Journal of Pathology
J Pathol 2017; **241:** 183–191
Published online 18 November 2016 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/path.4821

INVITED REVIEW

Measuring cancer evolution from the genome

Trevor A Graham[1,*] and Andrea Sottoriva[2,*]

[1] *Evolution and Cancer Laboratory, Barts Cancer Institute, Queen Mary University of London, London, UK*
[2] *Cancer Evolutionary Genomics and Modelling Laboratory, Centre for Evolution and Cancer, The Institute of Cancer Research, Sutton, UK*

**Correspondence to: TA Graham, Evolution and Cancer Laboratory, Barts Cancer Institute, Charterhouse Square, Queen Mary University of London, EC1M 6BQ, UK. E-mail: t.graham@qmul.ac.uk*
Or A Sottoriva, Cancer Evolutionary Genomics and Modelling Laboratory, Centre for Evolution and Cancer, The Institute of Cancer Research, Sutton, SM2 5NG, UK. E-mail: Andrea.Sottoriva@icr.ac.uk

Abstract

The temporal dynamics of cancer evolution remain elusive, because it is impractical to longitudinally observe cancers unperturbed by treatment. Consequently, our knowledge of how cancers grow largely derives from inferences made from a single point in time – the endpoint in the cancer's evolution, when it is removed from the body and studied in the laboratory. Fortuitously however, the cancer genome, by virtue of ongoing mutations that uniquely mark clonal lineages within the tumour, provides a rich, yet surreptitious, record of cancer development. In this review, we describe how a cancer's genome can be analysed to reveal the temporal history of mutation and selection, and discuss why both selective and neutral evolution feature prominently in carcinogenesis. We argue that selection in cancer can only be properly studied once we have some understanding of what the absence of selection looks like. We review the data describing punctuated evolution in cancer, and reason that punctuated phenotype evolution is consistent with both gradual and punctuated genome evolution. We conclude that, to map and predict evolutionary trajectories during carcinogenesis, it is critical to better understand the relationship between genotype change and phenotype change.

Keywords: clonal evolution of cancer; neutral evolution; selection; punctuated equilibrium; gradualism; saltation; hopeful monsters; subclones; intratumour heterogeneity; next-generation sequencing

Received 26 September 2016; Revised 5 October 2016; Accepted 7 October 2016

No conflicts of interest were declared.

Introduction

How do cancers grow? This basic question continues to be difficult to answer, for the obvious reason that longitudinal observation of tumour growth is nearly always impractical, both in humans and in model systems. Consequently, our understanding of tumour formation relies on historical inference based on the composition of excised tumours. In other words, our understanding of the temporal process of tumour evolution is largely derived from data collected at a single time point: the time point at the end of the process when the tumour ends up on the specimen table. However, this state of affairs is not as sorry as it may sound, as, fortunately, the tumour genome (or, more accurately, the genomes of all the cells in the tumour) provides a surreptitious but rich record of a tumour's growth.

Each time that a cell divides, errors during DNA replication mean that new mutations are introduced into the genomes of the daughter cells [1–3]. Epigenetic marks (e.g. DNA methylation) are also copied with limited fidelity [4]. Larger-scale chromosomal or part-chromosome losses or amplifications [somatic copy number alterations (SCNAs)] and other structural rearrangements also occur at an appreciable frequency in many cancers [5,6]. It is these naturally occurring (epi)genetic alterations that record the ancestry of the cells in the tumour, and, because tumours are clonally derived, all of the cells in the tumour will carry the mutations in the first cancer cell, whereas later-arising subclones are identifiable by their sharing of a particular unique set of variants; therefore, the order of clone development can be inferred by comparing the sets of mutations present in different cells of the tumour. The logic of this kind of analysis is at the heart of phylogenetics methods as applied to cancer [7]. Moreover, if a particular type of mutation accrues at a constant rate (for example, the same number of new mutations are introduced in each cell division; this appears to be the case for C > T transitions within specific three-base pair motifs, for instance - see ref. [1]), then counting the number of mutations of the type that are unique to a particular lineage gives an estimate of the relative time that the lineage arose. A constant mutation rate is referred to as a 'molecular clock', and, if the rate at which the molecular clock 'ticks' is known, then the absolute time of events (where time is measured in cell divisions elapsed)

can also be determined [8]. These methods have been applied to a wide variety of cancers and have provided new insights into the order and timing of mutation accumulation (for some examples, see [9–25]).

However, mutation is not the only force shaping the cancer genome: evolutionary selection also plays a critical role. Selection refers to the situation where one group of cells within the tumour is evolutionarily 'favoured' over another, such that the favoured cells have more offspring than the not-favoured cells. The favouring is a result of the cell evolving a new phenotypic trait that gives it an advantage in the current microenvironment (context) of the tumour; the trait is referred to as adaptive. For example, a cell with a low metabolic demand might grow faster than a cell with a high metabolic demand when both cells are together in a nutrient-poor microenvironment. The result of selection is that any mutation in the selected (favoured) population becomes more common in the tumour population as a whole, whereas negatively selected clones (not favoured) become relatively less common. Consequently, selection plays a central role in shaping the frequency distribution of mutations within a tumour.

To understand how a tumour has grown from its genome, we therefore need to understand both mutation and selection, and, critically, how these two processes together shape the pattern and frequency of mutations in the genome. Mutation and selection are deeply intertwined, as a new mutation may produce a new adaptive trait and therefore drive selection, and, conversely, a new microenvironmental selective pressure (such as targeted therapy) may mean that a pre-existing mutation becomes adaptive and so increases in frequency [26]. In general, mutation is considered to be a random process; any mutation may occur at any time with some (low and/or fluctuating) probability, whereas selection is non-random; only particular mutations are adaptive in a given context [27,28]. For example, loss of normal function of the *APC* gene provides a clear selective advantage to cells in the intestine [29–31], but not, say, to those in the lung, even though, presumably, *APC* mutation occurs at comparable rates in both tissues. Thus, the frequency at which particular mutations are observed across a tumour is a function both of the rate at which the mutations occur, and of the likelihood the mutation has of being adaptive and driving clonal expansion to a detectable level.

