherapies.

These results, along with other single-cell sequencing studies^{115,157-161}, show the great potential for this technology to uncover the genomic and transcriptomic diversity within tumors. Looking ahead, these methods may help to optimize therapeutic strategies by connecting individual cell properties to therapy response¹⁶²⁻¹⁶⁴, and they may enable noninvasive prognostics based on circulating tumor cell profiles¹⁶⁵⁻¹⁷⁰.

6.4. Spatial tumor heterogeneity

Another way to tease apart genetic heterogeneity is by analyzing spatially distinct samples of solid tumors. Driver mutations observed in 100% of cancer cells in one sample – and perhaps deemed good targets for therapy – can be subclonal, or even completely absent, in a different region of the tumor^{44,171-175}. Spatially distinct samples can be leveraged to infer aspects of a tumor's evolutionary history. In a study of 15 patients suffering from colorectal cancer, numerous mutations were observed to be subclonal in multiple glands sampled from the same patient (average of 23 glands per patient). Dubbed a “Big Bang” scenario (Figure 6A), this sharing of polymorphism across spatially separate sites suggests long-term maintenance of diversity initially established by early subclonal mutations⁴⁴.

Not all cancer types follow this scenario. If a mutation is clonal (present in all cancer cells) within a single sample, but completely absent from a spatially distant sample within the same patient, then it bears witness to a migration event producing physically separated lesions with distinct genetic fates. This scenario is reminiscent of speciation events following long-range migration or physical isolation of two (initially indistinguishable) subpopulations¹⁷⁶. In a study of 52 patients suffering from glioblastoma, 40% – 75% of mutations conformed to this pattern in patients who had samples taken from two physically

separated lesions. In contrast, for patients with two locally adjacent samples, only 14% – 34% of mutations conformed to this pattern¹⁷⁵. If we may further indulge in astrophysical metaphor: Each physically isolated lesion is akin to a separate “genetic universe” in a “multiverse” of different genetic possibilities (Figure 6B). The same glioblastoma study also proposed using Nei’s distance, a metric borrowed from population genetics¹⁷⁷, to quantify the difference between separate samples in a manner that fairly considers both clonal and subclonal mutations. Of particular clinical importance, cell lines initiated from genetically similar samples (measured by low Nei’s distance) responded similarly to *in vitro* treatment.

Identifying whether tumor evolution follows one scenario versus the other may provide information about evolutionary conditions in the ancestral tumor population. In a “Big Bang” scenario, polymorphisms are maintained in anatomically separate subpopulations. This scenario is similar to the phenomenon of incomplete lineage sorting, in which ancestral polymorphisms persist through a speciation event. This phenomenon explains why, for instance, the same polymorphism can exist within both the human species and the chimpanzee species^{178,179}. Incomplete lineage sorting is compatible with a large effective population size¹⁸⁰ and an absence of selective sweeps occurring after the mutation that gave rise to the shared polymorphism¹⁸¹. Conversely, in a “multiverse” scenario, polymorphisms are not shared between anatomical sites. This scenario suggests either that each site was founded by a very small population (i.e., a founder effect¹⁸²), or that selective sweep(s) occurred in the separate sites. It is important to note that the presence or absence of selective sweeps is *not* the same as the presence or absence of ongoing production of driver mutations. Fitness, demographic effects, cell motility, and the precise timing of a driver mutation interact to determine whether a driver mutation will sweep through an entire lesion — or hover at low frequencies, constrained by a crowded microenvironment^{44,183,184}.

7. Conclusions

Principles and models in evolutionary biology can help to unify and synthesize the avalanche of studies investigating intra-tumor heterogeneity. Our review leaves open three broad questions:

1. Is “hit hard, hit early” with combination drug therapy a feasible solution for cancers in general, or are there fundamental limits to this strategy, related to a tumor’s propensity to restore heterogeneity following environmental insult? Elegant and general mathematical models, experiments investigating single- and multiple-drug resistance, and the outstanding track records of combination therapy in treating pediatric leukemias and HIV, counsel strong emphasis on this strategy. When we consider how tumors maintain and exploit heterogeneity, though, we believe that careful evolutionary thinking will be needed to refine or alter this strategy in many cases.
2. Is modeling the tumor as an ecosystem of sometimes-competing, sometimes-cooperating components a useful approach for predicting disease progression and response to treatment? Since models that include cell-cell and cell-environment interaction can be complex, simplifying approaches will be needed to connect models to data. To what extent should dominant methods in population genetics

and evolutionary dynamics be modified to account for the possibility that tumors may self-regulate, like tightly integrated organs maintaining multiple differentiated functions? While we have introduced the distinction between “ecosystem thinking” and “organ thinking” as applied to tumors, we believe that further experiment and modeling are needed to clarify these concepts.

3. What technologies can grant us a detailed, single-cell view of intra-tumor heterogeneity, both at DNA and phenotypic levels? This view would advance understanding of the evolution of a tumor’s functionally distinct components. Bioinformatic techniques are available to handle the massive amounts of data that result from single-cell measurements, and theoretical approaches exist to understand population diversity and evolution at a single-cell level, but accurate measurement of single cells remains an extraordinary experimental challenge.

Acknowledgments

D.I.S.R. is supported by the US National Library of Medicine (T15 LM007079). P.G.C. thanks Arnold Levine, Suzanne Christen, and the Institute for Advanced Study for their hospitality. R.R. acknowledges funding from the NIH (U54 CA193313, R01 CA185486, and R01 CA179044).

References

1. Welch DR. Tumor Heterogeneity--A ‘Contemporary Concept’ Founded on Historical Insights and Predictions. *Cancer Res.* 2016; 76:4–6. [PubMed: 26729788]
2. Heppner GH, Miller BE. Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer Metastasis Rev.* 1983; 2:5–23. [PubMed: 6616442]
3. Heppner GH. Tumor heterogeneity. *Cancer Res.* 1984; 44:2259–2265. [PubMed: 6372991]
4. International Cancer Genome Consortium. International network of cancer genome projects. *Nature.* 2010; 464:993–998. [PubMed: 20393554]
5. Vogelstein B, et al. Cancer genome landscapes. *Science.* 2013; 339:1546–1558. [PubMed: 23539594]
6. Gatenby RA, Brown J, Vincent T. Lessons from applied ecology: cancer control using an evolutionary double bind. *Cancer Res.* 2009; 69:7499–7502. [PubMed: 19752088]
7. Gatenby RA. A change of strategy in the war on cancer. *Nature.* 2009; 459:508–509. [PubMed: 19478766]
8. Greaves M, Maley CC. Clonal evolution in cancer. *Nature.* 2012; 481:306–313. [PubMed: 22258609]
9. Aparicio S, Caldas C. The implications of clonal genome evolution for cancer medicine. *New England journal of medicine.* 2013; 368:842–851. [PubMed: 23445095]
10. Willyard C. Cancer therapy: an evolved approach. *Nature.* 2016; 532:166–168. [PubMed: 27075079]
11. Merlo LMF, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. *Nat Rev Cancer.* 2006; 6:924–935. [PubMed: 17109012]
12. Marusyk A, Polyak K. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta.* 2010; 1805:105–117. [PubMed: 19931353]
13. Swanton C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* 2012; 72:4875–4882. [PubMed: 23002210]
14. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer.* 2012; 12:323–334. [PubMed: 22513401]
15. Burrell RA, McGranahan N, Bartek J, Swanton C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature.* 2013; 501:338–345. [PubMed: 24048066]

16. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell*. 2015
17. Heppner GH, Shekhar M. Tumor heterogeneity is fundamental to the tumor ecosystem. *Oncology* (Williston Park, N Y). 2014; 28:780–781.
18. Tabassum DP, Polyak K. Tumorigenesis: it takes a village. *Nat Rev Cancer*. 2015; 15:473–483. [PubMed: 26156638]
19. Miller FR, Heppner GH. Cellular interactions in metastasis. *Cancer Metast Rev*. 1990; 9:21–34.
20. Neelakantan D, Drasin DJ, Ford HL. Intratumoral heterogeneity: Clonal cooperation in epithelial-to-mesenchymal transition and metastasis. *Cell Adh Migr*. 2015; 9:265–276. [PubMed: 25482627]
21. Zhou H, Neelakantan D, Ford HL. Clonal cooperativity in heterogenous cancers. *Seminars in Cell & Developmental* 2016; doi: 10.1016/j.semcdb.2016.08.028
22. Beerenwinkel N, Schwarz RF, Gerstung M, Markowitz F. Cancer evolution: mathematical models and computational inference. *Syst Biol*. 2015; 64:e1–25. [PubMed: 25293804]
23. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet*. 2015; 16:133–145. [PubMed: 25628217]
24. Bozic I, Allen B, Nowak MA. Dynamics of targeted cancer therapy. *Trends Mol Med*. 2012; 18:311–316. [PubMed: 22595628]
25. Altrock PM, Liu LL, Michor F. The mathematics of cancer: integrating quantitative models. *Nat Rev Cancer*. 2015
26. Goldie JH, Coldman AJ, Gudauskas GA. Rationale for the use of alternating non-cross-resistant chemotherapy. *Cancer Treat Rep*. 1982; 66:439–449. [PubMed: 7060033]
27. Grossi F, et al. Sequential, alternating, and maintenance/consolidation chemotherapy in advanced non-small cell lung cancer: a review of the literature. *Oncologist*. 2007; 12:451–464. [PubMed: 17470688]
28. Cheng TL, et al. Seven great achievements in pediatric research in the past 40 y. *Pediatr Res*. 2016; 80:330–337. [PubMed: 27556199]
29. Gatenby RA, Silva AS, Gillies RJ, Frieden BR. Adaptive Therapy. *Cancer Res*. 2009; 69:4894–4903. [PubMed: 19487300]
30. Egeblad M, Nakasone ES, Werb Z. Tumors as organs: complex tissues that interface with the entire organism. *Dev Cell*. 2010; 18:884–901. [PubMed: 20627072]
31. Losos JB. Adaptive Radiation, Ecological Opportunity, and Evolutionary Determinism. *The American Naturalist*. 2010; 175:623–639.
32. Losos JB, Mahler DL. Adaptive radiation: the interaction of ecological opportunity, adaptation, and speciation. *Evolution since Darwin: the first*. 2010
33. Drewes CD, Fournier CR. Morphallaxis in an aquatic oligochaete, *Lumbriculus variegatus*: reorganization of escape reflexes in regenerating body fragments. *Dev Biol*. 1990; 138:94–103. [PubMed: 2307291]
34. Agata K, Saito Y, Nakajima E. Unifying principles of regeneration I: Epimorphosis versus morphallaxis. *Dev Growth Differ*. 2007; 49:73–78. [PubMed: 17335428]
35. Luria SE, Delbrück M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics*. 1943; 28:491–511. [PubMed: 17247100]
36. Lenski RE, Slatkin M, Ayala FJ. Mutation and selection in bacterial populations: alternatives to the hypothesis of directed mutation. *Proc Natl Acad Sci U S A*. 1989; 86:2775–2778. [PubMed: 2704747]
37. Guerrero-Bosagna C, Skinner MK. Environmentally induced epigenetic transgenerational inheritance of phenotype and disease. *Mol Cell Endocrinol*. 2012; 354:3–8. [PubMed: 22020198]
38. Lackner MR, Wilson TR, Settleman J. Mechanisms of acquired resistance to targeted cancer therapies. *Future Oncol*. 2012; 8:999–1014. [PubMed: 22894672]
39. Drlica K. The mutant selection window and antimicrobial resistance. *J Antimicrob Chemother*. 2003; 52:11–17. [PubMed: 12805267]
40. Drlica K, Zhao X. Mutant selection window hypothesis updated. *Clin Infect Dis*. 2007; 44:681–688. [PubMed: 17278059]

