The fuel for cancer evolution: intratumor heterogeneity

Rachel Rosenthal*[1,2], Nicholas McGranahan*[1,2], Javier Herrero[3], Charles Swanton[1,2]

1 UCL Cancer Institute, CRUK Lung Cancer Centre of Excellence, Paul O'Gorman Building, Huntley Street, London WC1E 6DD, UK

2 The Francis Crick Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK

3 UCL Bill Lyons Informatics Center, Paul O'Gorman Building, Huntley Street, London WC1E 6DD, UK

Email addresses:

Rachel Rosenthal – rachel.rosenthal.14@ucl.ac.uk

Nicholas McGranahan – nicholas.mcgranahan@crick.ac.uk

Javier Herrero – javier.herrero@ucl.ac.uk

Charles Swanton – charles.swanton@crick.ac.uk

Corresponding author – Charles Swanton

Shortened running title: ITH and selection in cancer evolution

Table of Contents

Abstract

Cancer is a disease reliant on the generation of mutations and the subsequent selection of those subpopulations endowed with the greatest fitness advantage. Beginning with a heterogeneous landscape of somatic alterations, various selective pressures acting on the tumor can shape the way it evolves. In this review we first discuss the current bioinformatics tools available to tease apart the heterogeneous nature of the tumor, which underpin most of the clonal analyses performed, and second consider the impact that evolutionary forces have on sculpting the tumor. Neighboring subclones may alter the microenvironment cultivating either cooperation or competition between clonal populations. Additionally the harsh environment brought about by therapy and the immune system may force adaptation. Finally, we examine recent analyses focused on precancerous samples, which help to reveal clonal selection occurring during the earliest stages of tumor development, as well as work that has identified patterns of somatic evolution observed in normal tissues.

Key words

intratumor heterogeneity, bioinformatics tools, clonal selection, Darwinian evolution, chemotherapy

Introduction

Tumor development is an evolutionary process, involving the twin forces of mutation generation and selection. One hallmark of this Darwinian process is tumor heterogeneity, which provides the fuel upon which selection acts. While mutation generation necessarily leads to heterogeneity, micro-environmental, immune and therapeutic selection pressures can dynamically sculpt a tumor as it develops, fostering either homogeneity or heterogeneity. Recent studies have begun to shed light on the extent of diversity within solid and hematological tumors, revealing that most tumors develop through a process of branched clonal evolution, whereby distinct clonal populations co-exist, possibly competing or cooperating during a cancer's progression. The wealth of data from next-generation sequencing studies has also shed light on the processes involved in generating mutations (Alexandrov et al 2013, Helleday et al 2014, Lawrence et al 2013, Segovia et al 2015) and the dynamics of tumor clones during the disease course and through treatment (Calbo et al 2011, Keats et al 2012, Landau et al 2013, Marusyk & Polyak 2010, Murtaza et al 2013).

In this review, we focus on the role of selection and clonal evolution in generating, maintaining or eliminating tumor heterogeneity. We explore tumor evolution, intratumor heterogeneity (ITH) and clonal dynamics in context of pre-cancer, immune control, and through therapy. We also investigate the tools to analyze heterogeneous tumor genomes and the insights that these tools have provided to the field.

The scale of diversity in tumors

While macroscopic ITH has likely been evident ever since tumors were first excised many thousands of years ago (Mukherjee 2011), an understanding of the extent of diversity at the single-nucleotide level within tumors has only fully emerged as a result of next-generation sequencing.

Recent in depth studies of single tumor samples, as well as multiple and serial sampling techniques have revealed considerable variability in the extent of diversity both within patients and individual tumors. Next generation sequencing has allowed for the identification and characterization of genetic ITH across a wide range of cancer types including breast carcinomas (Navin et al 2011, Nik-Zainal et al 2012b), clear-cell renal carcinomas (Gerlinger et al 2014a, Gerlinger et al 2012), glioblastomas (Sottoriva et al 2013), gliomas (Johnson et al 2014), prostate cancers (Gundem et al 2015, Haffner et al 2013), non-small cell lung cancers (de Bruin et al 2014, Zhang et al 2014), head and neck squamous cell carcinomas (Mroz et al 2015), squamous cell melanomas (Ding et al 2014), high-grade serous ovarian cancer (Schwarz et al 2015), chronic lymphocytic leukemia (Landau et al 2013), acute myeloid leukemia (Ding et al 2012, Klco et al 2014), and multiple myeloma (Bolli et al 2014, Lohr et al 2014). Taken together, these studies have demonstrated that heterogeneity is observed to varying extents across a wide variety of cancers, with both clonal and subclonal driver mutations identified. However, the majority of studies considering heterogeneity in detail have either been limited to a small number of patients, or only investigated heterogeneity based on one sample

from each tumor, thereby potentially underestimating the true extent of diversity within tumors.

Bioinformatic tools to explore heterogeneity

Data available from next-generation sequencing experiments are suited to statistical analysis to assess diversity within tumors given that sequencing data represents a random sample of DNA molecules, and by extension cancer cell genomes, within a given tumor cell population. As such, the advent of next-generation sequencing has seen a surge in computational tools to explore the clonal architecture of tumors, both from single and multi-region studies.

The fraction of reads reporting a point mutation in a sample is dependent upon the copy number at that locus, the level of tumor purity (i.e. what proportion of cells that are sequenced are tumor cells) and finally, the cancer cell fraction, describing the fraction of cancer cells in the sample that harbor the mutation. The majority of tools to dissect clonal architecture rely on the relationship between these variables, to estimate whether mutations are likely clonal or subclonal (Figure 1).

In general, a first step in dissecting the clonal architecture of a tumor involves estimating its genomic copy number profile and also its purity. Tools, such as ASCAT (Van Loo et al 2010), ABSOLUTE (Carter et al 2012), OncoSNP (Yau et al 2010), PICNIC (Greenman et al 2010) and Sequenza (Favero et al 2015) utilize mathematical frameworks to decipher copy number and data purity by

assuming sequenced DNA represent a mixture of measurements from a population of at least two distinct cell types present at different proportions: tumors cells that contain an unknown amount of DNA, and an unknown proportion of normal cells with a known amount of DNA per cell. While the system of equations is undetermined, only a few combinations of purity and ploidy can result in biologically meaningful solutions (Van Loo & Campbell 2012).