Tissue architecture –which we can broadly think of as the 'mechanical microenvironment' – provides additional selective constraints on tumour evolution. Many epithelia have glandular architecture (e.g. the crypts in the colon, and ducts in the prostate and breast), and it is the abnormal growth of these glands (rather than the cells within them as such) that underlies neoplastic growth. Thus, cancer development requires evolution at multiple levels in epithelial tissues [32]: in the example of the colon, first a mutated cell must repopulate the crypt, and then the mutant crypt itself must divide to form a glandular adenoma [33]. Computational modelling suggests that these tissue architectures have (at least in part) evolved to suppress clonal evolution [34]. Tissue architecture means that, in solid tumours, clonal expansions are spatially delineated, so the indicators of mutation and selection within the genome are likely to show intratumour heterogeneity.

Table 1. Definition of terms

Neutral evolution	Evolution in which all individuals in the population have equal fitness. In a growing population (such as a newly formed tumour), this means that all cells grow at the same rate
Drift	Stochastic effects (e.g. random cell death in a tumour) can cause some 'lucky' individuals in a population to have more offspring than others, so the 'lucky lineage' increases in size. Consequently, drift can cause fluctuations in subclone size in the absence of selection
Selection	The process that results in one individual in a population, because of its particular well-adapted traits, having more offspring than another less well-adapted individual
Fitness	The relative ability of an individual to produce surviving offspring in a population
Punctuated equilibrium	The process whereby apparently abrupt changes in the phenotype of the population at large occur because of gradual evolution in small, spatially isolated niches
Hopeful monster/saltation	The process whereby abrupt changes in phenotype are cause by underlying (large-scale) punctuated changes in the genome. In cancer, massive genome alterations occurring in a single cell division are examples of saltatory genome evolution

It is clear that cancer formation requires the acquisition of a number of key driver alterations (mutations and epigenetic changes in cancer cells) [35,36]. The precise number of drivers per cancer is uncertain – and indeed, given the inherently contextual nature of selection, a comprehensive list of cancer-specific drivers is unlikely to exist in reality. It serves our purposes here to think of driver mutations as mutations that are positively selected in their (changing) microenvironmental context within the tumour. A central question in tumour evolution is: what is the temporal pattern of driver alteration acquisition? The competing theories are gradualism and punctuation (see Table 1 for a definition of terms). The gradualist theory proposes that cancer evolution happens via a steady accumulation of driver mutations and a concomitant steady series of selective clonal outgrowths, whereas the punctuated theory proposes that the evolution of cancer occurs in fits and starts.

In this review, we address how mutation and selection together shape the cancer genome, with particular reference to the manifestations of graduated and punctuated evolution.

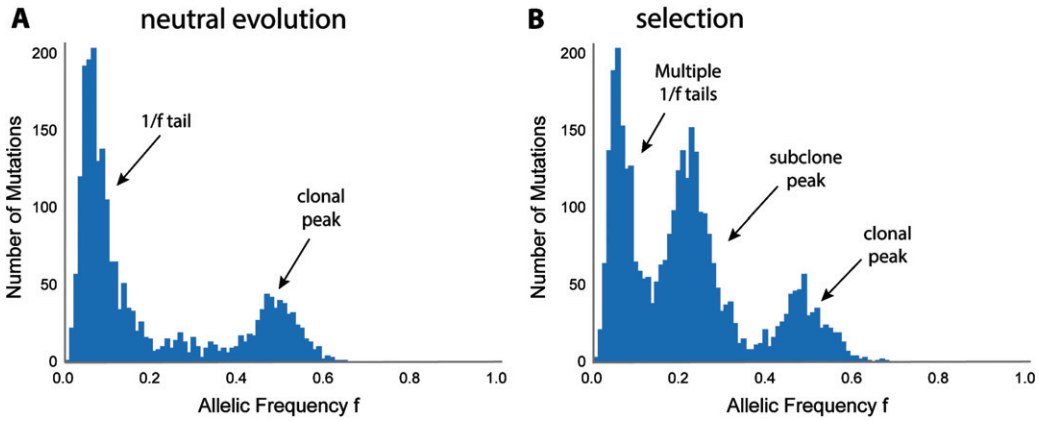

Figure 1. The influence of selection in the cancer genome. (A) A simulated distribution of VAFs in a cancer that is evolving neutrally. VAF distributions are naturally produced by next-generation genome sequencing. The distribution has a peak around 0.5 – these are the clonal variants present in all cancer cells. The distribution of subclonal variants (those at lower frequencies) follows a $1/f$ distribution, whereby there are ever more mutations at ever lower frequency. Neutral cancer evolution can be detected by comparing the observed distribution of mutation frequencies in a cancer with this expected theoretical distribution. (B) VAF distribution for an *in silico* model of a cancer in which a fitter subclone has clonally expanded within the tumour. The cluster of mutations within the subclone are 'passenger' mutations dragged along to higher frequency within the tumour during the clonal expansion. Even though there is selection, there is still a 'background' of neutral evolution – this is the neutral evolution within the selected subclone and in the residual tumour cells.

Detecting selection and neutrality from the cancer genome

Clonal selection, whatever the biological mechanism driving it, ultimately results in the relative outgrowth of the selected clone within the tumour. The clonal outgrowth appears in the cancer's genome via an 'over-representation' of the mutations in the selected clone, as compared with the 'null' case, where the genome evolved in the absence of selection (Figure 1). In principle, then, detecting selection just requires spotting the characteristic 'clonal outgrowths', and many different bioinformatics tools have been developed to spot the 'clusters' of mutations at similar frequency in tumour next-generation sequencing data that are characteristic of these outgrowths [22,37–39]. It is important to note here that the evolutionary dynamics of the selected clone are largely revealed by the passenger mutations in that clone, not the drivers themselves: as the selected clone grows out, all of the many passenger mutations in the clone are carried along to higher frequency, making the selected clone visible against the milieu of unselected mutations in the tumour. Therefore, both driver and passenger mutations in the clone are affected by selection, but passenger mutations are generally more informative, as they are more numerous [40]. This is just because, as evolution is a blind force, for every 'successful' driver mutation, many 'unsuccessful' mutations have occurred in a genome as large as the human one.