41. Rosenbloom DIS, Hill AL, Rabi SA, Siliciano RF, Nowak MA. Antiretroviral dynamics determines HIV evolution and predicts therapy outcome. *Nat Med*. 2012; doi: 10.1038/nm.2892
42. Hill AL, Rosenbloom DIS, Nowak MA. Evolutionary dynamics of HIV at multiple spatial and temporal scales. *J Mol Med (Berl)*. 2012; 90:543–561. [PubMed: 22552382]
43. Bozic I, Nowak MA. Timing and heterogeneity of mutations associated with drug resistance in metastatic cancers. *Proc Natl Acad Sci USA*. 2014; 111:15964–15968. [PubMed: 25349424]
44. Sottoriva A, et al. A Big Bang model of human colorectal tumor growth. *Nat Genet*. 2015; 47:209–216. [PubMed: 25665006]
45. Oshima K, et al. Mutational landscape, clonal evolution patterns, and role of RAS mutations in relapsed acute lymphoblastic leukemia. *Proc Natl Acad Sci U S A*. 2016; doi: 10.1073/pnas.1608420113
46. Landau DA, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature*. 2015; 526:525–530. [PubMed: 26466571]
47. Wang J, et al. Clonal evolution of glioblastoma under therapy. *Nat Genet*. 2016; 48:768–776. [PubMed: 27270107]
48. Sniegowski PD, Gerrish PJ, Johnson T, Shaver A. The evolution of mutation rates: separating causes from consequences. *Bioessays*. 2000; 22:1057–1066. [PubMed: 11084621]
49. De Jong G. The fitness of fitness concepts and the description of natural selection. *Quarterly Review of Biology*. 1994
50. Orr HA. Fitness and its role in evolutionary genetics. *Nat Rev Genet*. 2009; 10:531–539. [PubMed: 19546856]
51. Hughes D, Andersson DI. Evolutionary consequences of drug resistance: shared principles across diverse targets and organisms. *Nat Rev Genet*. 2015; 16:459–471. [PubMed: 26149714]
52. Goldie JH, Coldman AJ. The genetic origin of drug resistance in neoplasms: implications for systemic therapy. *Cancer Res*. 1984; 44:3643–3653. [PubMed: 6744284]
53. Attolini CS-O, Michor F. Evolutionary theory of cancer. *Ann N Y Acad Sci*. 2009; 1168:23–51. [PubMed: 19566702]
54. Tanaka S, Louis DN, Curry WT, Batchelor TT, Dietrich J. Diagnostic and therapeutic avenues for glioblastoma: no longer a dead end? *Nat Rev Clin Oncol*. 2013; 10:14–26. [PubMed: 23183634]
55. Diaz LA, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature*. 2012; 486:537–540. [PubMed: 22722843]
56. Misale S, et al. Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature*. 2012; 486:532–536. [PubMed: 22722830]
57. Bozic I, et al. Evolutionary dynamics of cancer in response to targeted combination therapy. *Elife*. 2013; 2:e00747. [PubMed: 23805382]
58. Greulich P, Waclaw B, Allen RJ. Mutational pathway determines whether drug gradients accelerate evolution of drug-resistant cells. *Phys Rev Lett*. 2012; 109:088101. [PubMed: 23002776]
59. Wu A, et al. Cell motility and drug gradients in the emergence of resistance to chemotherapy. *Proc Natl Acad Sci USA*. 2013; 110:16103–16108. [PubMed: 24046372]
60. Kepler TB, Perelson AS. Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proc Natl Acad Sci U S A*. 1998; 95:11514–11519. [PubMed: 9751697]
61. Moreno-Gamez S, et al. Imperfect drug penetration leads to spatial monotherapy and rapid evolution of multidrug resistance. *Proc Natl Acad Sci USA*. 2015; 112:E2874–83. [PubMed: 26038564]
62. Bivona TG, Doebele RC. A framework for understanding and targeting residual disease in oncogene-driven solid cancers. *Nat Med*. 2016; 22:472–478. [PubMed: 27149220]
63. Michor F, Beal K. Improving cancer treatment via mathematical modeling: surmounting the challenges is worth the effort. *Cell*. 2015; 163:1059–1063. [PubMed: 26590416]
64. Silva AS, Gatenby RA. A theoretical quantitative model for evolution of cancer chemotherapy resistance. *Biol Direct*. 2010; 5:25. [PubMed: 20406443]
65. Aur RJ, et al. Central nervous system therapy and combination chemotherapy of childhood lymphocytic leukemia. *Blood*. 1971; 37:272–281. [PubMed: 4322483]

66. Pinkel D, Simone J, Hustu HO, Aur RJ. Nine years' experience with "total therapy" of childhood acute lymphocytic leukemia. *PEDIATRICS*. 1972; 50:246–251. [PubMed: 4505343]
67. Rivera GK, Abromowitch M, Mirro J, Ochs JS, Pui CH. Improved outcome in childhood acute lymphoblastic leukaemia with reinforced early treatment and rotational combination chemotherapy. *The Lancet*. 1991; 337:61–66.
68. Ma H, Sun H, Sun X. Survival improvement by decade of patients aged 0–14 years with acute lymphoblastic leukemia: a SEER analysis. *Scientific Reports*. 2014; 4:4227. [PubMed: 24572378]
69. Hudson MM, Link MP, Simone JV. Milestones in the curability of pediatric cancers. *J Clin Oncol*. 2014; 32:2391–2397. [PubMed: 24733794]
70. Zook P, et al. Combination of Imatinib Mesylate and AKT Inhibitor Provides Synergistic Effects in Preclinical Study of Gastrointestinal Stromal Tumor. *Clin Cancer Res*. 2016; doi: 10.1158/1078-0432.CCR-16-0529
71. Floris G, et al. A potent combination of the novel PI3K Inhibitor, GDC-0941, with imatinib in gastrointestinal stromal tumor xenografts: long-lasting responses after treatment withdrawal. *Clin Cancer Res*. 2013; 19:620–630. [PubMed: 23231951]
72. Van Looy T, et al. Phosphoinositide 3-kinase inhibitors combined with imatinib in patient-derived xenograft models of gastrointestinal stromal tumors: rationale and efficacy. *Clin Cancer Res*. 2014; 20:6071–6082. [PubMed: 25316817]
73. Khamisipour G, Jadidi-Niaragh F, Jahromi AS, Zandi K, Hojjat-Farsangi M. Mechanisms of tumor cell resistance to the current targeted-therapy agents. *Tumour Biol*. 2016; 37:10021–10039. [PubMed: 27155851]
74. Liegl B, et al. Heterogeneity of kinase inhibitor resistance mechanisms in GIST. *J Pathol*. 2008; 216:64–74. [PubMed: 18623623]
75. Lai S, et al. KIT over-expression by p55PIK-PI3K leads to Imatinib-resistance in patients with gastrointestinal stromal tumors. *Oncotarget*. 2016; 7:1367–1379. [PubMed: 26587973]
76. Li J, et al. PI3K/AKT/mTOR pathway is activated after imatinib secondary resistance in gastrointestinal stromal tumors (GISTs). *Med Oncol*. 2015; 32:111. [PubMed: 25757539]
77. Wei W, et al. Single-Cell Phosphoproteomics Resolves Adaptive Signaling Dynamics and Informs Targeted Combination Therapy in Glioblastoma. *Cancer Cell*. 2016; 29:563–573. [PubMed: 27070703]
78. Lam FC, Yaffe MB. Kicking Genomic Profiling to the Curb: How Re-wiring the Phosphoproteome Can Explain Treatment Resistance in Glioma. *Cancer Cell*. 2016; 29:435–436. [PubMed: 27070697]
79. Hartman JL, Garvik B, Hartwell L. Principles for the buffering of genetic variation. *Science*. 2001; 291:1001–1004. [PubMed: 11232561]
80. Stelling J, Sauer U, Szallasi Z, Doyle FJ, Doyle J. Robustness of cellular functions. *Cell*. 2004; 118:675–685. [PubMed: 15369668]
81. Kitano H. A robustness-based approach to systems-oriented drug design. *Nature Reviews Drug Discovery*. 2007; 6:202–210. [PubMed: 17318209]
82. Goldberg DE, Siliciano RF, Jacobs WR. Outwitting evolution: fighting drug-resistant TB, malaria, and HIV. *Cell*. 2012; 148:1271–1283. [PubMed: 22424234]
83. Bose D, et al. Chemoresistant colorectal cancer cells and cancer stem cells mediate growth and survival of bystander cells. *Br J Cancer*. 2011; 105:1759–1767. [PubMed: 22045189]
84. Bewry NN, et al. Stat3 contributes to resistance toward BCR-ABL inhibitors in a bone marrow microenvironment model of drug resistance. *Mol Cancer Ther*. 2008; 7:3169–3175. [PubMed: 18852120]
85. Shain KH, Landowski TH, Dalton WS. The tumor microenvironment as a determinant of cancer cell survival: a possible mechanism for de novo drug resistance. *Current Opinion in Oncology*. 2000; 12:557. [PubMed: 11085455]
86. Westendorf JJ, Lammert JM, Jelinek DF. Expression and function of Fas (APO-1/CD95) in patient myeloma cells and myeloma cell lines. *Blood*. 1995; 85:3566–3576. [PubMed: 7540067]
87. Zhang H, et al. Paracrine SDF-1 α signaling mediates the effects of PSCs on GEM chemoresistance through an IL-6 autocrine loop in pancreatic cancer cells. *Oncotarget*. 2015; 6:3085–3097. [PubMed: 25609203]