Once the copy number and purity of a sample have been determined, the cancer cell fraction of each mutation can be estimated. A simple approach to determine the clonality of a given mutation is to determine whether the observed variant allele frequency differs from what would be expected given a clonal mutation (Carter et al 2012, Stephens et al 2012). More sophisticated methods rely on the assumption that mutations with similar variant allele frequencies may correspond to clonal or subclonal clusters, reflecting nodes on an evolutionary tree. For example, PyClone (Roth et al 2014, Shah et al 2012) integrates variant allele frequencies with allele specific integer copy number estimates in order to define the subclonal composition within individual biopsies. The method uses a Bayesian Dirichlet clustering process to jointly group mutations, and infers posterior density estimates over the cancer cell fraction for each mutation. Conveniently, modeling the number of subclones as coming from a Dirichlet process does not require knowledge of the number of subclones a priori (Roth et al 2014). A limitation of PyClone, however, is that it assumes all copy number events are clonal, and deep sequencing is recommended (Roth et al 2014).

The method adopted by Nik-Zainal (2012b) leverages data from whole genome sequencing to circumvent the need for deep sequencing and allows mutations to reside on subclonal copy numbers. By contrast, SciClone (Miller et al 2014), which also applies a Bayesian clustering method, focuses on single-nucleotide variants (SNVs) in balanced copy-number regions. Although this feature of SciClone circumvents issues associated with clonal and subclonal copy number aberrations, it means the clonality of every mutation cannot be determined.

More recent methods also use the fact that the structure of mutations should be hierarchical, with nested subclones (Deshwar et al 2015, Jiao et al 2014), as well as modeling different forms of data together (Fischer et al 2014). These methods have also been extended to allow multiple samples over space or time to be included in subclonal clustering and this may considerably improve the accuracy of subclonal reconstruction.

Importantly, it may be difficult to accurately de-convolve the subclonal structure in a tumor from a single sample. For example, if two subclonal populations have similar cancer cell fractions in one tumor region, they will appear as one clone. Analysis of another tumor region may enable their separation. Similarly, without single cell or multi-region sequencing an amplification to 8 copies occurring in half the tumor population may appear like an amplification to 4 copies in the tumor population as a whole. In addition to multi-region sequencing, further information regarding the clonal composition of tumors can be inferred based

on the mutual exclusivity or co-occurrence of mutations in cancer cells, either from single cell sequencing or from ‘phasing’. Phasing involves determining whether mutations co-occur or are mutually exclusive, allowing different subclones to be delineated; if two mutations never appear to occur together on the same haplotype they are likely to represent distinct subclones, whereas if two mutations can be phased in the same cancer cell they are necessarily present in the same lineage (Nik-Zainal et al 2012b). However, phasing approaches are currently limited to analysis of mutations in regions of hyper-mutation or high mutation burden.

Finally, it is important to recognize that the cancer cell fraction of a mutation need not directly reflect its timing or phylogenetic relationship with other mutations. For instance, given the dynamic nature of tumor evolution, an early ‘truncal’ mutation may later be lost as a result of copy number alterations (Murugaesu et al 2015). Equally, the identification of a mutation in every tumor region does not mean it is necessarily truncal as subclonal mutations may be present at low frequencies in multiple tumor regions.

Cancer, selection, and neutral evolution

First formally proposed as such by Nowell (1976), tumor progression represents an evolutionary process under continuous Darwinian selection (Greaves & Maley 2012). At the time of detection, a tumor will have undergone many rounds of cell division, with each generation of cells stochastically acquiring novel somatic mutations (Gerlinger et al 2014b). While most of these

mutations may have little impact on the overall fitness of the cell, a minority subset of these mutations (known as driver events) will endow a cell with an evolutionary advantage, allowing that cell and its progeny to flourish and outcompete others. Combined, the processes of clonal selection and clonal evolution acting on the cell may result in the outgrowth of multiple subclones, often with their own distinct driver events, leading to the branched evolutionary phylogeny that has been observed across many cancer types (de Bruin et al 2014, Gerlinger et al 2012, Gundem et al 2015, Nik-Zainal et al 2012b, Sottoriva et al 2013) (Figure 2A).

Generally, the number of driver events required for tumor initiation and progression and how this changes with time and tumor microenvironment remains unknown. Conventional theory has viewed tumorigenesis as a slow process of clonal evolution as driver mutations are gradually acquired, enabling various populations to expand, resulting in the formation of progressively more disordered clones (Merlo et al 2006, Stratton et al 2009). However recent studies have shown that the malignant transformation of normal mammary tissue may also be induced surprisingly quickly and efficiently, requiring the use of only *KRAS* G12D as an oncogene (Nguyen et al 2015). Additionally the analysis of tumor genomes has recently revealed novel catastrophic events such as chromothripsis, chromoplexy, and kataegis, which may result in a drastic shift in the evolutionary trajectory of a tumor (Baca et al 2013, Campbell et al 2010, Forment et al 2012, Nik-Zainal et al 2012a, Shen 2013, Stephens et al 2012).

While the majority of established driver events are clonal, indicating that they likely arise early in tumor evolution, subclonal driver events have been identified across many cancers and are thought to play a role in tumor maintenance and progression, potentially leading to subclonal expansions (McGranahan & Swanton 2015). Even the most common driver events may occur early in some tumors while occurring late in others (McGranahan et al 2015, Yates et al 2015). This is reflective of the fact that the delineation between driver and passenger mutations is context dependent. As selective pressures and the tumor microenvironment change, so do the requirements for tumor survival.

Subclonal populations of cancerous cells give rise to a heterogeneous environment within the tumor; however, each subclone is not an isolated entity. Interaction between subclones during tumor evolution may be competitive or cooperative. One subclone may outcompete another for vital resources, such as oxygen, nutrients, or space (Marusyk & Polyak 2010). Over time, different subclones may even alternate as the dominant clone in a population, indicative of the dynamic selective pressures influencing clonal competition (Egan et al 2012, Keats et al 2012).

Alternatively low frequency clones may support the growth of the dominant clone or promote resistance to therapy through paracrine signaling networks (Hobor et al 2014, Inda et al 2010). Additionally crosstalk between different cell populations is capable of shaping tumor properties, such as metastatic potential as has been shown in mouse models of small cell lung carcinoma (Calbo et al 2011). It is not necessary for a subclone to have an obvious fitness advantage

itself in order to affect tumor development, as the subclone may still drive tumor growth by inducing favorable changes in the microenvironment (Marusyk et al 2014). However if the subclone responsible for contributing to the advantageous growth conditions is outcompeted by another faster proliferating subclone, which itself is dependent on the current microenvironment, tumor collapse may occur (Maley et al 2004, Marusyk et al 2014).