Moreover, this means that clonal selection is always visible in the frequency distribution of mutations in cancer, irrespective of the biological mechanism that provides the selective advantage. For example, suppose that, rather than acquiring a new driver mutation, a clone gains a selective advantage because of a sudden change in microenvironmental context (such as a new non-cell-autonomous interaction within the tumour [41]); even though the clone's advantage is cell-extrinsically driven, its passenger mutations will still become over-represented.

However, we argue that, to be able to reliably spot clonal outgrowths in a cancer's genome, we first require an understanding of what the 'null case' – evolution in the absence of clonal selection – looks like. The absence of clonal selection is referred to as 'neutral evolution', and, by definition, neutral evolution (in a growing population such as a tumour) is the case when all cells grow at the same rate. The definition of neutral evolution also encompasses stochastic drift. In a drifting population, all cells grow at the same average rate, but, at any single point in time, any one lineage might, because of random effects, grow or shrink slightly faster than another. We note that, if a 'lucky' clone happens to drift to proportionally high frequency in a neutrally evolving asexual population such as cancer, it could appear indistinguishable from a selected clone.

Mathematical modelling (or perhaps more accurately put: population genetics theory) provides a formal description of the frequency of subclonal mutations within a neutrally growing tumour [42–44]. Under neutrality, the cumulative number of mutations at frequency f follows a '$1/f$' distribution: this means that the number of mutations at a particular frequency in the tumour will double each time that the frequency halves, or, more loosely, when a tumour is growing neutrally there will be ever more mutations at ever lower frequencies. This mathematical result can be understood intuitively: because a small number of new mutations are expected to accrue each time that a cell divides, then, as the tumour population increases in size, more and more new mutations are accrued by the population as a whole, and the $1/f$ distribution is reached because precisely twice as many new mutations are expected each time

that the population doubles in size. To test for selection in a growing cancer, it therefore suffices to ask whether or not the distribution of mutation frequencies observed in the cancer (as measured by next-generation sequencing) follows a $1/f$ distribution: in the case where it does not, we can reject the null hypothesis of neutrality in favour of recent selection.

This $1/f$ test must be applied with caution. Limited depth sequencing can blur the signal from the evolutionary dynamics (the signals of both selection and neutrality alike) [42]. It is conceivable that a particular 'just-right' combination of subclones could produce a variable allele frequency (VAF) distribution that masquerades as a $1/f$ distribution (if the selected subclones happened to reach a particular set of sizes and the 'noise' in the sequencing data blurred their passenger mutation VAFs appropriately). Although the allele frequencies alone cannot discount this possibility, neutral evolution nevertheless provides a much more parsimonious explanation of a $1/f$-like VAF distribution. Moreover, we note that the $1/f$ test provides an objective indication of the presence or absence of subclonal selection, which does not require prior knowledge of the identity of subclonal drivers.

The ratio of non-synonymous (NS) mutations (which are likely to alter fitness by changing protein structure and function) to synonymous (S) mutations (which are likely to be neutral) at a particular locus is another popular test for selection [45,46]. Typically, the NS/S ratio is normalized by the number of possible NS and S mutations that can occur at the locus of interest (the normalized ratio is referred to as dN/dS), and then deviations in the normalized ratio above 1 indicate positive selection (more NS mutations than expected by chance), whereas deviations below 1 indicate negative selection (fewer NS mutations than expected by chance). Applying dN/dS to cancer is complicated by the differential and evolving mutation rates of three-base pair motifs [47], which can potentially skew the dN/dS values, but, nevertheless, corrected dN/dS values within large cohorts of tumours has revealed evidence of positive selection in cancer for particular gene sets, such as the kinases [48]. We note that applying dN/dS to detect subclonal selection within a tumour is extremely challenging, because, if the selection is caused by a single base pair change (e.g. the common *KRAS* c.35G > T mutation), then the signal from this locus will be 'drowned out' by all the other passenger mutations within the clone, and applying dN/dS on a gene-by-gene basis in individual tumours is not possible, because of the relatively low numbers of detected somatic mutations in any individual cancer.

How often does selection occur?

We recently looked for evidence of clonal selection across cancer types by using the $1/f$ test described above. Remarkably, our analysis showed that, in ~30% of cancers of 14 different solid cancer types, we were unable to reject the null hypothesis of neutral evolution [42]. In cancer model systems, neutral drift of tumour cells is also observed [49]. Therefore, the signature of selection appears to be somewhat rarer than we might naively have expected from a gradualistic evolutionary perspective.

How often should we expect to see selection manifested in the cancer genome?

First, we only expect to see clonal outgrowth if the clone is 'caught in the act' of growing out – e.g. if the clone had already expanded to at least a minimal detectable size at the time when the tumour was sampled, but it has not yet expanded to repopulate the entire tumour [50]. The latter point is important because once a selected clone has taken over the whole tumour then all the cells within the clone are the same as one another – and so the population then again evolves neutrally. The duration of time during which a clone can be 'caught in the act' of expanding is determined by the selective advantage of the clone (relative to the residual 'host' cells in the tumour): fitter clones will grow out more quickly. Unfortunately, empirical measurements of selective advantages of clones in growing tumours are lacking, so our expectations of the likelihood of detecting a selected clone mid-expansion are largely based on guesswork. Within the intestinal crypt (which has a constant population size), empirical measurement of the selective advantages of the tumour suppressor gene *APC* and the proto-oncogene *KRAS* reveal almost 50% increases in the probability of stem cell replacement [29]. If tumour subclones experienced similarly large selective advantages, we might expect to only rarely see partially expanded clones. Intriguingly, however, abstract mathematical modelling of mutation accumulation in growing tumours suggests that very low selective advantages for new driver mutations (of the order of <1%) lead to reasonably long times before cancer occurs (in the models, cancer is defined by a subclone having accumulated a critical driver mutation burden) [51,52]. In addition, our own computational modelling shows that even sizeable selective advantages produce only slight changes in clone frequency in a growing population, and this result is even more pronounced when a new clone is formed in an already large tumour [14]. Together, these results would predict that partially expanded subclones would be commonplace if they were initiated at a sufficiently high rate. Clearly, empirical measurement of the differential fitness of tumour subclones is required.