88. Li Z-W, Dalton WS. Tumor microenvironment and drug resistance in hematologic malignancies. *Blood Rev.* 2006; 20:333–342. [PubMed: 16920238]
89. Azmi AS, Bao B, Sarkar FH. Exosomes in cancer development, metastasis, and drug resistance: a comprehensive review. *Cancer Metastasis Rev.* 2013; 32:623–642. [PubMed: 23709120]
90. Jones VS, et al. Cytokines in cancer drug resistance: Cues to new therapeutic strategies. *Biochim Biophys Acta.* 2016; 1865:255–265. [PubMed: 26993403]
91. Poste G, Doll J, Fidler IJ. Interactions among clonal subpopulations affect stability of the metastatic phenotype in polyclonal populations of B16 melanoma cells. *Proc Natl Acad Sci U S A.* 1981; 78:6226–6230. [PubMed: 6947225]
92. Berner D, Salzburger W. The genomics of organismal diversification illuminated by adaptive radiations. *Trends Genet.* 2015; 31:491–499. [PubMed: 26259669]
93. Rainey PB, Travisano M. Adaptive radiation in a heterogeneous environment. *Nature.* 1998; 394:69–72. [PubMed: 9665128]
94. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol.* 2014; 15:178–196. [PubMed: 24556840]
95. Calbo J, et al. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. *Cancer Cell.* 2011; 19:244–256. [PubMed: 21316603]
96. Kwon M-C, et al. Paracrine signaling between tumor subclones of mouse SCLC: a critical role of ETS transcription factor Pea3 in facilitating metastasis. *Genes Dev.* 2015; 29:1587–1592. [PubMed: 26215568]
97. Smaldino PE. The cultural evolution of emergent group-level traits. *Behav Brain Sci.* 2014; 37:243–254. [PubMed: 24970399]
98. Olejarz JW, Allen B, Veller C, Gadagkar R, Nowak MA. Evolution of worker policing. *J Theor Biol.* 2016; 399:103–116. [PubMed: 26976051]
99. Marusyk A, et al. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature.* 2014; 514:54–58. [PubMed: 25079331]
100. Archetti M, Ferraro DA, Christofori G. Heterogeneity for IGF-II production maintained by public goods dynamics in neuroendocrine pancreatic cancer. *Proc Natl Acad Sci USA.* 2015; 112:1833–1838. [PubMed: 25624490]
101. Abouheif E, Rafiqi AM. Sex combs find middle ground in evolution debate. *Proc Natl Acad Sci USA.* 2014; 111:14011–14012. [PubMed: 25228768]
102. Cleaver O, Melton DA. Endothelial signaling during development. *Nat Med.* 2003; 9:661–668. [PubMed: 12778164]
103. Mack SC, Hubert CG, Miller TE, Taylor MD, Rich JN. An epigenetic gateway to brain tumor cell identity. *Nature Neuroscience.* 2016; 19:10–19. [PubMed: 26713744]
104. Ikushima H, Miyazono K. TGFbeta signalling: a complex web in cancer progression. *Nat Rev Cancer.* 2010; 10:415–424. [PubMed: 20495575]
105. Junk DJ, Cipriano R, Bryson BL, Gilmore HL, Jackson MW. Tumor microenvironmental signaling elicits epithelial-mesenchymal plasticity through cooperation with transforming genetic events. *Neoplasia.* 2013; 15:1100–1109. [PubMed: 24027434]
106. Koizumi O, Bode HR. Plasticity in the nervous system of adult hydra. III. Conversion of neurons to expression of a vasopressin-like immunoreactivity depends on axial location. *J Neurosci.* 1991; 11:2011–2020. [PubMed: 2066772]
107. Bhat KPL, et al. Mesenchymal Differentiation Mediated by NF- κ B Promotes Radiation Resistance in Glioblastoma. *Cancer Cell.* 2013; 24:331–346. [PubMed: 23993863]
108. Kim S-H, et al. Serine/Threonine Kinase MLK4 Determines Mesenchymal Identity in Glioma Stem Cells in an NF- κ B-dependent Manner. *Cancer Cell.* 2016; 29:201–213. [PubMed: 26859459]
109. Zhou W, et al. Periostin secreted by glioblastoma stem cells recruits M2 tumour-associated macrophages and promotes malignant growth. *Nat Cell Biol.* 2015; 17:170–182. [PubMed: 25580734]