In addition to the ongoing clonal selection evident in some cancers, recent work has also highlighted the occurrence of neutral evolution during tumor development (Figure 2B). Using a theoretical model that determines the expected distribution of subclonal mutations under neutral evolutionary processes, Williams and colleagues demonstrated that the extensive heterogeneity observed in some tumors can be explained by neutral expansion and accumulation of passenger mutations after any early driver mutations required for tumor initiation have been acquired (Williams et al 2016). The absence of selective sweeps as predicted under a model of neutral evolution implies that once a mutation has arisen in a surviving lineage, it will persist and expand at the same rate as any other mutation. Thus, the entire historical record of tumor evolution can be observed in the current tumor genome. The earliest mutations that initiated tumor growth, would be present in all tumor cells, whereas more recently acquired mutations would be permanently constrained to a smaller tumor subpopulation, without any selective sweeps occurring to change their relative prevalence. One consequence of this model is that the malignant potential of a tumor may be determined very early in development (Sottoriva et al 2015). While the distribution of subclonal

mutations fits the neutral evolution model for a subset of tumors across many cancer types, it remains to be determined how prevalent the phenomena is with more comprehensive sampling strategies and how copy number corrections of the variant allele frequencies impacts the model.

Selection and somatic evolution in normal and precancerous cells

Stochastically acquired mutations in normal and precancerous tissues may be selected for, eventually resulting in the clinical presentation of a tumor. However before that stage, the analysis of noncancerous samples can shed light on the earliest stages of tumor development. Studies focusing on Barrett's esophagus, a precursor lesion to esophageal adenocarcinoma (EAC), have revealed that it is more likely to progress to cancer if there is a higher degree of clonal diversity present (Maley et al 2006, Merlo et al 2010). Indeed, by the time EAC is clinically diagnosed, there has often been such evolution that the percentage of mutations shared between EAC and the adjacent segments of Barrett's esophagus can be less than 20% (Ross-Innes et al 2015).

The order in which mutations arise can influence the outcome of subsequent selective pressures, restrict evolutionary paths (Papaemmanuil et al 2013), and affect the clinical behavior of disease presentation as well as response to therapy (Ortmann et al 2015). In a recent study on the evolution of melanoma from precursor lesions, a *BRAF* V600E alone was found to be sufficient to form a nevus; however precursor lesions with *NRAS* or alternative *BRAF* mutations also harbored additional oncogenic mutations. Moreover, it was observed that

different melanoma subtypes had evolved via separate means (Shain et al 2015).

Evidence of mutation acquisition and selection can also be observed in noncancerous cells. For instance, although human hematopoietic stem/progenitor cells (HSPCs) divide very rarely, they still stochastically acquire mutations that may confer a slight selective growth advantage. In an analysis of germline reference blood samples taken from TCGA patients diagnosed with many different solid cancer types, mutations associated with leukemia and/or lymphoma were identified. Importantly, none of these patients had any observable sign of hematological malignancies, and the fraction of patients harboring leukemia and/or lymphoma associate mutations increased with patient age. Thus these mutations appear to represent the earliest stages clonal selection and hematopoietic expansion (Xie et al 2014). Furthermore, apparently healthy individuals who carried driver mutations in their blood have been shown to be at a higher risk of developing blood cancers (Genovese et al 2014, Jaiswal et al 2014).

High numbers of mutations and extensive ITH has also been identified in normal skin tissue from the eyelids of middle-aged individuals. Positive selection of many known squamous skin cancer driver genes has been identified, with some clones acquiring multiple driver mutations without undergoing malignant transformation. This observation and the fact that in most cases, mutations were only detected in a small fraction of the cells suggests

that clonal outgrowth is not solely determined by driver gene acquisition or is somehow curtailed early on (Martincorena et al 2015).

Therapy, selection and heterogeneity

Longitudinal analyses of tumor samples have consistently identified shifts in the genomes of samples taken before and after treatment with chemotherapeutics (Ding et al 2012, Johnson et al 2014, Landau et al 2013, Mullighan et al 2008, Murugaesu et al 2015, Schuh et al 2012), indicating that cancer therapy can play a major role in altering the genomic landscape of a tumor.

One way therapy may impact the evolution of the tumor is by directly acting as a mutagenic agent. In this case, the genotoxic effects of chemotherapy may be reflected by changes to the mutational landscape and spectra of the tumor. Consistent with this, an analysis of primary and relapsed acute myeloid leukemia (AML) samples revealed an increase in transversion rate among relapse-specific mutations after cytotoxic therapy, indicating that at relapse chemotherapy had altered the mutational spectrum (Ding et al 2012).

In a multi-region exome sequencing analysis of a small cohort of 4 patients with esophageal adenocarcinoma taken before and after treatment with a platinum-containing chemotherapy, an increase in C>A transversions at CpC sites present in the post-chemotherapy samples of patients with residual disease was identified (Murugaesu et al 2015). Mutations in this particular context have been previously identified in *C. elegans* treated with cisplatin, a platinum based

chemotherapeutic (Meier et al 2014). The majority of the mutations observed in the platinum-associated mutational context by Murugaesu et al. were subclonal, consistent with those mutations occurring late in tumor evolution, as would be expected for chemotherapy-induced mutagenesis. Consistent with these findings, a larger study of 30 paired esophageal adenocarcinomas sampled before and after neo-adjuvant chemotherapy also found a significant increase in C>A transversions after treatment (Findlay et al 2016). As the presence of residual disease is indicative of an incomplete response to chemotherapy, the observation that some drugs leave behind clear signs of mutagenic activity in subclonal populations, highlights the need to better determine which patients are most likely to clinically benefit to therapy.

Conceivably, mutations generated from chemotherapeutics may not only leave scars in the genome but also may directly contribute to disease progression (Figure 3A). In a seminal study, Johnson and colleagues examined the mutational landscape of a cohort of initial low grade gliomas and their recurrences and found that 6 out of 10 patients treated with temozolomide recurred as hypermutated high-grade gliomas, with the majority of newly acquired mutations occurring in a temozolomide associated context (Johnson et al 2014). Additionally within this mutational context, they identified driver mutations in the RB and Akt-mTOR pathways. Their finding highlights how chemotherapy-induced mutagenesis is not limited to solely aiding genetic diversification, but that it can also influence the evolutionary path taken by the tumor, resulting in this case, in a tumor of higher histological grade with poor prognosis.

