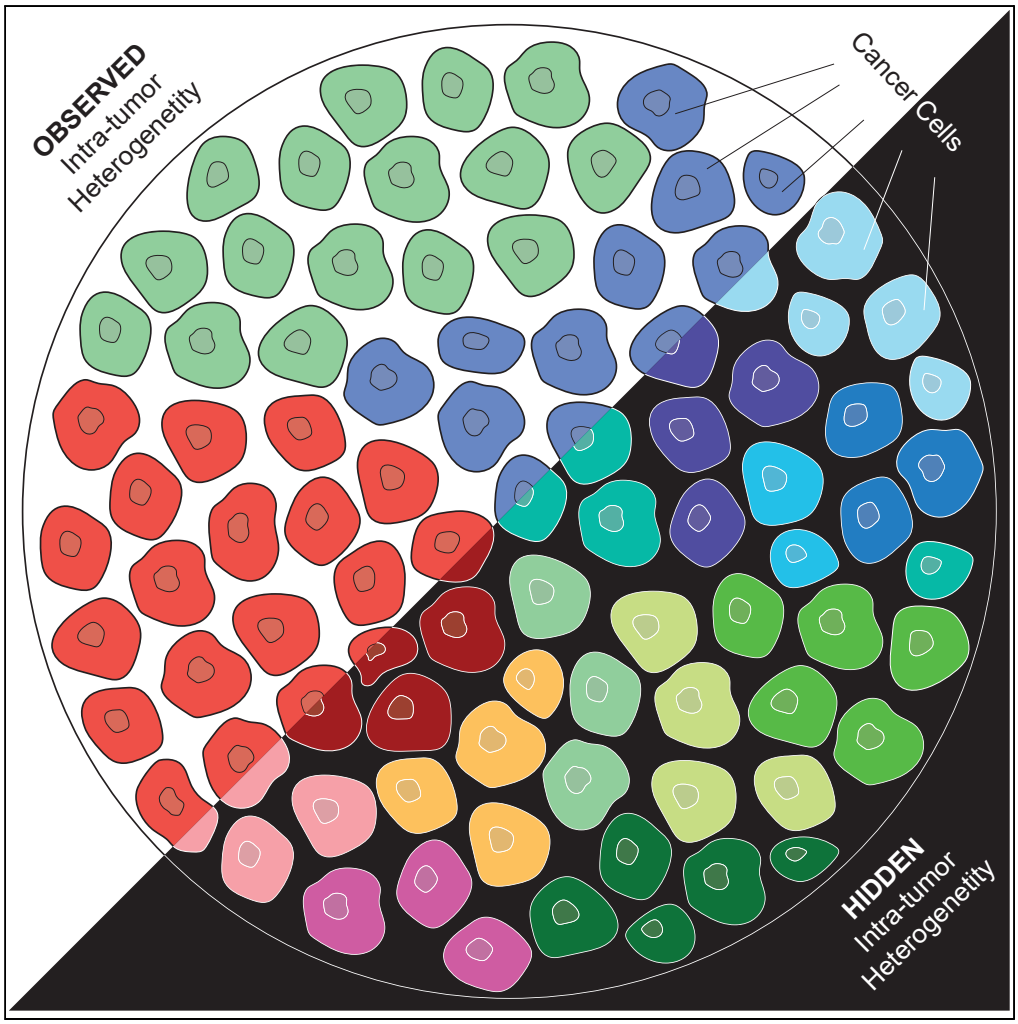
iScience

CellPress

Article

Dynamic Emergence of Observed and Hidden Intra-tumor Heterogeneity



Franck Raynaud,
Marco Mina,
Giovanni Ciriello

franck.raynaud@unige.ch (F.R.)
giovanni.ciriello@unil.ch (G.C.)

HIGHLIGHTS

Intra-tumor heterogeneity inferred from human samples is underestimated

Hidden and observed intra-tumor heterogeneity do not always correlate

Observable clones carry information to estimate the extent of hidden heterogeneity

Raynaud et al., iScience *21*, 157–167
November 22, 2019
https://doi.org/10.1016/j.isci.2019.10.018

iScience

CellPress

Article

Dynamic Emergence of Observed and Hidden Intra-tumor Heterogeneity

Franck Raynaud,[1,2,3,*] Marco Mina,[1,2] and Giovanni Ciriello[1,2,4,*]

SUMMARY

Intra-tumor heterogeneity is frequently observed in cancer patients, and it is associated with therapeutic resistance and disease relapse. However, its systematic assessment is still limited and often unfeasible. Here, we use a mathematical model of tumor progression to decipher how multiple clones emerge and organize into complex architectures. We found a trade-off between cancer cell alteration and proliferation rates that defines a transition between low and high heterogeneity, the latter characterized by branching tumor phylogenies. We predict the existence of observed and hidden intra-tumor heterogeneity, which challenges the correct estimation of intrinsic tumor complexity. Although the numbers of observed and hidden clones do not always correlate, we demonstrate that population frequencies of observed clones can be used to estimate the extent of hidden heterogeneity in both simulated and human tumors. The characterization of complex clonal architectures is a critical first step toward understanding their organizing principles and predicting their emergence.

INTRODUCTION

Individual tumors are often a composite of heterogeneous cancer cells, which have accumulated distinct molecular modifications over time. This dynamic process has been characterized according to evolutionary principles (Nowell, 1976) where abnormal cell proliferation is associated with the emergence and selection of alterations. Molecular alterations are in the most cases functionally neutral, but occasionally they confer a selective advantage to the cancer cell where they occurred, by promoting key oncogenic hallmarks (Hanahan and Weinberg, 2011). Typically referred to as "drivers", selected alterations promote clonal expansion and, thus, the evolution and diversification of the tumor.

