Biochimica et Biophysica Acta 1867 (2017) 151–161

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbacan

Tumor evolution: Linear, branching, neutral or punctuated?☆

Alexander Davis [a,b], Ruli Gao [a], Nicholas Navin [a,b,c,*]

[a] Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[b] Graduate School of Biomedical Sciences, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[c] Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

ARTICLE INFO

Article history:
Received 19 October 2016
Received in revised form 14 January 2017
Accepted 16 January 2017
Available online 19 January 2017

Keywords:
Single cell genomics
Intratumor heterogeneity
Tumor evolution
Cancer genomics
Genome evolution
Cancer biology

ABSTRACT

Intratumor heterogeneity has been widely reported in human cancers, but our knowledge of how this genetic diversity emerges over time remains limited. A central challenge in studying tumor evolution is the difficulty in collecting longitudinal samples from cancer patients. Consequently, most studies have inferred tumor evolution from single time-point samples, providing very indirect information. These data have led to several competing models of tumor evolution: linear, branching, neutral and punctuated. Each model makes different assumptions regarding the timing of mutations and selection of clones, and therefore has different implications for the diagnosis and therapeutic treatment of cancer patients. Furthermore, emerging evidence suggests that models may change during tumor progression or operate concurrently for different classes of mutations. Finally, we discuss data that supports the theory that most human tumors evolve from a single cell in the normal tissue. This article is part of a Special Issue entitled: Evolutionary principles - heterogeneity in cancer?, edited by Dr. Robert A. Gatenby.

1. Introduction

Tumor evolution begins when a single cell in the normal tissue transforms and expands to form a tumor mass. During this complex biological process, clonal lineages diverge and form distinct subpopulations, resulting in intratumor heterogeneity (ITH). ITH has long been observed by pathologists, such as Rudolf Virchow in the late 1800s who reported morphological differences between single tumor cells under the microscope [1]. Further development of karyotyping and cytogenetic technologies in the 1970s led to numerous studies reporting heterogeneity in amplifications of oncogenes and deletions of tumor suppressors within the same tumor [2–4]. The concept of ITH soon emerged, but was largely ignored in clinical practice, because it confounded the diagnosis and therapeutic treatment of cancer patients. In the late 1990s microarray technologies were developed [5], which were soon followed by the development of next-generation sequencing (NGS) technologies around 2005 [6,7]. These new genomic technologies led to a paradigm shift in the field, away from qualitative studies based on single markers, and towards large-scale quantitative ITH datasets. The subsequent application of NGS technologies to human tumors revealed that ITH is common in many human cancers [8–10]. However, despite the significant progress, a central question has remained: how did ITH emerge during tumor progression?

Tumor evolution is a field that applies knowledge of species evolution, ecology and population genetics to understand how tumor cell populations respond to selective pressures [11]. Formalizing the concept of tumor evolution is often accredited to Peter Nowell [12] and pioneers such as Isaiah Fidler who recognized the importance of clonal diversity in metastasis [13]. Over the following decades studies have showed that tumor cells encounter selective pressures in their microenvironment, including the immune system, pH changes, chemotherapy, radiation, nutrient deprivation and geographic barriers [14]. These selective pressures shape the evolutionary trajectory of the tumor and clonal lineages. Principles such as species richness, selection, fitness and population bottlenecks are useful concepts for understanding tumor evolution, however it is also important to note that many concepts from ecology and population genetics do not apply to tumors, most notably sexual selection and meiotic recombination [14,15].

Tumor evolution is difficult to study in human patients. The central problem is that patients cannot ethically be biopsied at multiple time points during the progression of the disease. As a consequence, most studies have inferred the evolutionary history from single time-point samples. This approach is conceptually feasible, because ITH provides a permanent record of the mutations that occurred during the natural history of the tumor [8,16]. Researchers can apply phylogenetic inference to reconstruct tumor cell lineages and order the chronology of mutations that occurred over time. However this approach provides an

☆ This article is part of a Special Issue entitled: Evolutionary principles - heterogeneity in cancer?, edited by Dr. Robert A. Gatenby.

* Corresponding author at: Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA.
E-mail address: nnavin@mdanderson.org (N. Navin).

http://dx.doi.org/10.1016/j.bbcan.2017.01.003
0304-419X/

incomplete picture of how tumor cells evolve, particularly when intermediate clones are not persistent during progression. Consequently there has been much debate regarding the general models of tumor evolution. Several competing models have been proposed: Linear Evolution (LE), Branching Evolution (BE), Neutral Evolution (NE) and Punctuated Evolution (PE) (Fig. 1). The evidence supporting these models will serve as the basis of discussion for this review, but first we will review the genomic methods that are used to study ITH and clonal evolution.

2. Methods for resolving intratumor heterogeneity

NGS methods can measure thousands of mutations and generate large-scale genomic datasets on tumors [6,7]. However standard NGS methods require bulk tissue and therefore provide limited information on the subclonal architecture of a tumor. To address this limitation, further methods were developed to delineate ITH: deep-sequencing, multi-region sequencing and single cell DNA sequencing (Fig. 2). Deep sequencing involves performing NGS at high coverage depth to measure mutant allele frequencies (MAFs) [17,18] (Fig. 2A). Using computational methods such as SciClone [19] or Pyclone [20], the mutation frequencies are then normalized and clustered to identify clonal subpopulations that are assumed to share similar MAFs. This approach is experimentally simple, but cannot accurately resolve clonal subpopulations when they share similar MAFs in the tumor. Another method is multi-region sequencing and involves sampling different geographical regions of the tumor for exome sequencing (Fig. 2B) [21–24]. This approach is experimentally straightforward, but has limited ability to resolve subclones that are intermixed within the same spatial regions. Another approach is single cell DNA sequencing (Fig. 2C) [25–29]. This approach involves isolating single tumor cells, performing whole genome amplification (WGA) and then sequencing and comparing multiple cells to resolve ITH and reconstruct clonal lineages [30]. The advantage of this approach is that it can fully resolve admixtures of clones, however due to cost and throughput, only a limited number of cells can be profiled, potentially leading to sampling bias [31].