A subset of tumors from patients who have been treated with alkylating agents, such as temozolomide, exhibit a significant increase in overall mutation burden, consisting of primarily C>T transitions at CpC and CpT dinucleotide sites (Alexandrov et al 2013, Johnson et al 2014). The proposed mechanism for resistance and hypermutation is via a selective advantage for clones that acquired inactivating somatic mutations of *MSH6*, rendering them resistant to alkylating agents, yet left to undergo accelerated mutagenesis due to the lack of effective mismatch repair (Hunter et al 2006).

Additionally chemotherapy-induced mutagenesis may confound conclusions drawn from genomic analyses. Somatic mosaic protein truncating variants in *PPM1D* were originally identified as associated with germline predisposition to breast cancer and ovarian cancer in women (Ruark et al 2013). However further analysis has revealed that that these variants are more commonly observed in post-chemotherapy cases rather than pretreatment cases, suggesting that they are somatic mutations caused by treatment (Pharoah et al 2016, Swisher et al 2016).

Even without directly inducing novel mutations, which may or may not be selected for, cancer therapy can alter the genomic landscape and heterogeneity of a tumor by applying new selective pressures, which can drive evolution reliant on the genetic variation that existed prior to the start of treatment. Within a heterogeneous tumor, some subclones may be present that originally had no obvious fitness advantage but impart a resistance to therapy

and are subsequently selected for (Figure 3B). Indeed, there have been numerous reports detailing the outgrowth of resistant subclonal populations in response to therapy across many cancer types including colorectal (Diaz et al 2012, Kreso et al 2013), glioblastoma (Cahill et al 2007, Yip et al 2009), melanoma (Shi et al 2014, Wagle et al 2011), non-small cell lung cancer (Kosaka et al 2006, Turke et al 2010), and CML (Shah et al 2002).

Recent research has also shown the utility of sequencing circulating tumor DNA (ctDNA) taken at multiple time points to non-invasively monitor clonal dynamics and identify emerging resistance to therapy (Diaz et al 2012, Murtaza et al 2015, Murtaza et al 2013, Russo et al 2016). Murtaza and colleagues followed a patient with metastatic ER-positive, HER2-positive breast cancer for more than 3 years and could infer the clonal evolution of the tumor over time in response to two lines of targeted therapy (2015). They detected both the presence of a *PIK3CA* mutation at the time the patient progressed on trastuzumab and tamoxifen therapy and its decline after lapatinib treatment was begun. This was followed by an increase in allele frequency of an *ERBB4* mutation in response to lapatinib therapy that was predicted to contribute to the patient's resistance to the drug. Additionally, in colorectal cancer integrative sequencing analyses of spatially and temporally separated tumor biopsies along with ctDNA has identified polyclonal mechanisms of resistance to *EGFR* blockade in distinct metastases from a single patient, highlighting the limits of relying on a single-lesion biopsy for decisions pertaining to the treatment of a heterogeneous tumor (Russo et al 2016).

Importantly, while the emerging mutations identified on the path to drug resistance are often initially undetectable above background levels, little remains known about the evolution of resistant clones during the course of drug therapy. In particular, the sensitivity of next generation sequencing is often not great enough to conclusively determine whether resistant clones existed before drug exposure, or if they evolved from the original clones to acquire resistance. In a recent study of *EGFR*-mutant NSCLC cell lines, Hata et al. used far more sensitive sequencing technologies to track *EGFR* T790M drug-resistant clones and found that both the intrinsic and acquired routes to resistance could be observed (2016). Intriguingly, they also found biological differences between those clones that had a pre-existing *EGFR* T790M mutation and those that were initially drug susceptible and evolved a T790M mutation de novo, which may help inform future therapeutic choices.

Clonal shifts without known resistance mechanisms have also been observed after treatment with cancer chemotherapy (Figure 3C). Studies performed by Landau and colleagues in CLL (2013, 2015) have shown clear evidence of clonal evolution over time, where minor subclones at the start of treatment expand during the course of treatment. Often these subclones are enriched for driver mutations, suggesting that the selective pressure of treatment could remove incumbent clones and result in the growth of more aggressive subclones (Landau et al 2013, Landau et al 2015). Furthermore, they observed that the presence of a subclonal driver was independently associated with poorer outcome and response to therapy. Evidence of changes in clonal composition in response to therapy has also been detailed in multiple myeloma

(Egan et al 2012, Weston-Bell et al 2013). Intriguingly, Landau and colleagues also observed a minority of CLL cases that relapsed after undergoing treatment yet maintained a stable clonal equilibrium (Figure 3D) (Landau et al 2013, Landau et al 2015).

Examining the clonal dynamics of a tumor over multiple points in evolutionary time may also provide a means for future therapeutic intervention. A recent analysis of brain metastases and their matched primary tumor samples observed that over half of the brain metastatic samples contained potentially clinically actionable alterations that were not ubiquitously present in the primary sample (Brastianos et al 2015). Hata et al also noted, with all the caveats of a limited sample size, that in their study of *EGFR* T790M mutant cell lines, the cell lines with a de novo *EGFR*-resistance mutation showed less evidence of a response to third-generation *EGFR* inhibitors, further highlighting the complexity of resistance mechanisms (2016). Such observations could potentially have important implications in the clinic.

Immune response and heterogeneity

There is much evidence supporting an interaction between the tumor and the immune system (Dunn et al 2002). Tumor cells may induce an immune response by expressing antigens that are recognized by a patient's T cells. Indeed therapeutic tumor regression has been observed through the collection, ex vivo expansion, and re-administration of autologous tumor infiltrating lymphocytes (TILs) resected from patients with metastatic melanoma. This

indicates that there are T cells present capable of recognizing tumor cells and mounting an immune response against them (Dudley et al 2002, Rosenberg 2012).

By recognizing some tumor specific antigens, the immune system may also produce an evolutionary pressure, shaping the antigenicity of the tumor and its diversity as it evolves. Driven by the process of immunoediting, which describes the interaction between the tumor and immune system wherein the immune system plays the dual role of protecting the host and sculpting the tumor, subclonal populations of tumor cells either lacking immunogenic antigens or able to withstand an immune response may be selected for (DuPage et al 2012, Matsushita et al 2012, Schreiber et al 2011).