Second, the likelihood of seeing selection is also determined by the rate at which new selected clones are generated, either by clone-intrinsic mutation or by the creation of a favourable microenvironment. This rate is directly related to the number of potential driver alterations that a clone can acquire: if there are many potential drivers, then new driver mutations are likely to occur frequently. Interestingly, genome sequencing studies on large cohorts (such as the Cancer Genome

Atlas; http://cancergenome.nih.gov/) consistently reveal fairly short lists of recurrently mutated genes in each cancer type: for example, in a cohort of 276 colorectal cancers, only 24 genes were mutated at a significantly greater than background frequency [53]. These studies suggest that the number of drivers may actually be quite limited, and hence neutral dynamics may be relatively common in cancers because of the low rate of driver mutation accrual.

Third, our ability to detect selection is, of course, limited by the resolution of our tools for finding it. The current standard of moderate-depth exome sequencing is 100× coverage, facilitating reasonably reliable detection of mutations at ~5% frequency. This means that low selective advantages that cause only slight changes in clone frequency are largely indistinguishable from the background neutral evolution. The 'mini-driver' hypothesis, which postulates that there are many mutations that each cause small fitness effects in cancers [54], would clearly be challenging to confirm or refute from moderate-depth sequencing data. Moreover, as a tumour grows, newly generated clones form ever lower proportions of the tumour cell population, so detecting them becomes ever more challenging as the tumour becomes larger: thus the ~5% sensitivity of sequencing provides a window to detect only those clones that form very early in a cancer's growth, or those that rapidly (e.g. within a small proportion of the lifetime of the cancer) grow to a detectable size.

Evolutionary dynamics and tumour progression

Manifest ongoing selection in cancers appears to be associated with a worse prognosis, because, across cancer types, tumours with three or more large clones have a worse prognosis than tumours with fewer clones [55,56], and putative subclonal driver mutations are also associated with a worse prognosis [55].

However, neutral evolution has a potential 'dark side' for prognosis, by virtue of it allowing huge variation to be generated and persist in a tumour. Whereas, by definition, the diversity in a neutrally evolving tumour is non-adaptive to the current microenvironment, if the microenvironment were to change – through the application of therapy, for instance – then variants within this reservoir of pre-existing variation could suddenly become adaptive. Thus, neutrally evolving tumours may be particularly prone to developing therapy resistance. The relationship between neutral and selective evolutionary dynamics and tumour progression should be the focus of future work.

Our recent analysis of the evolution of colorectal cancer led us to put forward the 'Big Bang' model of cancer growth, whereby the tumour mass grows as a single clonal expansion wherein differential clonal selection within the tumour has little influence on the subclonal composition of the tumour, and instead clonal mosaicism is determined largely by the time of clone generation [14]. As the clonal composition of a Big Bang tumour is determined simply by which clones were generated at the beginning of cancer growth, we speculated that a tumour's prognosis is similarly predetermined. In other words, in the absence of clonal selection, the phenotype of the 'first' cancer cells should determine the cancer's behaviour thereafter. Consequently, we speculate that reading these 'initial phenotypes' in the grown cancer may be prognostic, e.g. by looking for the degree of clonal mixing as a readout of cell migration ability, or the degree to which a clone coexists in multiple different microenvironments as a readout of plasticity.

Punctuated evolution

In the evolutionary biology literature, punctuated evolution is (loosely) defined as an apparently abrupt change in phenotype, and was originally suggested by Eldredge and Gould to explain the large morphological differences between species that appeared to occur without the presence of intermediate morphotypes in the fossil record [57]. The original descriptions of punctuated evolution put forward the idea that an ancestral species became subdivided into (spatially) isolated distinct niches, in which each subpopulation independently (and gradually) evolved until the point when one of those subspecies – by now grossly altered as compared with the ancestor – was able to escape the niche and expand its population significantly (Figure 2A). Because the isolating niche was small, the intermediate forms were lost to the fossil record, and only the widespread ancestral and then the new grossly altered populations were captured. The result was an apparently punctuated evolution of species – interspersed by long periods of time during which apparently no 'important evolution' happened. This pattern of events was described as punctuated equilibrium.

Punctuated equilibrium has frequently been conflated with saltation theory, although the two theories are distinct. The difference is that the two theories describe punctuated phenotype change and punctuated genotype change, respectively. Saltation theory suggests that new species can be generated rapidly because of sudden large-scale mutation(s) – in other words, the underlying genetic evolution causing the speciation event is itself punctuated (Figure 2B). Punctuated equilibrium, on the other hand, proposes that gross phenotypic change is the consequence of gradual (although perhaps rapid) genetic evolution in an isolated population. Richard Goldschmidt described the gross mutations as hopeful monsters – striving for 'perfection' in one big jump [58]; however, it is more likely that most gross genetic rearrangements will be maladaptive.

It is increasingly clear that the punctuated evolution of both phenotypes and genotypes occurs during cancer development.