110. Ricklefs F, et al. Extracellular Vesicles from High-Grade Glioma Exchange Diverse Pro-oncogenic Signals That Maintain Intratumoral Heterogeneity. *Cancer Res.* 2016; 76:2876–2881. [PubMed: 27013191]
111. Rejniak KA, Anderson ARA. Hybrid models of tumor growth. *Wiley Interdiscip Rev Syst Biol Med.* 2011; 3:115–125. [PubMed: 21064037]
112. Kouyos RD, Althaus CL, Bonhoeffer S. Stochastic or deterministic: what is the effective population size of HIV-1? *Trends Microbiol.* 2006; 14:507–511. [PubMed: 17049239]
113. Hartl, DL., Clark, AG. Principles of population genetics. Sinauer Associates Inc; 2006.
114. Navin NE. The first five years of single-cell cancer genomics and beyond. *Genome Research.* 2015; 25:1499–1507. [PubMed: 26430160]
115. Xu X, et al. Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell.* 2012; 148:886–895. [PubMed: 22385958]
116. Li Y, et al. Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. *GigaScience.* 2012; 1:69.
117. Hou Y, et al. Single-Cell Exome Sequencing and Monoclonal Evolution of a JAK2-Negative Myeloproliferative Neoplasm. *Cell.* 2012; 148:873–885. [PubMed: 22385957]
118. Voet T, Van Loo P. SNES makes sense? Single-cell exome sequencing evolves. *Genome Biol.* 2015; 16:86. [PubMed: 25926122]
119. Leung ML, Wang Y, Waters J, Navin NE. SNES: single nucleus exome sequencing. *Genome Biol.* 2015; 16:55. [PubMed: 25853327]
120. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. *Nat Genet.* 2016; 48:238–244. [PubMed: 26780609]
121. Deshwar AG, et al. PhyloWGS: reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol.* 2015; 16:35. [PubMed: 25786235]
122. Reiter JG, Makohon-Moore AP, Gerold JM, Bozic I. Reconstructing phylogenies of metastatic cancers. *bioRxiv.* 2016; doi: 10.1101/048157
123. Roth A, et al. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods.* 2014; 11:396–398. [PubMed: 24633410]
124. Miller CA, et al. SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput Biol.* 2014; 10:e1003665. [PubMed: 25102416]
125. Zare H, et al. Inferring Clonal Composition from Multiple Sections of a Breast Cancer. *PLoS Comput Biol.* 2014; 10:e1003703. [PubMed: 25010360]
126. Fischer A, Vázquez-García I, Illingworth CJR, Mustonen V. High-definition reconstruction of clonal composition in cancer. *Cell Rep.* 2014; 7:1740–1752. [PubMed: 24882004]
127. Ha G, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Research.* 2014; 24:1881–1893. [PubMed: 25060187]
128. Carter SL, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol.* 2012; 30:413–421. [PubMed: 22544022]
129. Greenman CD, et al. PICNIC: an algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics.* 2010; 11:164–175. [PubMed: 19837654]
130. Van Loo P, et al. Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci USA.* 2010; 107:16910–16915. [PubMed: 20837533]
131. Jost L. Entropy and diversity. *Oikos.* 2006; 113:363–375.
132. Maley CC, et al. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet.* 2006; 38:468–473. [PubMed: 16565718]
133. Merlo LMF, et al. A Comprehensive Survey of Clonal Diversity Measures in Barrett's Esophagus as Biomarkers of Progression to Esophageal Adenocarcinoma. *Cancer Prev Res.* 2010; 3:1388–1397.
134. Merlo LMF, Maley CC. The role of genetic diversity in cancer. *J Clin Invest.* 2010; 120:401–403. [PubMed: 20101092]
135. Gregson EM, Fitzgerald RC. Biomarkers for Dysplastic Barrett's: Ready for Prime Time? *World J Surg.* 2014; 39:568–577.

136. Hu M, et al. Regulation of In Situ to Invasive Breast Carcinoma Transition. *Cancer Cell*. 2008; 13:394–406. [PubMed: 18455123]
137. Gertz EM, et al. FISHtrees 3.0: Tumor Phylogenetics Using a Ploidy Probe. *PLoS ONE*. 2016; 11:e0158569. [PubMed: 27362268]
138. Park SY, Gönen M, Kim HJ, Michor F, Polyak K. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *J Clin Invest*. 2010; 120:636–644. [PubMed: 20101094]
139. Almendro V, et al. Genetic and phenotypic diversity in breast tumor metastases. *Cancer Res*. 2014; 74:1338–1348. [PubMed: 24448237]
140. Wangsa D, et al. Phylogenetic analysis of multiple FISH markers in oral tongue squamous cell carcinoma suggests that a diverse distribution of copy number changes is associated with poor prognosis. *Int J Cancer*. 2016; 138:98–109. [PubMed: 26175310]
141. Navin N, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011; 472:90–94. [PubMed: 21399628]
142. Zong C, Lu S, Chapman AR, Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science*. 2012; 338:1622–1626. [PubMed: 23258894]
143. Gawad C, Koh W, Quake SR. Single-cell genome sequencing: current state of the science. *Nat Rev Genet*. 2016; 17:175–188. [PubMed: 26806412]
144. Wang Y, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature*. 2014; 512:155–160. [PubMed: 25079324]
145. Tang F, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. 2009; 6:377–382. [PubMed: 19349980]
146. Islam S, et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Research*. 2011; 21:1160–1167. [PubMed: 21543516]
147. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. *Cell Rep*. 2012; 2:666–673. [PubMed: 22939981]
148. Ramsköld D, et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat Biotechnol*. 2012; 30:777–782. [PubMed: 22820318]
149. Picelli S, et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*. 2013; 10:1096–1098. [PubMed: 24056875]
150. Saliba A-E, Westermann AJ, Gorski SA, Vogel J. Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res*. 2014; 42:8845–8860. [PubMed: 25053837]
151. Patel AP, et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. 2014; 344:1396–1401. [PubMed: 24925914]
152. Verhaak RGW, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010; 17:98–110. [PubMed: 20129251]
153. Tirosh I, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. 2016; 352:189–196. [PubMed: 27124452]
154. Hoek KS, et al. In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Res*. 2008; 68:650–656. [PubMed: 18245463]
155. Müller J, et al. Low MITF/AXL ratio predicts early resistance to multiple targeted drugs in melanoma. *Nature Communications*. 2014; 5:5712.
156. Konieczkowski DJ, et al. A melanoma cell state distinction influences sensitivity to MAPK pathway inhibitors. *Cancer Discov*. 2014; 4:816–827. [PubMed: 24771846]
157. Hughes AEO, et al. Clonal architecture of secondary acute myeloid leukemia defined by single-cell sequencing. *PLoS Genet*. 2014; 10:e1004462. [PubMed: 25010716]
158. Yu C, et al. Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. *Cell Res*. 2014; 24:701–712. [PubMed: 24699064]
159. Francis JM, et al. EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. *Cancer Discov*. 2014; 4:956–971. [PubMed: 24893890]