While an immune response may be mounted against self-antigens that are aberrantly expressed in cancerous tissues, recently there has been much interest in elucidating how the mutational landscape of a tumor may result in tumor-specific neoantigens that have arisen from non-silent mutations. These neoantigenic mutations result in novel epitopes that may be recognized as foreign by a patient's tumor-infiltrating T lymphocytes (Rajasagi et al 2014).

Through a pan-cancer study of patients from The Cancer Genome Atlas (TCGA), Rooney and colleagues defined an expression-based measure of cytolytic activity and observed that it was positively correlated with the number of putative neoantigens in multiple tumor types. Furthermore, by comparing the observed and expected numbers of predicted neoantigens generated per non-

silent mutation, they found evidence of immunoediting, T cell-mediated surveillance and elimination of subclones containing neoantigens (Rooney et al 2015). Additional studies of TCGA patients have shown that both the presence and number of predicted neoantigens associates with overall survival (Brown et al 2014, McGranahan et al 2016).

While the impact of the clonal architecture of the tumor on the neoantigen repertoire remains unclear (Heemskerk et al 2013), there are some early indications that considering ITH might contribute towards characterizing the immunogenicity of a tumor. It has been shown that, among non-hypermutated tumors, the most homogenous tumors have the largest number of predicted neoantigens (Angelova et al 2015, McGranahan et al 2016). More heterogeneous tumors have been shown to have an overall greater depletion of immune subpopulations (Angelova et al 2015) and lower levels of immune infiltration (Morris et al 2016). Additionally, patients with homogeneous tumors exhibit longer overall survival than patients with more heterogeneous tumors (Angelova et al 2015, McGranahan et al 2016). These observations implicate ITH as a potentially important factor in determining the immune response elicited by a tumor.

Immunotherapy and heterogeneity

The influence of ITH on the immune response may also have implications for personalized immune therapies, particularly those which target tumor neoantigens by modulating the activity of the immune system. The use of

immune checkpoint inhibitors, such as antibodies directed against programmed cell death-1 (PD-1) or cytotoxic T lymphocyte-associated antigen-4 (CTLA-4), has shown great promise across a wide variety of cancers; however, clinical benefit has only been seen within a subset of the patient population (Hodi et al 2010, Pardoll 2012, Topalian et al 2012). To date, molecular determinants capable of predicting patient response to checkpoint blockade have been hard to establish. In seminal work studying NSCLC (Rizvi et al 2015) and melanoma (Snyder et al 2014, Van Allen et al 2015) overall mutation burden has been shown to correlate with response to anti-PD-1 therapy and anti-CTLA-4 therapy, respectively, while in colorectal cancers (Le et al 2015), mismatch repair status could predict clinical benefit to PD-1 blockade.

Based on the observations made in the treatment-naïve setting, it has been suggested that tumors with a low level of ITH and a high clonal neoantigen repertoire might further contribute to checkpoint inhibitor response. Such patients may respond better due to the higher level of immune infiltration in more homogeneous tumors or because a highly homogeneous tumor itself may be evidence of extensive immunoediting of antigenic subclonal populations by a functional and active immune system (Angelova et al 2015, McGranahan et al 2016, Morris et al 2016). Recent research, further exploring the same checkpoint blockade treated NSCLC and melanoma cohorts (Rizvi et al 2015, Snyder et al 2014, Van Allen et al 2015) has found that response to PD-1 and CTLA-4 antibodies was particularly improved in patients with tumors enriched for clonal neoantigens. Among the patients treated with an anti-PD-1 antibody, patients without durable benefit were found to have a significantly more

heterogeneous neoantigen repertoire than patients experiencing a durable clinical benefit. Additionally, incorporating a measure of heterogeneity, rather than considering total neoantigen burden alone, could more accurately stratify these patients into durable clinical benefit or no clinical benefit groups (McGranahan et al 2016).

These observations suggest that the clonal structure of neoantigenic mutations may play a role in immune surveillance and raises the question of whether subclonal mutations are sufficient to generate a tumor wide immune response. Furthermore, among the anti-CTLA4 treated cohort, a small subset of patients previously treated with an alkylating agent had an increased mutational load of subclonal mutations, consistent with therapy-induced hypermutation, but were among the poor or limited responders, suggesting that while therapy may induce potentially immunogenic mutations, those might not be always sufficient to elicit an efficient anti-tumor immune response (McGranahan et al 2016).

It is possible that in the presence of many antigenic subclonal mutations, an immune response is only mounted against a few them, and even then, when lymphocytes are generated that recognize subclonal neoantigens, it is possible they would not be able to target the whole tumor. Thus as more data becomes available, it will be of vital importance to determine the full extent of the interplay between the clonal architecture of a tumor and the potential reaction of the immune system.

Future perspectives

Here we have considered the selective forces shaping a tumor as it develops, including the response to its environment and other extrinsic factors such as therapy. As the continuous acquisition of mutations results in a heterogeneous landscape that confers tumors with increased resilience, it will be of particular importance to determine the impact underlying clonal architecture has on tumor evolution. Most studies to date have been constrained by sequencing depth and single sample biopsies; therefore calculations of tumor heterogeneity are bound to reflect underestimates.

Moving forward, improved computational methods to better dissect clonal architecture and model tumor clonal dynamics, as well as studies examining tumors with finer resolution through the use of deeper sequencing depth, single cell samples, or multiple tumor regions, will help to elucidate the extent of impact that ITH has on a tumor's evolutionary path and potential. One such study, TRACERx (TRAcking non-small cell lung Cancer Evolution through therapy), which uses multi-region and longitudinal ultra-deep exome sequencing to prospectively track the evolution of primary non-small cell lung cancer from diagnosis through treatment and relapse, has already been commenced (Jamal-Hanjani et al 2014). Additionally, this study along with others have begun to analyze tumor immune infiltrates in order to assess how the immune system may shape tumor development and determine how heterogeneity impacts antitumor immunity both with and without the aid of immunotherapy (Angelova et al 2015, Jamal-Hanjani et al 2014, Llosa et al 2015). Beyond genomic analyses of ITH, DNA methylation studies have also begun to identify extensive epigenetic ITH, adding yet another layer of

complexity to the understanding of tumor evolution (Brocks et al 2014, Mazor et al 2015, Oakes et al 2014).