Starting from this premise, evolutionary theory has provided an accurate framework to understand the emergence of cancer and its progression. Seminal works established this connection for neoplasms and metastasis (Fidler, 1978; Nowell, 1976). Ever since, multiple models of cancer evolution have been proposed, mostly falling in two major categories: models of tumor initiation and models of tumor progression. Tumor initiation requires the fixation of a first selective event, a gain-of-function mutation in an oncogene (Michor et al., 2004), a loss-of-function alteration in a tumor suppressor gene (Knudson, 1971; Komarova, 2007), or early incremental increases of cell fitness (Michor et al., 2011). Tumor progression, instead, is characterized by a succession of selected mutations and corresponding waves of clonal expansion. In this context, models of clonal expansion proved to be useful to characterize disease progression, resistance to therapy (Durrett and Moseley, 2010; Michor et al., 2005), mutation burden (Bozic et al., 2010; McFarland et al., 2013, 2014), and mutation timing (Beerenwinkel et al., 2007, 2014).

Despite the numerous insights provided by these models, the link between evolutionary parameters modeling tumor growth, such as alteration and/or replication rates, and the resulting extent of intra-tumor heterogeneity remains largely unexplored (Bozic et al., 2016; Chowell et al., 2018), especially in relation to actual human tumors. Indeed, a robust experimental estimation of the parameters themselves is currently missing, and the clonal composition of a tumor is mainly assessable through algorithmic inferences from bulk tumor sequencing rather than direct observations (Deshwar et al., 2015; Jiang et al., 2016; Miller et al., 2014; Raynaud et al., 2018; Roth et al., 2014). Recently, distinct types of clonal architectures have been inferred for multiple tumor types analyzed by multi-regional sequencing and/or longitudinal sample cohorts (McGranahan and Swanton, 2017; Yates et al., 2017). Nonetheless, whether and how these architectures depend on properties of tumor growth remains to be investigated.

Here we numerically explored the emergence of diverse clonal architectures as a function of alteration and proliferation rates. Based on extensive simulations using a wide range of parameters, we explored

[1]Department of Computational Biology, University of Lausanne, Lausanne, 1015 Lausanne, Switzerland

[2]Swiss Institute of Bioinformatics, Lausanne, 1015 Lausanne, Switzerland

[3]Department of Computer Science, University of Geneva, 1205, Geneva, Switzerland

[4]Lead Contact

*Correspondence: franck.raynaud@unige.ch (F.R.), giovanni.ciriello@unil.ch (G.C.)

https://doi.org/10.1016/j.isci.2019.10.018

fundamental features of intra-tumor heterogeneity, such as how many clones can typically be found in a tumor sample, whether they have similar sizes or one accounts for most of the cell population, how did they descend from each other, and, finally, what are the evolutionary determinants of these features. Our results from simulated tumors directly link clonal architectures to characteristics of tumor progression and indicate that intra-tumor heterogeneity is frequently underestimated. Importantly, we could determine features that can be computed for both simulated and human tumors that are predictive of this underestimation.

Model of Cancer Evolution

To investigate the properties of cancer evolution, tumor clonal architecture, and intra-tumor heterogeneity, we adapted a previously proposed model of tumor progression (Bozic et al., 2010). In this model, tumors evolve according to a discrete time process in which cells, simultaneously, at each time step either replicate or die. When a cell replicates, one of the daughter cells can gain a new alteration with a probability μ (alteration rate). We consider two types of alterations: *driver* alterations, which increase the selective advantage of the cell by reducing its probability to die, and passenger alterations, which have instead a neutral effect. To which extent the probability of dying is reduced by the acquisition of a new driver alteration is modulated by the parameter s (fitness). The alteration rate μ and fitness s are input parameters; they are identical for all cells and remain constant during the simulation. For a given fitness parameter s, a cell that has accumulated k driver alterations has a probability to die d_k given by:

$$d_k = \frac{1}{2}(1-s)^k \qquad \text{(Equation 1)}$$

The probability to replicate b_k satisfies the following relationship:

$$b_k + d_k = 1 \qquad \text{(Equation 2)}$$

Human tumors typically exhibit highly heterogeneous proliferation rates and alteration loads. Here, to mimic such heterogeneity, we did not estimate specific model parameters, instead we chose to explore a wide range of possible parameter values for μ and s. First, we set the fitness parameter $s \in [10^{-4} - 5 \times 10^{-1}]$ in agreement with previously proposed values (Bozic et al., 2010; Chowell et al., 2018; Levy et al., 2015; McFarland et al., 2014; Vermeulen et al., 2013). Then, we estimated a range of alteration rates based on previously reported mutation burden across multiple human cancers (Lawrence et al., 2013). By analyzing tumor cohorts from The Cancer Genome Atlas (TCGA), the number of mutations per nucleotide was estimated between 7×10^{-8} to 10^{-4} (assuming that the exome accounts for ~2% of the genome). Therefore, we set for our simulations $\mu \in [10^{-7} - 10^{-3}]$, also consistent with recent numerical studies (Bozic et al., 2010; Chowell et al., 2018; McFarland et al., 2014). New driver mutations occur at a rate $\mu \times \mu_d$, with $\mu_d = 0.025$. We chose this value based on an estimated number of 500 cancer-associated alterations (Bailey et al., 2018; Mina et al., 2017). Notably, a similar number of cancer-associated mutations was reported in the COSMIC Cancer Census, http://cancer.sanger.ac.uk/census.

In our simulations, we defined a clone as a population of cells with identical mutational composition (i.e., genotype): after each replication step, if no alteration has occurred then the two daughter cells will remain in the same clone, otherwise the sibling with the new alteration will create a new clone, whether the new alteration is a driver or a passenger (Figure 1A). Tracking exact lineages between clones of each simulated tumors is computationally intensive, as all cells have to be monitored at each replication step. To improve efficiency, we simulated the evolution of clones, rather than cells. Specifically, for a clone with N identical cells, the number of replicating cells n_r was drawn from a binomial distribution with a success probability b_k. Then among n_r, we determined the number of mutating cells n_m from a binomial distribution with a probability μ. Finally, for each newly altered cell, the probability for the new mutation to be a driver mutation was μ_d.