160. Gawad C, Koh W, Quake SR. Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. *Proc Natl Acad Sci USA*. 2014; 111:17947–17952. [PubMed: 25425670]
161. Paguirigan AL, et al. Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. *Sci Transl Med*. 2015; 7:281re2–281re2.
162. Suzuki A, et al. Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment. *Genome Biol*. 2015; 16:66. [PubMed: 25887790]
163. Kim K-T, et al. Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. *Genome Biol*. 2015; 16:127. [PubMed: 26084335]
164. Kim K-T, et al. Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. *Genome Biol*. 2016; 17:80. [PubMed: 27139883]
165. Powell AA, et al. Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. *PLoS ONE*. 2012; 7:e33788. [PubMed: 22586443]
166. Heitzer E, et al. Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res*. 2013; 73:2965–2975. [PubMed: 23471846]
167. Ni X, et al. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci USA*. 2013; 110:21083–21088. [PubMed: 24324171]
168. Lohr JG, et al. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol*. 2014; 32:479–484. [PubMed: 24752078]
169. Ting DT, et al. Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. *Cell Rep*. 2014; 8:1905–1918. [PubMed: 25242334]
170. Aceto N, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*. 2014; 158:1110–1122. [PubMed: 25171411]
171. Yachida S, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*. 2010; 467:1114–1117. [PubMed: 20981102]
172. Sottoriva A, et al. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci USA*. 2013; 110:4009–4014. [PubMed: 23412337]
173. Kumar A, et al. Deep sequencing of multiple regions of glial tumors reveals spatial heterogeneity for mutations in clinically relevant genes. *Genome Biol*. 2014; 15:530. [PubMed: 25608559]
174. Yates LR, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med*. 2015; 21:751–759. [PubMed: 26099045]
175. Lee J-K, et al. Modeling genomic architecture informs targeted therapy in glioblastoma. In prep.
176. Mayr E. Ecological factors in speciation. *Evolution*. 1947; doi: 10.2307/2405327
177. Nei M. Genetic Distance between Populations. *The American Naturalist*. 1972; 106:283–292.
178. Leffler EM, et al. Multiple instances of ancient balancing selection shared between humans and chimpanzees. *Science*. 2013; 339:1578–1582. [PubMed: 23413192]
179. Azevedo L, Serrano C, Amorim A, Cooper DN. Trans-species polymorphism in humans and the great apes is generally maintained by balancing selection that modulates the host immune response. *Hum Genomics*. 2015; 9:21. [PubMed: 26337052]
180. Maddison WP, Knowles LL. Inferring phylogeny despite incomplete lineage sorting. *Syst Biol*. 2006; 55:21–30. [PubMed: 16507521]
181. Dutheil JY, Munch K, Nam K, Mailund T, Schierup MH. Strong Selective Sweeps on the X Chromosome in the Human-Chimpanzee Ancestor Explain Its Low Divergence. *PLoS Genet*. 2015; 11:e1005451. [PubMed: 26274919]
182. Provine WB. Ernst Mayr: Genetics and speciation. *Genetics*. 2004; 167:1041–1046. [PubMed: 15280221]
183. Reiter JG, Bozic I, Allen B, Chatterjee K, Nowak MA. The effect of one additional driver mutation on tumor progression. *Evol Appl*. 2013; 6:34–45. [PubMed: 23396615]
184. Waclaw B, et al. A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature*. 2015; 525:261–264. [PubMed: 26308893]

185. Landau DA, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013; 152:714–726. [PubMed: 23415222]
186. Burger JA, et al. Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nature Communications*. 2016; 7:11589.
187. Bozic I, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci USA*. 2010; 107:18545–18550. [PubMed: 20876136]
188. Godfrey-Smith, P. *Darwinian Populations and Natural Selection*. OUP; Oxford: 2009.
189. Wakeley J, Sargsyan O. Extensions of the coalescent effective population size. *Genetics*. 2009; 181:341–345. [PubMed: 19001293]
190. Song S, et al. qpure: A tool to estimate tumor cellularity from genome-wide single-nucleotide polymorphism profiles. *PLoS ONE*. 2012; 7:e45835. [PubMed: 23049875]
191. Larson NB, Fridley BL. PurBayes: estimating tumor cellularity and subclonality in next-generation sequencing data. *Bioinformatics*. 2013; 29:1888–1889. [PubMed: 23749958]
192. Yu Z, Li A, Wang M. CloneCNA: detecting subclonal somatic copy number alterations in heterogeneous tumor samples from whole-exome sequencing data. *BMC Bioinformatics*. 2016; 17:310. [PubMed: 27538789]
193. Wang B, Zhu J, Pierson E, Batzoglou S. Visualization and analysis of single-cell RNA-seq data by kernel-based similarity learning. *bioRxiv*. 2016; doi: 10.1101/052225
194. duVerle DA, Yotsukura S, Nomura S, Aburatani H, Tsuda K. CellTree: an R/bioconductor package to infer the hierarchical structure of cell populations from single-cell RNA-seq data. *BMC Bioinformatics*. 2016; 17:363. [PubMed: 27620863]
195. Jackman TR, Irschick DJ, de Queiroz K, Losos JB, Larson A. Molecular phylogenetic perspective on evolution of lizards of the *Anolis grahami* series. *J Exp Zool*. 2002; 294:1–16. [PubMed: 11932945]
196. Losos JB, Ricklefs RE. Adaptation and diversification on islands. *Nature*. 2009; 457:830–836. [PubMed: 19212401]

Box 1**How can we measure tumor evolution?**

Frameworks for reasoning about evolution, drug resistance, and heterogeneity in tumors make recurring use of certain quantitative concepts and measurements. Their frequent use may lead readers to believe that clear definitions have been settled upon for these concepts, but this is not typically the case. Instead, these concepts have operational definitions that may shift, depending on usage or evolving technology for measurement.

Tumor size, typically measured as a diameter or volume in the clinic, often must be translated to a number of cells in order for modeling approaches to make sense. Since solid tumors are typically infiltrated by non-cancerous tissue or have a necrotic core, the actual number of viable cancer cells is not always clear. Statistical methods that estimate *tumor purity* – the proportion of cancer cells in a sample – identify the most frequent cluster of somatic mutations in the sample^{128,130,190,191}. Since it is often assumed that tumors have one or more founder mutations present throughout the entire tumor, the size of this cluster is a proxy for purity. Inference is possible, but may be more difficult in polyclonal tumors, in which this founder assumption fails¹²¹.

It is important to note that not all cancer cells may be actively proliferating, meaning that model-relevant “effective” population sizes may be far less than the total number of cancer cells.

Population frequency of a particular mutation or tumor subclone refers to the fraction of cancer cells that possess a particular mutation or genotype. Changes in this frequency over time are often used to measure fitness. Determining this frequency from bulk sequencing is difficult, as the available technology reports the number of amplified reads containing individual variants, rather than the actual genotypes of single cells. As a result, two mutations that differ considerably in their observed allele frequency may actually be present in the same fraction of cells.

A host of statistical models, some of which rely on explicit evolutionary reasoning, have been developed to group mutations together into subclones (i.e., cells sharing all the mutations in the group) and infer cellular frequency of these subclones (reviewed in²²). Additionally, some methods attempt to reconstruct the ancestral relationships between these subclones from whole-exome or whole-genome data^{22,121,122}. Resolving copy number alterations into subclones is generally more difficult than resolving small substitution and indel mutations, and it requires different statistical models¹⁹².