Punctuated phenotype change is clearly seen in the development of neoplasia: typically, a neoplastic lesion

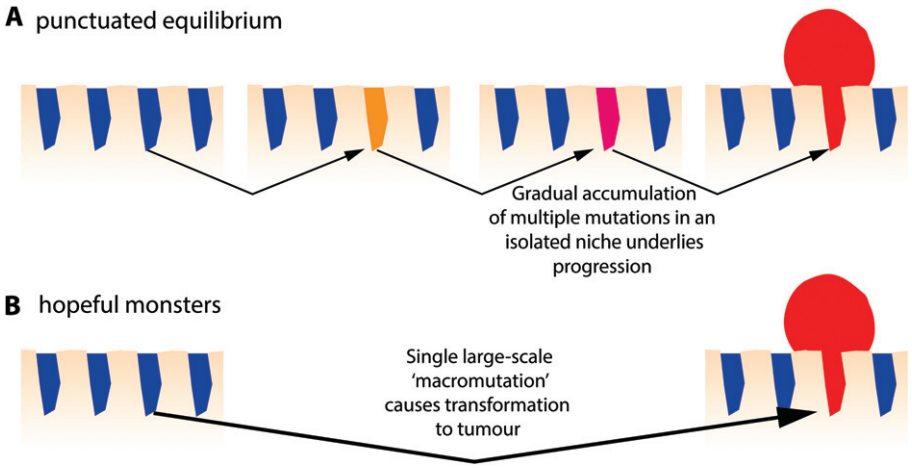

Figure 2. Punctuated equilibrium and hopeful monsters. (A) Mutations accumulate within small spatially isolated niches (here, an intestinal crypt is depicted), and only after a sufficient number of adaptive mutations are acquired is clonal expansion initiated. From a macroscopic perspective, the evolution of the neoplasia appears to be punctuated, even though the driver mutations were acquired gradually within the crypt. (B) The generation of a hopeful monster – a clone with a grossly altered genotype – in a single cell division produces a neoplasm in a single 'catastrophic' step.

(such as a colorectal adenoma) – with a grossly different phenotype to that of the normal cells – arises 'abruptly' without intermediate partial neoplastic forms (although we acknowledge that one could argue that crypt hyperplasia may sometimes be an intermediate form in the intestine). It is important to recognize that such punctuated phenotype change (normal to malignant cells) may be underpinned by gradual genotype evolution. In the example of intestinal neoplasia, it is clear that loss of the normal function of the *APC* gene is sufficient to generate adenomas [30,31]. Loss of normal *APC* function can be caused by the 'gradual' accumulation of the two mutational hits on each of the *APC* alleles [59], and it is clear that the gross changes in phenotype (normal to neoplastic) need not be accompanied by large-scale genetic alteration [60]. In leukaemia, 'intermediate' clone genotypes are present at only very low frequency, potentially indicating punctuated equilibrium-like evolutionary dynamics [61]. In follicular lymphoma, disease transformation is associated with an increased mutation burden, and often also the acquisition of mutations in key 'driver' genes, although the underlying temporal pattern of mutation accumulation remains undetermined [21].

The genotype–phenotype map describes the relationship between genetic change and phenotypic traits. *APC* loss in the intestine demonstrates how slight changes in genotype (e.g. single base pair changes) can cause large changes in phenotype: this is an example where mapping between the space of possible genotypes and phenotypes is not smooth. Moreover, phenotypic change may not occur until multiple independent mutations in a number of key genes have accumulated and act in tandem to cause phenotypic alteration (this is called epistasis). Epistasis can underlie punctuated equilibrium in cancer – an individual lineage may steadily acquire individual driver mutations but not clonally expand until it has the full complement of drivers necessary to enhance its fitness. Intriguingly, in colorectal [62,63] and lung [20] cancers, the majority of the driver mutations often appear to be clonal throughout the cancer, perhaps implying that the growth of these cancers can be initiated only when a complete epistatically-interacting complement of drivers is obtained. Epistasis clearly adds much complexity to the relationship between genotypes and phenotypes. Resolving the genotype–phenotype map is key to understanding evolutionary trajectories in cancer, although, given the inherently contextual definition of phenotypes and the near-infinite space of possible genotypes, the resolution will be extremely challenging to achieve.

It is increasingly clear that Goldschmidt's hopeful monsters – punctuated changes in genotype – are frequently found in cancer (Figure 3). Chromothripsis – chromosome shattering and reassembly in an aberrant manner – has now been reported in many different cancer types [64–66], and this saltatory mutation occurs following a single 'catastrophic' mitosis [67]. Chromoplexy – the interleaving of different chromosomal regions into one aberrant block – has been reported in prostate cancer, and is likely to occur in a single cell division [68]. More generally, genome doubling is a relatively common saltatory mutation type observed across cancer types [69], and, furthermore, the tolerance of genome doubling facilitates subsequent chromosomal instability [70]. In breast cancer, sequencing of individual nuclei detects clones with dramatic copy number deviation from the diploid genome, and no evidence of cells with intermediate patterns of copy number alteration [71,72] (Figure 3). Similarly relatively homogeneous intratumour patterns of grossly deviant copy number alterations are observed in many cancer types, including colorectal [62] and ovarian [17] cancer and the premalignant disease Barrett's oesophagus [25,73], suggesting underlying saltatory mutational mechanisms.

How often are these hopeful monsters formed? The monsters that we sample in cancer are the