To estimate intra-tumor heterogeneity (the mean number of clones) and quantify the tumor clonal architecture, we retained only clones with a number of cells greater or equal to 1% of the total population (here referred to as *observable clones*). This is consistent with the fraction of sequencing reads typically required by cancer exome-sequencing studies, e.g., from TCGA, to retain a somatic mutation (Figure S1A).

The model of clonal evolution is implemented in Python, using the environment for tree exploration (ETE) (Huerta-Cepas et al., 2010).

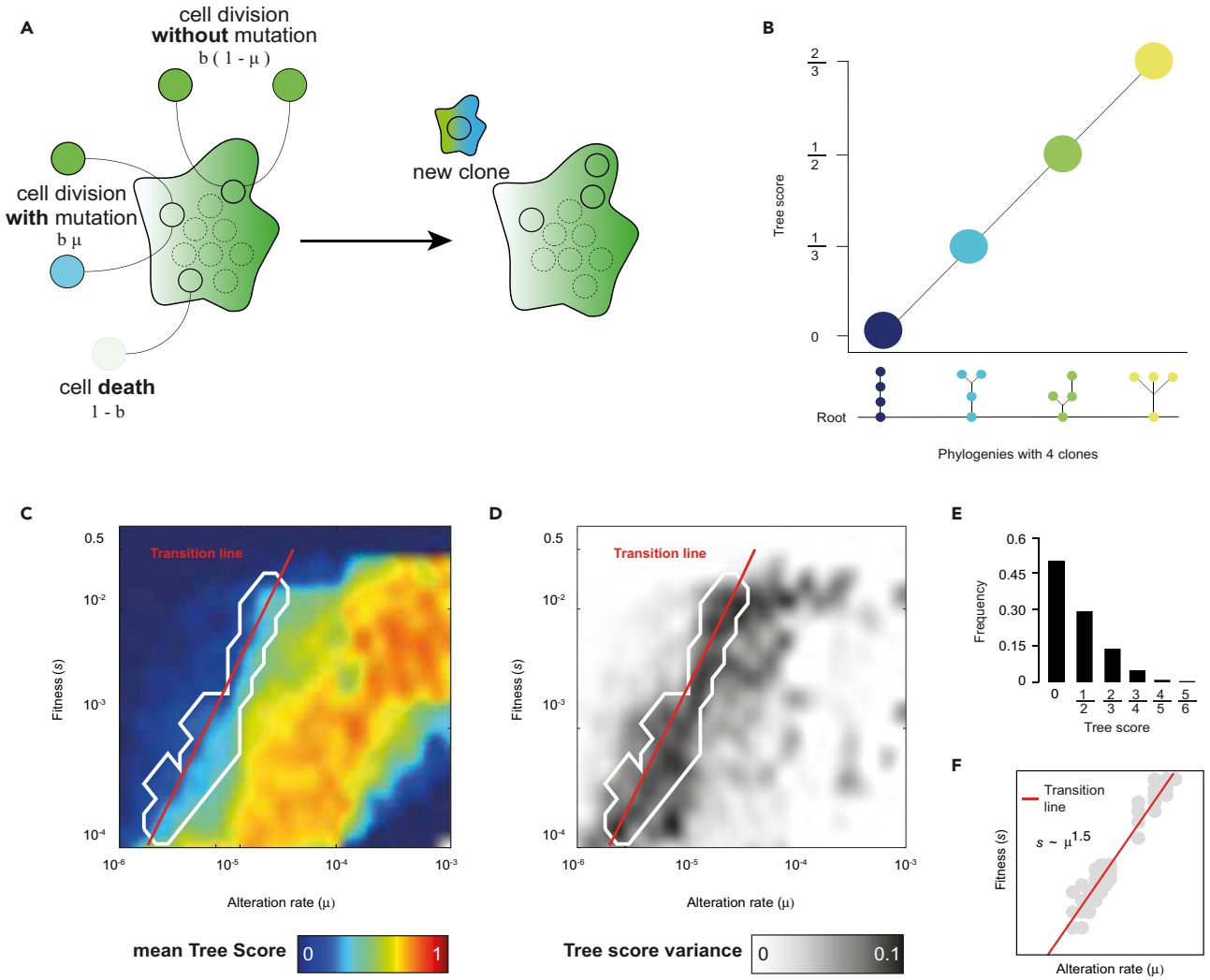


Figure 1. Tumor Clonal Architectures and Transition between Linear and Branched Phylogenies

(A) Representation of the mutational process. Top, cell division without mutational event. The size of the clone is just increased by one unit. Middle, an alteration occurs during replication. A new clone is created (blue) by one cell, and the size of the original clone does not change. Bottom, cell death. In practice, more than one cell replicates and/or acquires a new mutation in the same time, resulting in the formation of more than one new clone at each replication. We pictured one event for the sake of simplicity.

(B) Example of Tree score values for phylogenies with four clones. Tree scores increase with increasing divergence.

(C) Tree scores of simulated tumors as a function of the alteration rate μ (x axis) and fitness *s* (y axis). Tree scores are color coded (cold colors for low scores, warm colors for high scores). The region (white contour) and fitted line (red line) corresponding to the transition between linear and divergent phylogenies are highlighted.

(D) Tree score variance values of simulated tumors as a function of the alteration rate μ (x axis) and fitness *s* (y axis). Tree score variances are color coded (white to black corresponding to low to high variance). The region (white contour) and fitted line (red line) corresponding to the transition between linear and divergent phylogenies are highlighted.

(E) Tree score distribution of tumors within the region of transition.

(F) The transition line (red line) was derived by fitting the points with maximal variance (gray dots).

See also Figure S1.

RESULTS

In total, we simulated ~40,000 tumors, with a similar number of replicates for each pair of parameters (μ, s). For each simulation, we tracked the number of clones as well as their size and lineage, thus to be able to reconstruct the exact architecture of each simulated tumor. We stopped the simulation when a tumor reached a size of 5×10^8 cells.