Continuing such studies will lead to an improved understanding of ITH and clonal dynamics, endowing us with the ability to more fully decipher the evolutionary history of a tumor as well as a greater understanding of which events are truly clonal and how the tumor may respond to therapy. This will potentially translate into novel therapeutic approaches and inform new ways to best stratify patient groups for maximal treatment efficacy.

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, et al. 2013. Signatures of mutational processes in human cancer. *Nature* 500: 415-21

Angelova M, Charoentong P, Hackl H, Fischer ML, Snajder R, et al. 2015. Characterization of the immunophenotypes and antigenomes of colorectal cancers reveals distinct tumor escape mechanisms and novel targets for immunotherapy. *Genome Biol* 16: 64

Baca SC, Prandi D, Lawrence MS, Mosquera JM, Romanel A, et al. 2013. Punctuated evolution of prostate cancer genomes. *Cell* 153: 666-77

Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, et al. 2014. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun* 5: 2997

Brastianos PK, Carter SL, Santagata S, Cahill DP, Taylor-Weiner A, et al. 2015. Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets. *Cancer Discov* 5: 1164-77

Brocks D, Assenov Y, Minner S, Bogatyrova O, Simon R, et al. 2014. Intratumor DNA methylation heterogeneity reflects clonal evolution in aggressive prostate cancer. *Cell Rep* 8: 798-806

Brown SD, Warren RL, Gibb EA, Martin SD, Spinelli JJ, et al. 2014. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Res* 24: 743-50

Cahill DP, Levine KK, Betensky RA, Codd PJ, Romany CA, et al. 2007. Loss of the mismatch repair protein MSH6 in human glioblastomas is associated with tumor progression during temozolomide treatment. *Clin Cancer Res* 13: 2038-45

Calbo J, van Montfort E, Proost N, van Drunen E, Beverloo HB, et al. 2011. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. *Cancer Cell* 19: 244-56

Campbell PJ, Yachida S, Mudie LJ, Stephens PJ, Pleasance ED, et al. 2010. The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* 467: 1109-13

Carter SL, Cibulskis K, Helman E, McKenna A, Shen H, et al. 2012. Absolute quantification of somatic DNA alterations in human cancer. *Nature biotechnology* 30: 413-21

de Bruin EC, McGranahan N, Mitter R, Salm M, Wedge DC, et al. 2014. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 346: 251-6

Deshwar AG, Vembu S, Yung CK, Jang GH, Stein L, Morris Q. 2015. PhyloWGS: reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol* 16: 35

Diaz LA, Jr., Williams RT, Wu J, Kinde I, Hecht JR, et al. 2012. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 486: 537-40

Ding L, Kim M, Kanchi KL, Dees ND, Lu C, et al. 2014. Clonal architectures and driver mutations in metastatic melanomas. *PLoS One* 9: e111153

Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, et al. 2012. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481: 506-10

Dudley ME, Wunderlich JR, Robbins PF, Yang JC, Hwu P, et al. 2002. Cancer regression and autoimmunity in patients after clonal repopulation with antitumor lymphocytes. *Science* 298: 850-4

Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. 2002. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol* 3: 991-8

DuPage M, Mazumdar C, Schmidt LM, Cheung AF, Jacks T. 2012. Expression of tumour-specific antigens underlies cancer immunoediting. *Nature* 482: 405-9

Egan JB, Shi CX, Tembe W, Christoforides A, Kurdoglu A, et al. 2012. Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood* 120: 1060-6

Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, et al. 2015. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol* 26: 64-70

Findlay JM, Castro-Giner F, Makino S, Rayner E, Kartsonaki C, et al. 2016. Differential clonal evolution in oesophageal cancers in response to neo-adjuvant chemotherapy. *Nat Commun* 7: 11111

Fischer A, Vazquez-Garcia I, Illingworth CJ, Mustonen V. 2014. High-definition reconstruction of clonal composition in cancer. *Cell Rep* 7: 1740-52

Forment JV, Kaidi A, Jackson SP. 2012. Chromothripsis and cancer: causes and consequences of chromosome shattering. *Nat Rev Cancer* 12: 663-70

Genovese G, Kahler AK, Handsaker RE, Lindberg J, Rose SA, et al. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* 371: 2477-87

Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, et al. 2014a. Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* 46: 225-33

Gerlinger M, McGranahan N, Dewhurst SM, Burrell RA, Tomlinson I, Swanton C. 2014b. Cancer: evolution within a lifetime. *Annu Rev Genet* 48: 215-36

Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, et al. 2012. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 366: 883-92

Greaves M, Maley CC. 2012. Clonal evolution in cancer. *Nature* 481: 306-13

Greenman CD, Bignell G, Butler A, Edkins S, Hinton J, et al. 2010. PICNIC: an algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics* 11: 164-75

Gundem G, Van Loo P, Kremeyer B, Alexandrov LB, Tubio JM, et al. 2015. The evolutionary history of lethal metastatic prostate cancer. *Nature* 520: 353-7

Haffner MC, Mosbruger T, Esopi DM, Fedor H, Heaphy CM, et al. 2013. Tracking the clonal origin of lethal prostate cancer. *The Journal of clinical investigation* 123: 4918-22

Hata AN, Niederst MJ, Archibald HL, Gomez-Caraballo M, Siddiqui FM, et al. 2016. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat Med* 22: 262-9

Heemskerk B, Kvistborg P, Schumacher TN. 2013. The cancer antigenome. *EMBO J* 32: 194-203

Helleday T, Eshtad S, Nik-Zainal S. 2014. Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet* 15: 585-98

Hobor S, Van Emburgh BO, Crowley E, Misale S, Di Nicolantonio F, Bardelli A. 2014. TGF-alpha and amphiregulin paracrine network promotes resistance to EGFR blockade in colorectal cancer cells. *Clin Cancer Res*

Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, et al. 2010. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 363: 711-23

Hunter C, Smith R, Cahill DP, Stephens P, Stevens C, et al. 2006. A hypermutation phenotype and somatic MSH6 mutations in recurrent human malignant gliomas after alkylator chemotherapy. *Cancer Res* 66: 3987-91

Inda MM, Bonavia R, Mukasa A, Narita Y, Sah DW, et al. 2010. Tumor heterogeneity is an active process maintained by a mutant EGFR-induced cytokine circuit in glioblastoma. *Genes Dev* 24: 1731-45

Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, et al. 2014. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* 371: 2488-98

Jamal-Hanjani M, Hackshaw A, Ngai Y, Shaw J, Dive C, et al. 2014. Tracking genomic cancer evolution for precision medicine: the lung TRACERx study. *PLoS Biol* 12: e1001906

Jiao W, Vembu S, Deshwar AG, Stein L, Morris Q. 2014. Inferring clonal evolution of tumors from single nucleotide somatic mutations. *BMC Bioinformatics* 15: 35

Johnson BE, Mazor T, Hong C, Barnes M, Aihara K, et al. 2014. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. *Science* 343: 189-93

Keats JJ, Chesi M, Egan JB, Garbitt VM, Palmer SE, et al. 2012. Clonal competition with alternating dominance in multiple myeloma. *Blood*

Klco JM, Spencer DH, Miller CA, Griffith M, Lamprecht TL, et al. 2014. Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. *Cancer Cell* 25: 379-92

Kosaka T, Yatabe Y, Endoh H, Yoshida K, Hida T, et al. 2006. Analysis of epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer and acquired resistance to gefitinib. *Clin Cancer Res* 12: 5764-9

Kreso A, O'Brien CA, van Galen P, Gan OI, Notta F, et al. 2013. Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* 339: 543-8

Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, et al. 2013. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* 152: 714-26

Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, et al. 2015. Mutations driving CLL and their evolution in progression and relapse. *Nature* 526: 525-30

Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, et al. 2013. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499: 214-8

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, et al. 2015. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 372: 2509-20

Llosa NJ, Cruise M, Tam A, Wicks EC, Hechenbleikner EM, et al. 2015. The vigorous immune microenvironment of microsatellite instable colon cancer is

balanced by multiple counter-inhibitory checkpoints. *Cancer Discov* 5: 43-51

Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, et al. 2014. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell* 25: 91-101

Maley CC, Galipeau PC, Finley JC, Wongsurawat VJ, Li X, et al. 2006. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet* 38: 468-73

Maley CC, Reid BJ, Forrest S. 2004. Cancer prevention strategies that address the evolutionary dynamics of neoplastic cells: simulating benign cell boosters and selection for chemosensitivity. *Cancer Epidemiol Biomarkers Prev* 13: 1375-84

Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, et al. 2015. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 348: 880-6

Marusyk A, Polyak K. 2010. Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta* 1805: 105-17

Marusyk A, Tabassum DP, Altrock PM, Almendro V, Michor F, Polyak K. 2014. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* 514: 54-8

Matsushita H, Vesely MD, Koboldt DC, Rickert CG, Uppaluri R, et al. 2012. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* 482: 400-4

Mazor T, Pankov A, Johnson BE, Hong C, Hamilton EG, et al. 2015. DNA Methylation and Somatic Mutations Converge on the Cell Cycle and Define Similar Evolutionary Histories in Brain Tumors. *Cancer Cell* 28: 307-17

McGranahan N, Favero F, de Bruin EC, Birkbak NJ, Szallasi Z, Swanton C. 2015. Clonal status of actionable driver events and the timing of mutational processes in cancer evolution. *Sci Transl Med* 7: 283ra54

McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, et al. 2016. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* 351: 1463-9

McGranahan N, Swanton C. 2015. Biological and Therapeutic Impact of Intratumor Heterogeneity in Cancer Evolution. *Cancer Cell* 27: 15-26

Meier B, Cooke SL, Weiss J, Bailly AP, Alexandrov LB, et al. 2014. C. elegans whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. *Genome Res*

Merlo LM, Pepper JW, Reid BJ, Maley CC. 2006. Cancer as an evolutionary and ecological process. *Nat Rev Cancer* 6: 924-35

Merlo LM, Shah NA, Li X, Blount PL, Vaughan TL, et al. 2010. A comprehensive survey of clonal diversity measures in Barrett's esophagus as biomarkers of progression to esophageal adenocarcinoma. *Cancer Prev Res (Phila)* 3: 1388-97

Miller CA, White BS, Dees ND, Griffith M, Welch JS, et al. 2014. SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput Biol* 10: e1003665

Morris LG, Riaz N, Desrichard A, Senbabaoglu Y, Hakimi AA, et al. 2016. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget*

Mroz EA, Tward AD, Hammon RJ, Ren Y, Rocco JW. 2015. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. *PLoS Med* 12: e1001786

Mukherjee S. 2011. *The Emperor of All Maladies*: 4th Estate

Mullighan CG, Phillips LA, Su X, Ma J, Miller CB, et al. 2008. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 322: 1377-80

Murtaza M, Dawson SJ, Pogrebniak K, Rueda OM, Provenzano E, et al. 2015. Multifocal clonal evolution characterized using circulating tumour DNA in a case of metastatic breast cancer. *Nat Commun* 6: 8760

Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, et al. 2013. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature*

Murugaesu N, Wilson GA, Birkbak NJ, Watkins T, McGranahan N, et al. 2015. Tracking the genomic evolution of esophageal adenocarcinoma through neoadjuvant chemotherapy. *Cancer Discov*

Navin N, Kendall J, Troge J, Andrews P, Rodgers L, et al. 2011. Tumour evolution inferred by single-cell sequencing. *Nature* 472: 90-4

Nguyen LV, Pellacani D, Lefort S, Kannan N, Osako T, et al. 2015. Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. *Nature* 528: 267-71

Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, et al. 2012a. Mutational processes molding the genomes of 21 breast cancers. *Cell* 149: 979-93

Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, et al. 2012b. The life history of 21 breast cancers. *Cell* 149: 994-1007

Nowell PC. 1976. The clonal evolution of tumor cell populations. *Science* 194: 23-8

Oakes CC, Claus R, Gu L, Assenov Y, Hullein J, et al. 2014. Evolution of DNA methylation is linked to genetic aberrations in chronic lymphocytic leukemia. *Cancer Discov* 4: 348-61

Ortmann CA, Kent DG, Nangalia J, Silber Y, Wedge DC, et al. 2015. Effect of mutation order on myeloproliferative neoplasms. *N Engl J Med* 372: 601-12

Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, et al. 2013. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* 122: 3616-27; quiz 99

Pardoll DM. 2012. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer* 12: 252-64