Heterogeneity or diversity of a tumor may refer to genetic diversity measures, such as pairwise genetic distance, as is common in the field of population genetics. If discrete tumor subclones are identified, measurements of subclone diversity such as the Shannon index or Simpson’s index may be used, borrowed from the field of ecology¹³¹. Transcriptomic and proteomic measures of diversity are more complex, as they depend on expression levels; each cell can itself be considered a diverse population of functional elements. Summarizing heterogeneity from single-cell measurements is a current bioinformatic challenge^{23,77,175,193,194}.

Definitions

Adaptive radiation: Diversification of an evolving population into multiple lineages, each adapted to a separate ecological niche^{31,32}.

Morphallaxis: Anatomical regeneration following an amputation, relying on cellular redifferentiation and migration to restore lost organism function^{33,34}.

Secondary mutation: A somatic mutation that typically is observed only after one or more specific initial mutations. The appearance and growth of the initial mutation may establish conditions that favor the growth of the secondary mutation.

Compensatory mutation: A **secondary mutation** that compensates for a fitness cost incurred by one or more initial mutations. Mutations that confer drug resistance sometimes incur fitness costs; subsequent compensatory mutations can allow resistant cells to grow unimpeded by this cost.

Fitness of a genotype describes the growth potential of cells or organisms carrying that genotype, within a particular environment. Fitness “of an individual” is shorthand for describing fitness of that individual’s genotype. Fitness of individual mutations is often discussed, with the genetic background implied by context. *In vivo* fitness estimates may be performed directly, using longitudinal samples to track subclone frequency over time^{46,185,186}. With additional evolutionary modeling assumptions, indirect *in vivo* fitness estimates may be made even without longitudinal data^{183,187}.

Absolute fitness is expressed in physical units, such as net growth rate per day, or per cellular generation.

Relative fitness is a quantity that has been normalized against a standard measure of absolute fitness, such as the average fitness of a population or the fitness of a reference genotype or “wild type”. In discussions of drug resistance, a drug-sensitive ancestor is usually treated as this reference.

Neutral mutation: Strictly speaking, a mutation that carries no fitness consequences. More broadly, mutations with very minor fitness consequences, not likely to affect long-term evolution of a population, are also described as neutral or “nearly neutral”.

Driver mutation: A somatic mutation that enhances fitness of the cancer cell carrying it, potentially driving tumor growth and development.

Cross-resistance: Resistance to multiple drugs conferred by a single somatic mutation. Drug combinations to which no cross-resistance is possible are highly favored.

“Hit hard, hit early” (the “Goldie-Coldman Hypothesis”): The strategy of early treatment with multiple high-dose therapies to which no cross-resistant mutations are possible. The Goldie-Coldman Hypothesis proposes that this strategy can prevent a large, genetically diverse tumor from forming, thus preventing emergence of resistant subclones^{24,26,27}.

Genetic bottleneck: An event that drastically reduces genetic diversity of a population. Such events include appearance and rapid growth of a selectively advantageous mutant

(leading to death of all individuals not carrying the mutant) or an external force causing a sharp temporary decline in population.

Functional bottleneck: Analogous to a **genetic bottleneck**, an event that drastically reduces phenotypic diversity of a population.

Group-level / community traits are expressed by a population, but not by any one member of the population. Examples include human institutional structures and, we propose, metastatic capability of a tumor. A trait such as average swiftness of a herd of deer is not a group-level trait, as it is a mere summary of speeds of individual animals⁹⁷. The precise definition of group-level traits is a philosophically thorny issue related to arguments about group selection¹⁸⁸.

Frequency-dependent fitness is a framework for understanding how the fitness of an individual depends on the presence and frequency of other genotypes or phenotypes in a population.

Hybrid / multiscale mathematical models represent evolution of a population at multiple temporal and spatial scales. These scales can include intracellular signaling, competition and signaling among cells, diffusion of environmental resources, and establishment of spatially separate environments. Such models allow integration of population genetic theory with principles of tumor cell biology¹¹¹.

Effective population size: A single quantity summarizing demographic properties (e.g., population growth or migration patterns) that determine a population's capacity to main genetic diversity. It is *not* an approximation of the population's true or "census" size. This concept is not mysterious: it is a practical tool modelers use in order to ignore unwanted complications. Effective population size answers the question, "If I want to simulate this population using my favorite idealized mathematical model that ignores inconvenient demographic details, what size population should I simulate¹¹²?" For some models, the concept of an effective size is not useful at all¹⁸⁹.

Passenger mutation: Neutral mutations that occur in a tumor, alongside **driver mutations**.

Founder mutation: Mutations present in a cell that founds a particular clonal lineage. These mutations can be used to distinguish all descendants in the lineage.

Selective sweep: A **genetic bottleneck** caused by natural selection.

Incomplete lineage sorting: Maintenance of ancestral genetic diversity within two or more genetically isolated descendant lineages. These isolated lineages may be separate species or, in the cancer context, two physically separated lesions. Individuals within two different lineages can share a mutation, making them genetically closer (at that locus) to each other than they are to other individuals in their own lineage.

Figure 1. Competition between sensitive and resistant strains

The fitness landscape of drug-resistant strains, as well as the timing of drug resistance mutations, can determine the outcome of therapy. (A,B) Schematics of fitness of sensitive and drug-resistant cancer subclones as a function of drug concentration. In (A), the two subclones have equal fitness in the absence of drug, while in (B), the resistant subclone pays a fitness “cost” and is less fit than the sensitive clone at low drug concentrations. (C,D) Possible outcomes of therapy, given the fitness profile and drug concentration indicated by arrow in panel A. (C) The drug resistance mutation occurs early, allowing the resistant clone to grow to a large size by the time therapy begins. As a result, the slight decrease in tumor size caused by therapy may go undetected, and the tumor may be deemed intrinsically resistant to therapy. (D) The drug resistance mutation occurs late, and the resistant clone is small at the time of treatment. Substantial time may pass before the tumor regrows. The tumor is deemed to acquire resistance during therapy, even though the resistant subclone existed before treatment began. (E,F) Possible outcomes of therapy, given the fitness profile and drug concentrations indicated by arrows in panel B. (E) Drug resistance mutations occur multiple times before therapy, but cannot grow substantially due to the fitness cost. If a resistance mutation occurs near the time that treatment begins, elimination of drug-sensitive cells releases the resistant subclone from competition, allowing it to grow. (F) Same scenario

as (E), with a higher drug concentration, which even the resistant subclone cannot tolerate. Although this resistant subclone has high fitness relative to its drug-sensitive ancestor, its absolute fitness is below the viability threshold, and it eventually goes extinct.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 2. Evolution of single- and multiple-drug resistance