Properties of Tumor Clonal Architectures

To capture the evolutionary history of a tumor, we need to understand how clone lineages are organized, i.e., how individual clones descend from one another (clone phylogeny). Clone phylogenies can be conveniently represented as trees wherein nodes are the clones and two clones are connected if one is the offspring of the other. The founder clone is the root of the tree, and clones without descendant are the leaves. At the extremes of the possible tree configurations are *linear* phylogenies, i.e., sequential generation of clones along a single lineage, and *star-like* phylogenies, i.e., each new clone descends directly from the root. Intuitively, the more branching (i.e., star-like) a phylogeny, the shorter is the path from the leaf to the root; in contrast, linear phylogenies will have a single leaf at the maximal distance from the root. To quantify the extent of branching in a phylogeny we introduced the following quantity:

$$Tree\ Score = 1 - \frac{\langle d_l \rangle_l}{N-1} \qquad \text{(Equation 3)}$$

where N is the number of observable clones and $<d_l>_l$ is the average length of the path from the root to the leaves. Linear phylogenies obtain a Tree score equal to 0, whereas the Tree score will increase with greater branching (Figure 1B). We should note that Tree scores are not uniquely associated with phylogenies, i.e., different phylogenies can receive the same Tree score. However, in the simulations, this degeneracy remained moderate: for each pair of parameter values, μ and s, we observed on average less than two different phylogenies with the same Tree score (Figure S1B). Moreover, phylogenies receiving the same Tree score were significantly more similar than phylogenies with different Tree scores (Figures S1C–S1E and Transparent Methods), indicating that the Tree score provides a good representation of the tree structure.

Next, we assessed how Tree scores varied in our simulated tumors based on input parameters. The resulting distribution (Figure 1C) showed an opposite effect of the fitness and alteration rate on the Tree score. At fixed fitness the mean Tree score increased with increasing alteration rates (i.e., multiple independent linages emerge when increasing the alteration rate), whereas at a fixed alteration rate, it decreased with increasing fitness. The parameter space was largely split into two regions: one with low branching or linear phylogenies (low or zero Tree score) and one with highly branching phylogenies (high Tree score). Interestingly, where the mean Tree score shifted to non-zero values, we observed an increase of its variance (Figure 1D). This result is reminiscent of a first-order phase transition in physical systems, where the variance of the order parameter diverges at the transition and different phases coexist, as shown by the distribution of Tree scores at the transition (Figure 1E). In our case, the transition was characterized by two phases: linear and branching evolution, separated by the transition line $s_t \sim \mu_t^{\alpha}$ (with $\alpha = 1.5 \pm 0.05$ Figure 1F). To further assess the robustness of this result, we estimated the transition line with an alternative approach. Briefly, for every fitness values s we identified the mutation rate μ^* that generated the highest number of observable clones. We then fit the measured mutation rate μ^* as function of s. The best fit showed that $\mu^*(s)$ followed the same power law as the transition estimated using the variance of the Tree score, and the rescaling of the mean number of clones confirmed that the transition held for diverse model parameters (Figures S2A–S2F).

Overall, Tree scores of simulated phylogenies were highly correlated with their number of clones, indicating that highly heterogeneous tumors typically exhibited a branching phylogeny (Figure S2G). Intriguingly, we recently reported a similar association between Tree scores and number of clones in human tumors (Raynaud et al., 2018) (Figure S2H). Indeed, the set of phylogenies generated by our simulations and inferred from human tumors were highly consistent and, importantly, they represented only a subset of all possible phylogenies for a given number of clones (Figure S2I). These results indicate that the growth dynamics of our model and human tumors prevent highly polyclonal phylogenies with limited branching from emerging.

In the parameter space defined by μ and s, the Tree score surprisingly decreased from high to low values in a region characterized by low fitness and high alteration rates (Figure 1C, bottom right corner). However, here we did not observe high variance of scores, suggesting a different mechanism for the reemergence of linear phylogenies in this region of parameters. A similar trend was observed for the overall distribution of the number of clones (Figure 2A). Indeed, here two transitions were observed from low to high number of clones (blue to red in Figure 2A) and from high to low in the region with high alteration rates and relatively low fitness (red to blue in Figure 2A, bottom right corner). Given we initially restricted the analysis to clones with a number of cells corresponding to at least 1% of the total cell population (observable clones),

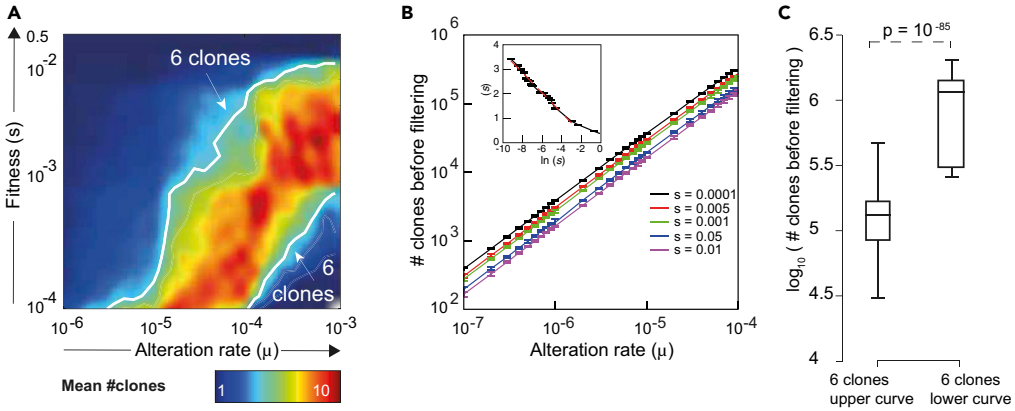


Figure 2. Observed and Total Number of Clones in Simulated Tumors

(A) Mean number of clones with a size (number of cells) greater than 1% of the total cell population obtained by simulated tumors as a function of their alteration rate μ (x axis) and fitness *s* (y axis). For each pair of coordinates, μ and *s*, the mean number of clones observed in simulations corresponding to those coordinates is color coded (cold colors for low number, warm colors for high numbers). Simulations with six clones (white lines) can be found for different ranges of parameters.