Pharoah PD, Song H, Dicks E, Intermaggio MP, Harrington P, et al. 2016. PPM1D Mosaic Truncating Variants in Ovarian Cancer Cases May Be Treatment-Related Somatic Mutations. *J Natl Cancer Inst* 108

Rajasagi M, Shukla SA, Fritsch EF, Keskin DB, DeLuca D, et al. 2014. Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood* 124: 453-62

Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, et al. 2015. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348: 124-8

Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. 2015. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160: 48-61

Rosenberg SA. 2012. Raising the bar: the curative potential of human cancer immunotherapy. *Sci Transl Med* 4: 127ps8

Ross-Innes CS, Becq J, Warren A, Cheetham RK, Northen H, et al. 2015. Whole-genome sequencing provides new insights into the clonal architecture of Barrett's esophagus and esophageal adenocarcinoma. *Nat Genet* 47: 1038-46

Roth A, Khattra J, Yap D, Wan A, Laks E, et al. 2014. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods* 11: 396-8

Ruark E, Snape K, Humburg P, Loveday C, Bajrami I, et al. 2013. Mosaic PPM1D mutations are associated with predisposition to breast and ovarian cancer. *Nature* 493: 406-10

Russo M, Siravegna G, Blaszkowsky LS, Corti G, Crisafulli G, et al. 2016. Tumor Heterogeneity and Lesion-Specific Response to Targeted Therapy in Colorectal Cancer. *Cancer Discov* 6: 147-53

Schreiber RD, Old LJ, Smyth MJ. 2011. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* 331: 1565-70

Schuh A, Becq J, Humphray S, Alexa A, Burns A, et al. 2012. Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood* 120: 4191-6

Schwarz RF, Ng CK, Cooke SL, Newman S, Temple J, et al. 2015. Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis. *PLoS Med* 12: e1001789

Segovia R, Tam AS, Stirling PC. 2015. Dissecting genetic and environmental mutation signatures with model organisms. *Trends Genet* 31: 465-74

Shah NP, Nicoll JM, Nagar B, Gorre ME, Paquette RL, et al. 2002. Multiple BCR-ABL kinase domain mutations confer polyclonal resistance to the tyrosine kinase inhibitor imatinib (STI571) in chronic phase and blast crisis chronic myeloid leukemia. *Cancer Cell* 2: 117-25

Shah SP, Roth A, Goya R, Oloumi A, Ha G, et al. 2012. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* 486: 395-9

Shain AH, Yeh I, Kovalyshyn I, Sriharan A, Talevich E, et al. 2015. The Genetic Evolution of Melanoma from Precursor Lesions. *N Engl J Med* 373: 1926-36

Shen MM. 2013. Chromoplexy: a new category of complex rearrangements in the cancer genome. *Cancer Cell* 23: 567-9

Shi H, Hugo W, Kong X, Hong A, Koya RC, et al. 2014. Acquired Resistance and Clonal Evolution in Melanoma during BRAF Inhibitor Therapy. *Cancer Discov* 4: 80-93

Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, et al. 2014. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med* 371: 2189-99

Sottoriva A, Kang H, Ma Z, Graham TA, Salomon MP, et al. 2015. A Big Bang model of human colorectal tumor growth. *Nat Genet* 47: 209-16

Sottoriva A, Spiteri I, Piccirillo SG, Touloumis A, Collins VP, et al. 2013. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci U S A*

Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, et al. 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486: 400-4

Stratton MR, Campbell PJ, Futreal PA. 2009. The cancer genome. *Nature* 458: 719-24

Swisher EM, Harrell MI, Norquist BM, Walsh T, Brady M, et al. 2016. Somatic Mosaic Mutations in PPM1D and TP53 in the Blood of Women With Ovarian Carcinoma. *JAMA Oncol* 2: 370-2

Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, et al. 2012. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *The New England journal of medicine* 366: 2443-54

Turke AB, Zejnullahu K, Wu YL, Song Y, Dias-Santagata D, et al. 2010. Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. *Cancer Cell* 17: 77-88

Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, et al. 2015. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science* 350: 207-11

Van Loo P, Campbell PJ. 2012. ABSOLUTE cancer genomics. *Nat Biotechnol* 30: 620-1

Van Loo P, Nordgard SH, Lingjaerde OC, Russnes HG, Rye IH, et al. 2010. Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci U S A* 107: 16910-5

Wagle N, Emery C, Berger MF, Davis MJ, Sawyer A, et al. 2011. Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. *J Clin Oncol* 29: 3085-96

Weston-Bell N, Gibson J, John M, Ennis S, Pfeifer S, et al. 2013. Exome sequencing in tracking clonal evolution in multiple myeloma following therapy. *Leukemia* 27: 1188-91

Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. 2016. Identification of neutral tumor evolution across cancer types. *Nat Genet* 48: 238-44

Xie M, Lu C, Wang J, McLellan MD, Johnson KJ, et al. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med* 20: 1472-8

Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, et al. 2015. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med* 21: 751-9

Yau C, Mouradov D, Jorissen RN, Colella S, Mirza G, et al. 2010. A statistical approach for detecting genomic aberrations in heterogeneous tumor samples from single nucleotide polymorphism genotyping data. *Genome Biol* 11: R92

Yip S, Miao J, Cahill DP, Iafrate AJ, Aldape K, et al. 2009. MSH6 mutations arise in glioblastomas during temozolomide therapy and mediate temozolomide resistance. *Clin Cancer Res* 15: 4622-9

Zhang J, Fujimoto J, Zhang J, Wedge DC, Song X, et al. 2014. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* 346: 256-9

Figure legends:

Figure 1: Correcting for copy number to determine cancer cell fraction. The variant allele frequency (VAF) of each mutation will depend on its local copy number. By correcting for mutation copy number and then clustering mutations, clonal and subclonal mutations can be determined.

Figure 2: Modes of tumor evolution. A) When cancer cells are subject to positive selection their frequency may increase, meaning the variant allele frequency of a mutation will not remain constant. B) During neutral evolution the frequency of a cancer cell clone will directly reflect its timing, with the variant allele frequency remaining constant.

Figure 3: Modes of clonal evolution in response to therapy. Beginning with a heterogeneous population of cells, therapy can generate a mutation (depicted in red) that may endow a subclonal population with a fitness advantage (A), select for a pre-existing resistant clone (B), result in clonal shifts without a known resistance mechanism (C), or not substantially alter the clonal composition (D).