Each box represents a patient's population of cancer cells at a single point in time, either before or after the start of targeted therapy. (A) Single-drug therapy. Before the start of therapy, a small subpopulation of cells (shown in red) are resistant to the drug. While therapy initially decreases tumor size, the resistant subpopulation eventually regrows. (B) Dual-drug therapy. Before the start of therapy, small subpopulations are sensitive to each drug individually (shown in red and blue), but it is assumed that mutations conferring resistance to both drugs have not occurred within the same cell. As a result, application of both drugs leads to eventual eradication of the tumor. (C) Discordant penetration of both drugs throughout the tumor can lead to evolution of multiple-drug resistance. If small compartments exist that are subject to only a single drug (light red and blue strips on either side of the diagram), single-drug resistance can evolve in each compartment. Eventually, resistance to the second drug may evolve in one of these compartments, allowing migration of the doubly-resistant cell (shown in purple) into the larger region in which both drugs penetrate, leading to complete regrowth of the tumor. (D) Secretion of soluble factors conferring resistance can lead to evolution of multiple-drug resistance. In this scenario, cells resistant to a drug produce resistance-conferring factors that can be shared among nearby cells. As a result, cells may protect one another from the effects of both drugs. If single-drug

resistant subclones can grow to sufficient levels, regrowth of the tumor can occur. The precise balance of sensitive and resistant subclones will depend on the fitness costs and benefits of producing the soluble factors.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. Phenotypic diversification in the tumor ecosystem

(A,B) Simplified illustration of results of experiments described in Fig. 4 of⁹¹. Each panel represents 60 serial passages of melanoma-derived cells in mixed culture (x-axis, see original article for scale); drug resistance to trifluorothymidine (TFT), ouabain (Oua), and diaminopurine (DAP) were used as biomarkers to distinguish lineages in the culture. At different passage numbers, homogeneous subclones were isolated from the culture and tested for drug resistance. Single-cell suspensions of each subclone were then injected intravenously into mice, and metastatic capability of each aliquot was measured as the number of lung tumor colonies formed 18 days after injection (y-axis range: 0 – 350 colonies; see original article for scale). Each symbol (circle or diamond) represents the result of a single subclone. (A) Corresponding to Fig. 4(B,C) of⁹¹. Metastatic capability of three subclonal lineages remains stable until introduction of drug eradicating one of these lineages of intermediate metastatic capability; in subsequent mixed culture this intermediate phenotype is recovered by the remaining two lineages. (B) Corresponding to Fig. 4(D,E)

of⁹¹. A third clone of low metastatic capability (DAP-resistant) was added to the result of the previous mixed culture, resulting in stable propagation of metastatic capability in each lineage. After the two other clones are eradicated, the phenotype distribution of this DAP-resistant lineage expands. (C) To the extent that this niche exploration is caused by Darwinian selective pressures operating on heritable cellular traits, it is a form of **adaptive radiation**, in which a single lineage, upon gaining access to an under-utilized habitat, explores new niches and gives rise to descendants with diverse phenotypes. The evolutionary history of anole lizards in the Caribbean has witnessed recurrent adaptive radiations, as the founding lineages on each island have diversified to make greater use of the arboreal landscape. Different species make use of different parts of trees – “trunk-ground” species live near the bottom, while “crown-giants” occupy the upper canopy. This panel recapitulates a possible history of Jamaican anole species^{195,196}. The x-axis represents evolutionary time, and the y-axis represents relative perch height. (D) Instead of being caused by selective pressures, this phenotypic diversification may be caused by the coordinating effects of signaling processes, analogous to cellular processes occurring in development and regeneration or insect caste differentiation in a hive. For instance, the presence of highly metastatic cells in the mixed culture might inhibit metastatic capability of other cells by production of a secreted factor, and their removal may induce regrowth of the highly metastatic portion of the culture from other cells. The analogy that we consider is **morphallaxis**, or recreation of amputated body parts by remodeling of differentiated cells³⁴. This panel depicts regeneration of a hydra head by conversion of both epithelial cells and neurons near the wound surface, including restoration of the apical ganglia¹⁰⁶.

Figure 4. Scenarios of frequency-dependent tumor dynamics

In each scenario, a mutant subclone (purple) appears against the wild-type background (blue) at the time indicated by the star. (A) No frequency-dependence, apart from competition for limiting resources that caps the total tumor size. The fitter mutant outgrows the less-fit wild type. (B) The mutant produces a soluble growth factor that also benefits the wild type. Although the mutant is faster-growing than the wild type in isolation, this interaction allows the two subclones to coexist. (C) The mutant can only grow in the presence of an environmental resource produced by the wild type. After outcompeting the wild type to extinction, depletion of this resource leads to tumor collapse. (B,C) Arrows indicate the direction of resource flow between subclones.

Figure 5. Single-cell sequencing analyses of intra-tumor heterogeneity

Single-cell sequencing methods are powerful tools for studying clonal diversity and mutational evolution of tumors, as well as for characterizing the cellular composition and transcriptional state of their microenvironments. (A) Representation of nucleotide variants, copy number alterations, and expression data in individual cells. (B) Reconstructed mutational history of a poly-clonal triple-negative ductal breast carcinoma, inferred from single-cell copy-number profiling and exome sequencing data¹⁴⁴. (C) Expression profiles of individual cells from a metastatic melanoma tumor, displaying heterogeneity in the expression of MITF and AXL gene programs¹⁵³. Upregulation of the AXL gene program is known to be linked to resistance to RAF and MEK inhibition. (D) Two-dimensional visualization of single-cell expression data of nonmalignant cells from metastatic melanoma tumors¹⁵³. Single-cell expression data permits characterizing the cellular composition and transcriptional state of the tumor micro-environment.

Figure 6. Tumor evolutionary history determines spatial organization of somatic mutations Each panel depicts progression of a tumor over time (left to right), with height of each portion indicating relative size of clones. In each panel, a subclonal mutation occurs early in development and persists in a portion of the tumor (gray hatchmarks). A portion of the tumor, some of which contains this subclonal mutation, migrates to a different region, indicated as downward branching of the tumor. (A) In colorectal tumors studied in⁴⁴, multiple sampled regions of a tumor often contained the same mutation subclonally, indicating persistence of the mutation without occurrence of selective sweeps that dramatically change its cellular frequency. (B) In glioblastomas studied in¹⁷⁵, subclonal mutations were rarely observed in multiple lesions. If a subclonal mutation did in fact exist and migrate between regions of the brain, then a driver mutation occurring in one region could have swept it to 100% cancer cell fraction (green clone, top branch) or could have removed it completely (pink clone, bottom branch). Occurrence of these selective sweeps would prevent the shaded mutation from being detected subclonally in both branches.