(B) The total number of clones, irrespective of their size, increases linearly with the mutation rate. The slope of the curves depends on the fitness *s* (curves for five representative *s* values are shown) and decreases with the logarithm of *s* (inset). Data are presented as mean ± SD.

(C) Boxplot of the number of clones before filtering for simulated tumors with six detectable clones found in the area of low hidden heterogeneity, upper white curve in (A), and of high hidden heterogeneity, lower white curve in (A); p value is calculated with a two-tailed t test.

See also Figure S2.

we wondered whether the number of clones below such threshold was significantly different in the two areas characterized by low number of clones and low Tree scores. In our simulations, we could count all clones independently of their sizes and we found that the total number of clones increased linearly with the alteration rate and never decreased (Figures 2B and S2J). Indeed, for samples with the same number of observable clones (e.g., n = 6), the number of "hidden" clones was significantly higher in the second transition (high to low) than in the first one (low to high) (Figure 2C) and confirmed that the decrease in the number of clones and Tree score was simply associated with the filtering process. The immediate consequence of these results is that highly heterogeneous tumors could exhibit the same number of observable clones as less heterogeneous ones, "hiding" their true complexity.

The total number of clones, irrespective of their size, could reach values orders of magnitude higher than those typically observed in human datasets (Andor et al., 2016). The difference in the number of clones due to the 1% filtering indicated that simulated tumors were composed by few large clones (up to ~10 clones on average) and a wide array of small clones with size below the detectability threshold (Figures 2A and 2B). A similar observation was recently made in Chowell et al. (2018) for a different detection threshold (10%), suggesting that independent of the threshold values, the majority of the true heterogeneity remains hidden. Given that clones are here defined by their genotype, independent of whether they actually exhibit greater fitness, we expect many of these small hidden clones to disappear. Hence, we asked what percentage of the hidden cell population is predicted to expand and form a new clonal lineage. Where hidden heterogeneity is the greatest (Figure 2B, bottom right corner), the number of driver mutations per hidden clone increased (Figure S2K). However, given the low values of the fitness parameter in this space, less than 0.1% of the hidden cell population was predicted expanding (Figure S2L). Vice versa, for high values of both μ and *s*, this value reached 5% (Figure S2L). Although multiple factors could limit the growth of newly generated clones in human tumors, this *in silico* observation suggests that filtering mutations with low frequencies may result in underestimating intra-tumor heterogeneity. Importantly, in human tumors where alteration rates and fitness are not given it remains impossible to estimate the extent of hidden heterogeneity.

Given each simulated tumor is composed of the same number of cells, the difference between the number of observable and hidden clones suggests that the relative size of detectable clones could provide an

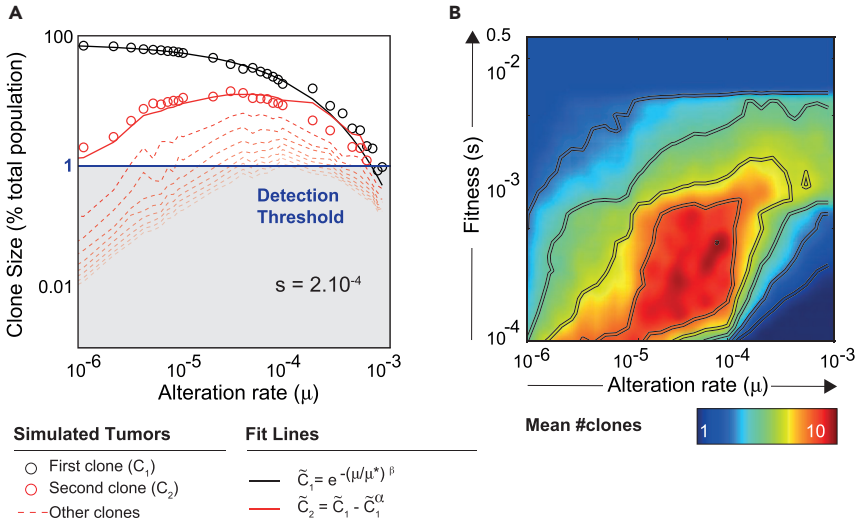


Figure 3. Clone Size Distribution and Analytical Prediction of Intra-tumor Heterogeneity

(A) The size, as a percentage of the total population, of the first clone (black dots) across simulations at different alteration rates and fixed fitness ($s = 0.0002$) decreases with increasing alteration rates μ (x axis) and is fitted by a stretched exponential (black line). The size of the second clone (red dots) initially grows with increasing alteration rates but eventually decreases. This trend can be fitted as a function of the size of the first clone and the model parameters (red line, see Methods). Subsequent clones follow the same trend (red dotted lines). Clones with sizes greater than 1% (blue line) are detectable.

(B) Analytically derived number of clones as a function of mutation rate and fitness.

See also Figure S3.

indication of the actual extent of heterogeneity. We found that the normalized size of the biggest clone C_1 decreased with the alteration rate as a stretched exponential (Figures 3A, S3A, and S3B and Transparent Methods)

$$C_1(s,\mu) = e^{-\left(\frac{\mu}{\mu^*}\right)^{\beta}} \qquad \text{(Equation 4)}$$

where μ^* and β are both functions of the fitness s (Transparent Methods and Figures S3C and S3D). Vice versa, the size of the other clones C_i ($i = 2,3,..,N$) increased with the alteration rate while $\mu << s$ (corresponding to an increase of the number of detectable clones), but as μ approaches s, the curves reached a maximum and then collapsed below the detection threshold (resulting in a decrease of the number of observable clones) (Figures 3A and S3E). We can write the sizes of the subsequent clones C_i using the following ansatz:

$$C_i = C_1 - C_1^{\alpha_i} \qquad \text{(Equation 5)}$$

where α_i is a function of the rank i and the fitness s (Transparent Methods and Figures S3F–S3H). Using our numerical results, we estimated the value of μ^*, β, and α_i and analytically reconstructed the distribution of the number of observable clones. The resulting distribution closely mimicked what we observed in our simulations (Figure 3B), and, importantly, it can be estimated independent of any detection threshold because μ^*, β, and α_i do not depend on it.