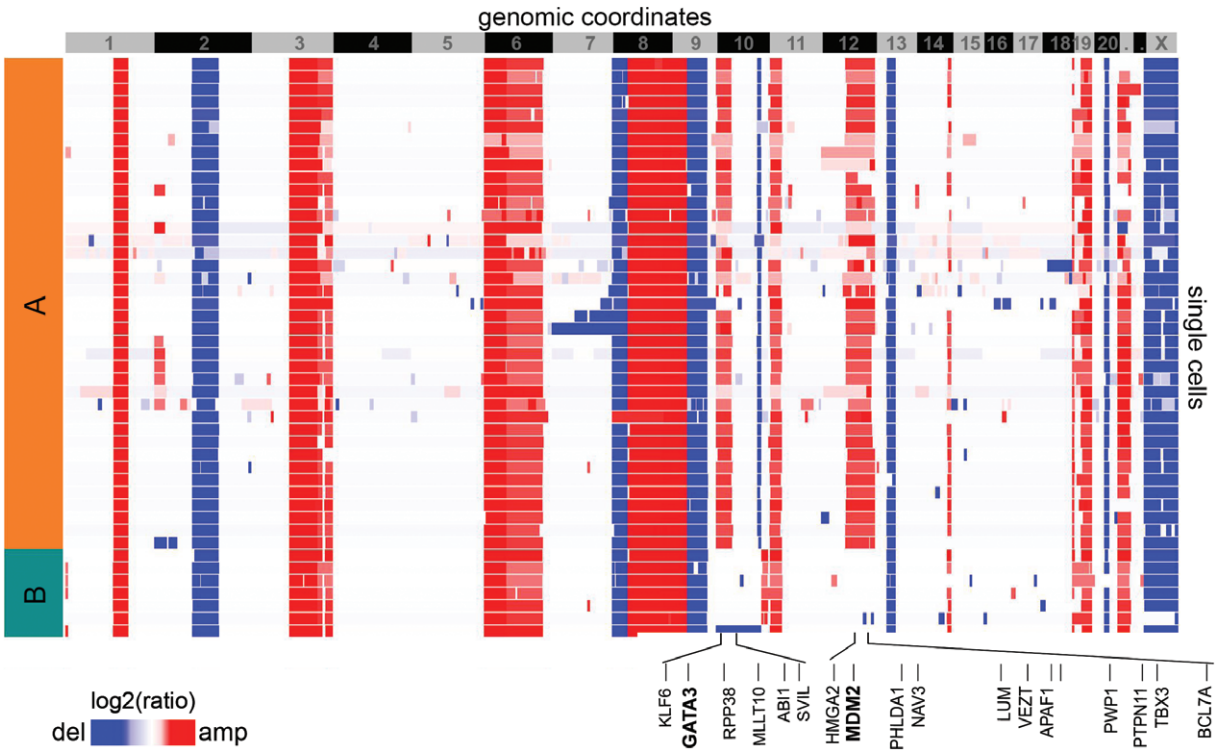

Figure 3. Primary data indicating punctuated copy number evolution. Copy number profiles of individual breast cancer cells from a single breast cancer case showing the same grossly altered genomes are observed in all cells sampled, and no intermediate forms are detected. Image scaled, cropped and reprinted with permission from [72].

ones that have fortuitously stumbled upon an adaptive genotype–phenotype combination. However, many saltatory mutations presumably lead to altered maladapted phenotypes, or phenotypes that are lethal – indeed, this is often the case for chromosomal instability [74]. Logically, therefore, this means that, for every saltatory mutation that produces an adaptive phenotype, there are probably many more saltatory mutations that produce maladapted phenotypes. This likely abundance of 'maladapted monsters' in cancer types (or their premalignant precursors) that frequently show saltatory mutation is a testable prediction, and, furthermore, the detection of maladapted monsters could prove to be a useful prognostic biomarker in premalignant diseases such as Barrett's oesophagus, in which large-scale genome alteration appears to be a key punctuated event in cancer formation [25,73].

An important aside is the potential for punctuated evolution of the *rate* of single-nucleotide alterations (SNAs). SNAs accumulate according to a relatively small number of underlying mutational processes associated with natural replication errors, defective DNA replication and repair machinery, and mutagen exposures [47]. Although the accumulation of SNAs is a clearly a gradual process (although, as noted above, individual SNAs can cause punctuated change in phenotype), we note that the abrupt 'switching on' of a new mutational process can cause punctuated changes in the SNA mutation rate. For example, mutation of the mismatch repair machinery causes a sudden increase in a cell's point mutation rate [75].

Conclusion: neutrality and selection, and punctuation and gradualism, are each two sides of the same coin

Cancer genomes show frequent evidence of both neutral evolution and clonal selection. As neutral evolution is just the evolution that happens between selection events – e.g. the evolution that happens within a clone – the frequent detection of neutral evolution in cancer should come as no surprise. It is our opinion that, in fact, it would be more surprising if a signature of neutral evolution was never seen in cancer, because this would mean that new 'driver' mutations accrue all the time in our cells – an implication that appears to be at odds with the relatively low age-dependent incidence of cancer [76] and the small number of drivers with respect to passengers [40].

At a molecular level, cancers unquestionably show both gradual (the steady accumulation of single-nucleotide variants) and punctuated (large-scale copy number alterations) genotype change. However, whether or not phenotype change is similarly punctuated depends on the relationship between the genotype and phenotype, and also on the microenvironment context. Thus, to be able to predict and manipulate the evolutionary trajectories of cancer for respective prognostic and therapeutic benefits, it is critical that we understand the genotype–phenotype map and the associated transitions around genotype–phenotype space. To achieve this, we critically need to understand exactly which phenotypes

in cancer are selected and why – genetics help us to understand the accessibility of the space of different phenotypes, but genetics alone cannot give us the full picture of cancer evolution.

Finally, a word of caution is necessary. Studies have shown that treatment frequently selects for rare subclones in a tumour – and sometimes subclones that were so rare that they went undetected in the pretreatment samples [26,77–80]. Thus, although the evolutionary dynamics of large tumour subclones – the focus of this review – are clearly of much interest for understanding the basic biology of cancer evolution, we must ask ourselves if these dynamics directly relate to a patient's prognosis. It is our conviction that these evolutionary dynamics are clinically relevant, because only by learning the 'rules of cancer evolution' can we hope to effectively intervene and change the evolutionary course.

Author contributions statement

TG and AS co-wrote the manuscript.

Acknowledgements

TG and AS are grateful for support from the Wellcome Trust. TG is also a Cancer Research UK Career Development Fellow. AS is also supported by the Chris Rokos Fellowship in Evolution and Cancer. The funders had no influence on the content of this review. The authors are grateful to Marc Williams for providing images for Figure 1, and to Nicholas Navin for providing Figure 3.