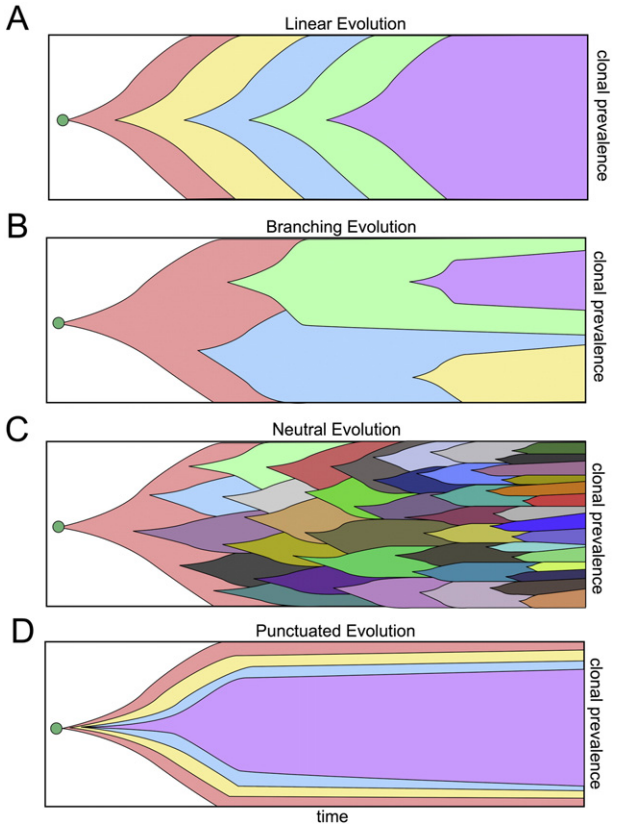

Fig. 1. Models of tumor evolution. Illustration of tumor evolution models showing dynamic changes in clonal frequencies over time. This figure is based on the original publication by Marusyk and Polyak [8]. (A) Linear Evolution (B) Branching Evolution (C) Neutral Evolution (D) Punctuated Evolution. Colors indicate clones with different genotypes.

3. Reconstructing tumor evolution from intratumor heterogeneity

After resolving ITH, the data can be used to reconstruct clonal lineages using phylogenetic inference to understand tumor evolution. In phylogenetic tumor trees, the internal nodes represent common ancestors, whose genotype can be deduced from the commonalities between their descendants. A phylogenetic tree thus provides a window into the past, by estimating the order in which mutations occurred as clones diverged in lineages and formed subpopulations. Phylogenetic trees can be constructed from ITH using different algorithms. The units of heterogeneity that appear at the tips of the tree are called *taxons*, and represent either clones, single cells, or spatial regions, depending on the experimental method that was used. The tree is often constructed using an algorithm to satisfy a *parsimony criterion*, in which the tree with the minimum number of changes leading to the observed data is inferred. For deep sequencing data, clones are inferred by clustering MAFs and are arranged into a tree using the *sum condition* that MAFs of child nodes must sum to those of their parents, and the *ancestry condition* that descendants have all the mutations in their parents. Many computational algorithms have been developed for this purpose [32–40]. There are also specialized algorithms for inferring phylogenetic trees from multiregion sequencing data [41,42]. Using single-cell data, it is possible to order mutations and attach cells to the mutation trees [43], or to additionally cluster cells into clones and construct a clone tree similar to those produced by deep sequencing analysis methods [44,45]. In summary, these methods enable tumor evolution to be reconstructed from ITH using single time-point samples. However these trees are based on the *infinite sites assumption* [46] which implies that mutations accumulate and are never lost. This assumptions is often violated in tumors, where chromosome deletions and LOH are common.

4. Evolutionary concepts and definitions

To understand models of tumor evolution, several concepts and definitions are necessary. A clone is defined as a group of tumor cells that shares a highly similar genotype and mutational profile, while a subclone is a group of tumor cells that diverged in the evolutionary lineage and has acquired additional mutations [9]. Truncal mutations are ancestral mutations in the *trunk* of the phylogenetic tree that are shared by all clones, while *subclonal* mutations define a lineage that has diverged from the trunk [47]. *Private mutations* refer to mutations that are only detected in a single taxon. Another important concept is *fitness*, which refers to the ability of a tumor cell to survive and proliferate, so that it can propagate its genotype to the gene pool in the tumor. Tumor clones with increased fitness will become more prevalent in the tumor mass over time. *Driver mutations* confer a fitness advantage, while *passenger mutations* have no effect on fitness. Increased fitness can lead to clonal expansions in which one genotype expands in frequency in the tumor mass. A *selective sweep* describes the process in which a genotype emerges with an extremely high fitness that it outcompetes all other clones in the tumor [14].

5. Linear tumor evolution

One of the most renowned models for tumor evolution posited that mutations were acquired linearly in a step-wise process leading to more malignant stages of cancer [48]. In this linear evolution (LE) model new driver mutations provide such a strong selective advantage, that they outcompete all previous clones via selective sweeps that occur during