Based on Equations 4 and 5, we can predict that for low alteration rates the first clone is significantly bigger than the second one (Figure 3A, left side), whereas at high alteration rates their sizes become comparable (Figure 3A, right side). As clone sizes can be inferred in human tumors, this observation could be used to discriminate between low and high hidden heterogeneity and help estimating its extent in human samples without any information on the underlying alteration rate and fitness.

A convenient way to analyze the difference in size among subsequent clones is to rely on the concept of population frequency (PF), which is independent of the final size of the tumor. For a clone i, its PF, PF_i, corresponds to the fraction of cells in the tumor that exhibits the set of alterations in i, but not necessarily only

those, i.e., the fraction of cells in i and in all the clones descending from i. Formally, the PF of clone i in a tumor T is defined as:

$$PF_i = \frac{1}{\sum_{j \in T} |j|} \left(|i| + \sum_{j \in \{s_i\}} |j| \right) \quad \text{(Equation 6)}$$

where $|i|$ is the size of the clone i and $\{s_i\}$ is the set of clones descending from i (Figure S4A). The distribution of sorted PF values can recapitulate the differences among clone sizes that we observed at varying alteration rates (Figure 4A). Intuitively, a sharply decreasing distribution indicates that the first clone had time to grow before the emergence of new clones, hence the first clone is considerably bigger than the others (Figure 4A, green line). By contrast, a slowly decreasing distribution indicates that new clones rapidly emerged giving rise to observable clones of similar size (Figure 4A, red line).

We computed the PF distribution for all simulated tumors with 5–10 observable clones and scored each simulation by the area under the curve of the PF distribution (PF-AUC) (Figure 4A, gray area). PF-AUC scores increased with alteration rates and decreased with fitness (Figure 4B) and were highly correlated with the extent of hidden heterogeneity (Spearman's correlation = −0.95, p-value = 10^{-165}, Figures 4B and 4C). This observation confirmed that PF-AUC could be used to discriminate among tumors with the same number of observable clones but different extent of hidden heterogeneity. Interestingly, PF distributions in our simulated dataset, separated into three classes corresponding to different parameters and clonal heterogeneity (Figure 4D). Tumors characterized by low hidden heterogeneity (Figure 4D, region A) had sharply decreasing PF distributions, consistent with the presence of the dominant first clone, and those with high hidden heterogeneity (Figure 4D, region C) were characterized by PF distributions decreasing slower, and thus by multiple clones of more similar size. A third group corresponded to simulated tumors with a high mean number of clones (see Figure 2B) and intermediate level of hidden heterogeneity (Figure 4D, region B). To assess whether these regions of hidden heterogeneity are independent of the adopted growth model, first we implemented a second model of tumor evolution, similar to a previously proposed one wherein passenger mutations induce a small, yet not null, deleterious effect on cell fitness (see Transparent Methods) (McFarland et al., 2014). The results showed the same three distributions independently of the chosen model (Figure S4C). Next, we ran simulations with variable rates of driver mutations (μ_d), and again three regions of hidden heterogeneity emerged, consistent with the previous results (Figure S4D). Overall, these regions were robust to modifications of the adopted growth model.

Finally, to estimate the extent of hidden heterogeneity in human tumors, we collected the cancer genomes of 6,000 samples from 32 tumor types profiled by TCGA and inferred clone population frequencies from the phylogeny of each human sample estimated from its set of somatic mutations and copy number alterations (Raynaud et al., 2018). We assigned each human sample, with 5–10 predicted clones, to one of the three groups (A, B, and C) by matching its PF distribution to the closest mean distribution obtained in each group by simulated tumors (Figure 4E). Importantly, distributions of PF-AUC values in human tumors were neither uniformly nor normally distributed but rather consistent with a trimodal distribution whose modes were significantly different against sub- and bootstrap sampling (Figure 4F, Transparent Methods, Tables S1 and S2). The identification of three classes with a different amount of hidden heterogeneity was thus independently confirmed by simulations using different growth models and inferred phylogenies from human samples.

In the TCGA dataset, we found that different percentages of samples within each tumor type were assigned to the three classes, with almost all tumor types including samples in the high hidden heterogeneity class and 17 of 32 cancer types having more than 10% of their samples in this category (Figure 4G). Despite the fact that this analysis could be done on a limited number of samples (only samples with 5–10 clones were analyzed), preliminary observations indicated that hidden heterogeneity classes were often associated with different overall survival (Figure S4E). Overall, these results suggest that clone population frequencies and PF-AUC scores can provide a simple but effective way to discriminate among human tumors with similar observable intra-tumor heterogeneity, those with high and low hidden heterogeneity.