References

1. Alexandrov LB, Jones PH, Wedge DC, *et al.* Clock-like mutational processes in human somatic cells. *Nat Genet* 2015; **47:** 1402–1407.
2. Xue Y, Wang Q, Long Q, *et al.* Human Y chromosome base-substitution mutation rate measured by direct sequencing in a deep-rooting pedigree. *Curr Biol* 2009; **19:** 1453–1457.
3. 1000Genomes Project Consortium, Abecasis GR, Altshuler D, *et al.* A map of human genome variation from population-scale sequencing. *Nature* 2010; **467:** 1061–1073.
4. Laird CD, Pleasant ND, Clark AD, *et al.* Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc Natl Acad Sci USA* 2004; **101:** 204–209.
5. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998; **396:** 643–649.
6. Rajagopalan H, Lengauer C. Aneuploidy and cancer. *Nature* 2004; **432:** 338–341.
7. Schwarz RF, Trinh A, Sipos B, *et al.* Phylogenetic quantification of intra-tumour heterogeneity. *PLoS Comput Biol* 2014; **10:** e1003535.
8. Shibata D, Tavare S. Counting divisions in a human somatic cell tree: how, what and why? *Cell Cycle* 2006; **5:** 610–614.
9. de Bruin EC, McGranahan N, Mitter R, *et al.* Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *Science* 2014; **346:** 251–256.
10. Gerlinger M, Horswell S, Larkin J, *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* 2014; **46:** 225–233.
11. Gerlinger M, Rowan AJ, Horswell S, *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012; **366:** 883–892.
12. Thirlwell C, Will OCC, Domingo E, *et al.* Clonality assessment and clonal ordering of individual neoplastic crypts shows polyclonality of colorectal adenomas. *Gastroenterology* 2010; **138:** 1441–1454.
13. Siegmund K, Marjoram P, Woo Y, *et al.* Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. *Proc Natl Acad Sci USA* 2009; **106:** 4828–4833.
14. Sottoriva A, Kang H, Ma Z, *et al.* A Big Bang model of human colorectal tumor growth. *Nat Genet* 2015; **47:** 209–216.
15. Jones S, Chen W-D, Parmigiani G, *et al.* Comparative lesion sequencing provides insights into tumor evolution. *Proc Natl Acad Sci USA* 2008; **105:** 4283–4288.
16. Yachida S, Jones S, Bozic I, *et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* 2010; **467:** 1114–1117.
17. Schwarz RF, Ng CK, Cooke SL, *et al.* Spatial and temporal heterogeneity in high-grade serous ovarian cancer: a phylogenetic analysis. *PLoS Med* 2015; **12:** e1001789.
18. Sottoriva A, Spiteri I, Piccirillo SG, *et al.* Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. *Proc Natl Acad Sci USA* 2013; **110:** 4009–4014.
19. Sottoriva A, Spiteri I, Shibata D, *et al.* Single-molecule genomic data delineate patient-specific tumor profiles and cancer stem cell organization. *Cancer Res* 2013; **73:** 41–49.
20. Zhang J, Fujimoto J, Zhang J, *et al.* Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science* 2014; **346:** 256–259.
21. Okosun J, Bodor C, Wang J, *et al.* Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet* 2014; **46:** 176–181.
22. Nik-Zainal S, Van Loo P, Wedge DC, *et al.* The life history of 21 breast cancers. *Cell* 2012; **149:** 994–1007.
23. Yates LR, Gerstung M, Knappskog S, *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med* 2015; **21:** 751–759.
24. Gundem G, Van Loo P, Kremeyer B, *et al.* The evolutionary history of lethal metastatic prostate cancer. *Nature* 2015; **520:** 353–357.
25. Li X, Galipeau PC, Paulson TG, *et al.* Temporal and spatial evolution of somatic chromosomal alterations: a case-cohort study of Barrett's esophagus. *Cancer Prev Res (Phila)* 2014; **7:** 114–127.
26. Diaz LA, Jr, Williams RT, Wu J, *et al.* The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* 2012; **486:** 537–540.
27. Sieber OM, Tomlinson SR, Tomlinson IPM. Tissue, cell and stage specificity of (epi)mutations in cancers. *Nat Rev Cancer* 2005; **5:** 649–655.
28. Lipinski KA, Barber LJ, Davies MN, *et al.* Cancer evolution and the limits of predictability in precision cancer medicine. *Trends Cancer* 2016; **2:** 49–63.
29. Vermeulen L, Morrissey E, van der Heijden M, *et al.* Defining stem cell dynamics in models of intestinal tumor initiation. *Science* 2013; **342:** 995–998.
30. Lamlum H, Papadopoulou A, Ilyas M, *et al.* APC mutations are sufficient for the growth of early colorectal adenomas. *Proc Natl Acad Sci USA* 2000; **97:** 2225–2228.
31. Sansom OJ, Reed KR, Hayes AJ, *et al.* Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev* 2004; **18:** 1385–1390.
32. Aktipis CA, Boddy AM, Gatenby RA, *et al.* Life history trade-offs in cancer evolution. *Nat Rev Cancer* 2013; **13:** 883–892.
33. Humphries A, Wright NA. Colonic crypt organization and tumorigenesis. *Nat Rev Cancer* 2008; **8:** 415–424.