DISCUSSION

Intra-tumor heterogeneity severely hinders durable and complete response to cancer therapy (Jamal-Hanjani et al., 2017; Siravegna et al., 2015). Its systematic assessment is often unfeasible in the clinic, and

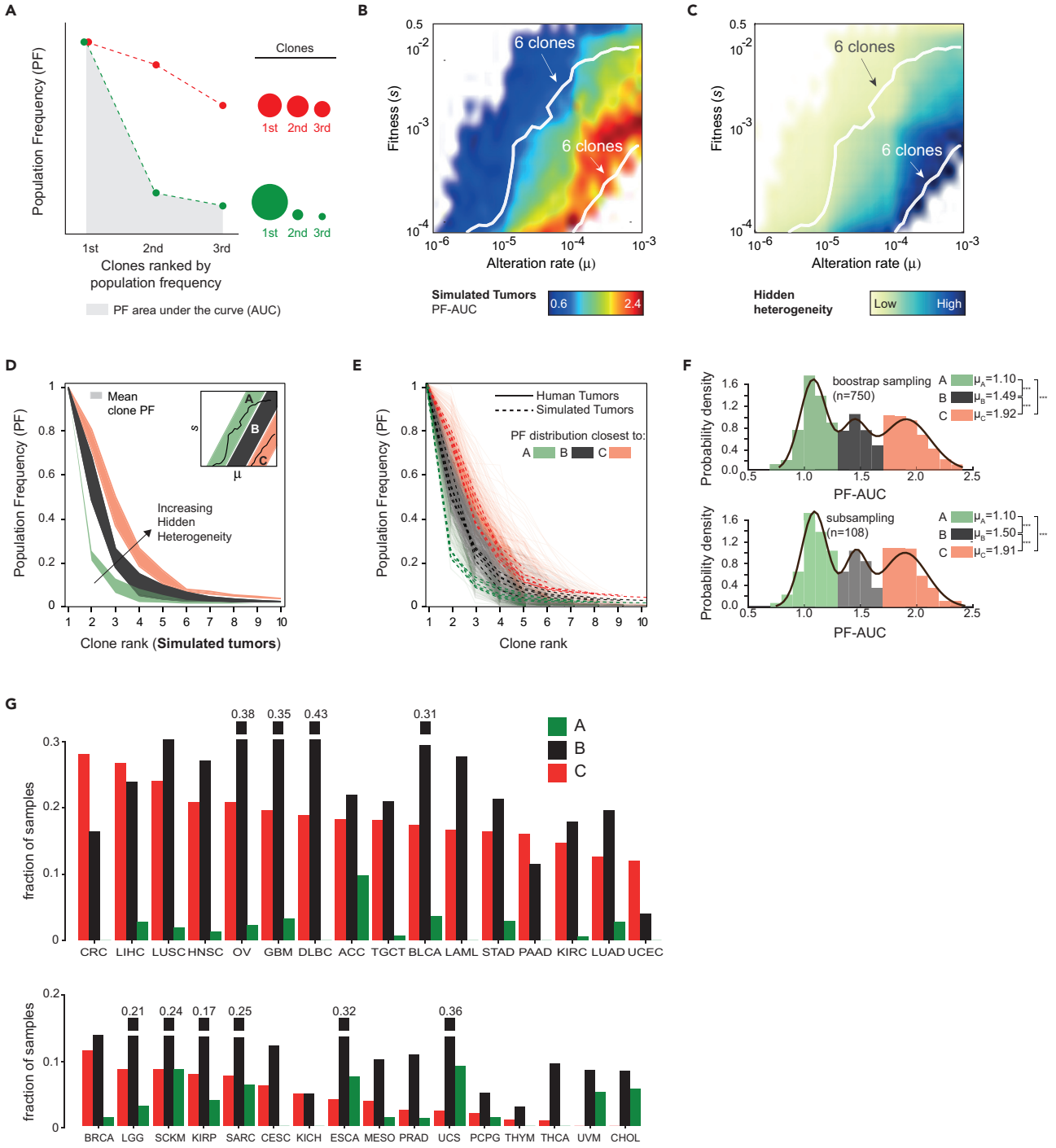


Figure 4. Estimation of Hidden Intra-tumor Heterogeneity Using Clone Population Frequencies and Comparison with Human Tumor Samples

(A) Schematic of the curves of ranked clone population frequencies. A sharply decreasing curve (green dotted line) corresponds to a large initial clone followed by small clones (green dots with size indicative of the clone size). A slowly decreasing curve (red dotted line) corresponds to clones of a similar size (red dots with size indicative of the clone size). Each curve can be scored by its area under the curve (AUC, e.g., the gray area below the green line).

(B) Clone population frequency AUC (PF-AUC) values of simulated tumors with the number of clones between 5 and 10 as a function of the alteration rate μ (x axis) and fitness s (y axis). PF-AUC values are color coded (cold colors for low values, warm colors for high values). Simulation with six clones (white lines) has different PF-AUC values based on the range of parameters.

Figure 4. *Continued*

(C) Hidden heterogeneity values (1 − fraction of cells in detectable clones) of simulated tumors with the number of clones between 5 and 10 as a function of the alteration rate μ (x axis) and fitness *s* (y axis). Values are color coded (light colors for low values, dark colors for high values). Simulation with six clones (white lines) has different extent of hidden heterogeneity.

(D) Ranked clone population frequency curves for simulated tumors with the number of clones between 5 and 10. PF curves were separately derived for simulated tumors in three ranges of parameters (μ and *s*) as shown in the inset on the top right corner (group A in green, group B in black, group C in red). For each group, the corresponding curves were aggregated and the range of values are displayed.

(E) Ranked clone population frequency curves for human (continuous lines in the background) and simulated (dotted lines on top) tumors with the number of clones between 5 and 10. PF curves of human tumors were assigned to groups A, B, or C (lines are colored with the color of the corresponding group) based on the closest curve of simulated tumors with the same number of clones.

(F) Distribution of PF-AUC values for bootstrap (upper panel, n = 501 samples) and subsampling (lower panel, n = 108 samples). Histograms represent the empirical distributions; black lines are the best fit from a mixture of three Gaussian distributions. Histograms are color coded for clarity and to help to visualize regions A–C.

(G) Fraction of samples within each category of hidden heterogeneity in human data. Actual values are written on the top of the bar plots cut for readability. See also Figure S4 and Tables S1 and S2.

markers of its emergence and extent are far from being characterized. Evolutionary parameters, such as alteration and proliferation rates, remain poorly documented in human tumors, and evolutionary architectures can at best be algorithmically inferred. From this perspective, mathematical modeling of cancer progression provides an ideal framework to explore diverse evolutionary features and their impact on intra-tumor heterogeneity. Here, we investigated through extensive numerical simulations how clonal diversity emerges under a wide range of alteration rate (μ) and fitness (s) parameters, the latter positively associated with a cell replicative potential.