34. Pepper JW, Sprouffske K, Maley CC. Animal cell differentiation patterns suppress somatic evolution. *PLoS Comput Biol* 2007; **3:** e250.
35. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009; **458:** 719–724.
36. Vogelstein B, Papadopoulos N, Velculescu VE, *et al.* Cancer genome landscapes. *Science* 2013; **339:** 1546–1558.
37. Roth A, Khattra J, Yap D, *et al.* PyClone: statistical inference of clonal population structure in cancer. *Nat Methods* 2014; **11:** 396–398.
38. Andor N, Harness JV, Muller S, *et al.* EXPANDS: expanding ploidy and allele frequency on nested subpopulations. *Bioinformatics* 2014; **30:** 50–60.
39. Fischer A, Vazquez-Garcia I, Illingworth CJ, *et al.* High-definition reconstruction of clonal composition in cancer. *Cell Rep* 2014; **7:** 1740–1752.
40. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009; **458:** 719–724.
41. Marusyk A, Tabassum DP, Altrock PM, *et al.* Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* 2014; **514:** 54–58.
42. Williams MJ, Werner B, Barnes CP, *et al.* Identification of neutral tumor evolution across cancer types. *Nat Genet* 2016; **48:** 238–244.
43. Bozic I, Gerold JM, Nowak MA. Quantifying clonal and subclonal passenger mutations in cancer evolution. *PLoS Comput Biol* 2016; **12:** e1004731.
44. Durrett R. Population genetics of neutral mutations in exponentially growing cancer cell populations. *Ann Appl Probab* 2013; **23:** 230–250.
45. Ostrow SL, Barshir R, DeGregori J, *et al.* Cancer evolution is associated with pervasive positive selection on globally expressed genes. *PLoS Genet* 2014; **10:** e1004239.
46. Martincorena I, Roshan A, Gerstung M, *et al.* Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 2015; **348:** 880–886.
47. Alexandrov LB, Nik-Zainal S, Wedge DC, *et al.* Signatures of mutational processes in human cancer. *Nature* 2013; **500:** 415–421.
48. Greenman C, Stephens P, Smith R, *et al.* Patterns of somatic mutation in human cancer genomes. *Nature* 2007; **446:** 153–158.
49. Driessens G, Beck B, Caauwe A, *et al.* Defining the mode of tumour growth by clonal analysis. *Nature* 2012; **488:** 527–530.
50. Merlo LMF, Pepper JW, Reid BJ, *et al.* Cancer as an evolutionary and ecological process. *Nat Rev Cancer* 2006; **6:** 924–935.
51. Beerenwinkel N, Antal T, Dingli D, *et al.* Genetic progression and the waiting time to cancer. *PLoS Comput Biol* 2007; **3:** e225.
52. Bozic I, Antal T, Ohtsuki H, *et al.* Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci USA* 2010; **107:** 18545–18550.
53. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012; **487:** 330–337.
54. Castro-Giner F, Ratcliffe P, Tomlinson I. The mini-driver model of polygenic cancer evolution. *Nat Rev Cancer* 2015; **15:** 680–685.
55. Andor N, Graham TA, Jansen M, *et al.* Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat Med* 2016; **22:** 105–113.
56. Morris LG, Riaz N, Desrichard A, *et al.* Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget* 2016; **7:** 10051–10063.
57. Eldredge N, Gould SJ. Punctuated equilibria: an alternative to phyletic gradualism. In *Models in Paleobiology*, Schopf T (ed.), Freeman, Cooper & Co., San Francisco, 1972; 82–115.
58. Goldschmidt R. *The Material Basis of Evolution.* Yale University Press: New Haven, 1940.
59. Rowan AJ, Lamlum H, Ilyas M, *et al.* APC mutations in sporadic colorectal tumors: a mutational 'hotspot' and interdependence of the 'two hits'. *Proc Natl Acad Sci USA* 2000; **97:** 3352–3357.
60. Jones A, Thirlwell C, Howarth K, *et al.* Analysis of copy number changes suggests chromosomal instability in a minority of large colorectal adenomas. *J Pathol* 2007; **213:** 249–256.
61. Campbell PJ, Pleasance ED, Stephens PJ, *et al.* Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proc Natl Acad Sci USA* 2008; **105:** 13081–13086.
62. Kim TM, Jung SH, An CH, *et al.* Subclonal genomic architectures of primary and metastatic colorectal cancer based on intratumoral genetic heterogeneity. *Clin Cancer Res* 2015; **21:** 4461–4472.
63. Jesinghaus M, Wolf T, Pfarr N, *et al.* Distinctive spatiotemporal stability of somatic mutations in metastasized microsatellite-stable colorectal cancer. *Am J Surg Pathol* 2015; **39:** 1140–1147.
64. Stephens PJ, Greenman CD, Fu B, *et al.* Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 2011; **144:** 27–40.
65. Kloosterman WP, Hoogstraat M, Paling O, *et al.* Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. *Genome Biol* 2011; **12:** R103.
66. Molenaar JJ, Koster J, Zwijnenburg DA, *et al.* Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* 2012; **483:** 589–593.
67. Zhang CZ, Spektor A, Cornils H, *et al.* Chromothripsis from DNA damage in micronuclei. *Nature* 2015; **522:** 179–184.
68. Baca SC, Prandi D, Lawrence MS, *et al.* Punctuated evolution of prostate cancer genomes. *Cell* 2013; **153:** 666–677.
69. Carter SL, Cibulskis K, Helman E, *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol* 2012; **30:** 413–421.
70. Dewhurst SM, McGranahan N, Burrell RA, *et al.* Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Discov* 2014; **4:** 175–185.
71. Wang Y, Waters J, Leung ML, *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* 2014; **512:** 155–160.
72. Gao R, Davis A, McDonald TO, *et al.* Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet* 2016; **48:** 1119–1130.
73. Stachler MD, Taylor-Weiner A, Peng S, *et al.* Paired exome analysis of Barrett's esophagus and adenocarcinoma. *Nat Genet* 2015; **47:** 1047–1055.
74. Gordon DJ, Resio B, Pellman D. Causes and consequences of aneuploidy in cancer. *Nat Rev Genet* 2012; **13:** 189–203.
75. Tsao JL, Yatabe Y, Salovaara R, *et al.* Genetic reconstruction of individual colorectal tumor histories. *Proc Natl Acad Sci USA* 2000; **97:** 1236–1241.
76. Armitage P, Doll R. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br J Cancer* 1954; **8:** 1–12.
77. Ding L, Ley TJ, Larson DE, *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012; **481:** 506–510.
78. Anderson K, Lutz C, van Delft FW, *et al.* Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature* 2011; **469:** 356–361.
79. Findlay JM, Castro-Giner F, Makino S, *et al.* Differential clonal evolution in oesophageal cancers in response to neo-adjuvant chemotherapy. *Nat Commun* 2016; **7:** 11111.
80. Siravegna G, Mussolin B, Buscarino M, *et al.* Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. *Nat Med* 2015; **21:** 795–801.