By tracking clonal lineages for all simulated tumors, we first demonstrated that intra-tumor heterogeneity increases with the alteration rate and decreases with fitness and that high intra-tumor heterogeneity invariably corresponds to branching phylogenies. Interestingly, the transition from linear to branched evolution could be accurately parameterized in the space of parameters (μ, s) by $s_t \sim \mu_t^{\frac{3}{4}}$. Given the correlation observed between the number of clones and the extent of branching measured by Tree score, this transition effectively describes the emergence of intra-tumor heterogeneity as a function of the evolutionary parameters, alteration rate, and fitness.

Recently, rare passenger (Bozic et al., 2016) and driver mutations (Chowell et al., 2018) have been predicted to characterize small subclones that could go undetected by standard molecular profiling. Indeed, our simulations revealed that in tumors with high alteration rates and low fitness, the number of clones with a size greater than 1% of the total cell population decreases concomitantly with an increasing presence of small clones below the detectability threshold. Importantly, the broad range of parameters here analyzed allowed to show that the number of observable and hidden clones do not always correlate. This result implies that tumors with the same number of observable clones, for example, inferred from the analysis of human molecular profiles, could exhibit instead very different extent of intra-tumor heterogeneity. To address this limitation, we found that the extent of hidden heterogeneity can be estimated by clone population frequencies, providing a unique opportunity to re-assess inferred clonal diversity in human tumors. Indeed, our results show that all the analyzed cancer types include samples with high extent of hidden heterogeneity, potentially associated with survival differences. It should be noted, that distinct clone population frequencies were observed in simulated tumors with the same number of clones but different values of alteration rate and fitness. This association could thus prove useful to infer evolutionary parameters in human tumors.

Finally, our approach relied on a simple model of clonal evolution that ignored spatial constraints, such as interactions with the microenvironment and mechanical constraints between cells, and it assumed global and constant parameters in each simulation. On the one hand, modifications of the adopted growth model including deleterious effects of passenger mutations, a variable number of expected drivers, or a variable alteration rate during a single simulation all led to concordant results. On the other hand, it will be important in the future to explore alternative models including, for example, spatial models of tumor evolution that could assess the effect of cell-cell interactions or different cell phenotypes on intra-tumor heterogeneity. We envision that with extended models and a growing availability of data from detailed intra-tumor molecular profiling, it will be possible to provide qualitative and quantitative endpoints to systematically characterize the tumor clonal architecture of each patient.

Limitations of the Study

We have limited our modeling framework to a discrete time branching process disregarding spatial and mechanical constraints, as well as the phenotype of subclonal populations. Such aspects are worth considering in the future. Comparison with human tumor samples requires reliable datasets to resolve phylogenetic reconstructions. Tumor phylogenies have been inferred using PhyloWGS whose accuracy depends on the number of mutations as well as the sequencing read depth.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2019.10.018.

ACKNOWLEDGMENTS

The project was supported by the Gabriella Giorgi-Cavaglieri Foundation.

AUTHOR CONTRIBUTIONS

Conceptualization, F.R. and G.C.; Methodology, F.R. and G.C.; Formal Analysis, F.R., M.M., and G.C.; Writing, F.R. and G.C.; Visualization, F.R. and G.C.; Software, F.R.; Supervision, G.C.; Project Administration, G.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 5, 2018
Revised: July 30, 2019
Accepted: October 7, 2019
Published: November 22, 2019

REFERENCES

Andor, N., Graham, T.A., Jansen, M., Xia, L.C., Aktipis, C.A., and Petritsch, C. (2016). Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. Nat. Med. *22*, 10–12.

Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M.C., Kim, J., Reardon, B., et al. (2018). Comprehensive characterization of cancer driver genes and mutations. Cell *173*, 371–385.e18.

Beerenwinkel, N., Antal, T., Dingli, D., Traulsen, A., Kinzler, K.W., Velculescu, V.E., Vogelstein, B., and Nowak, M.A. (2007). Genetic progression and the waiting time to cancer. PLoS Comput. Biol. *3*, e225.

Beerenwinkel, N., Schwarz, R.F., Gerstung, M., and Markowetz, F. (2014). Cancer evolution: mathematical models and computational inference. Syst. Biol. *64*, e1–e25.

Bozic, I., Antal, T., Ohtsuki, H., Carter, H., Kim, D., Chen, S., Karchin, R., Kinzler, K.W., Vogelstein, B., and Nowak, M.A. (2010). Accumulation of driver and passenger mutations during tumor progression. Proc. Natl. Acad. Sci. U S A *107*, 18545–18550.

Bozic, I., Gerold, J.M., and Nowak, M.A. (2016). Quantifying clonal and subclonal passenger mutations in cancer evolution. PLoS Comput. Biol. *12*, e1004731.

Chowell, D., Napier, J., Gupta, R., Anderson, K.S., Maley, C.C., and Sayres, M.A.W. (2018). Modeling the subclonal evolution of cancer cell populations. Cancer Res. *78*, 830–839.

Miller, C.A., White, B.S., Dees, N.D., Griffith, M., Welch, J.S., Griffith, O.L., Vij, R., Tomasson, M.H., Graubert, T.A., Walter, M.J., et al. (2014). SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. PLoS Comput. Biol. *10*, e1003665.

Deshwar, A.G., Vembu, S., Yung, C.K., Jang, G.H., Stein, L., and Morris, Q. (2015). PhyloWGS: reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. Genome Biol. *16*, 35.

Durrett, R., and Moseley, S. (2010). Evolution of resistance and progression to disease during clonal expansion of cancer. Theor. Popul. Biol. *77*, 42–48.

Fidler, I.J. (1978). Tumor heterogeneity and the biology of cancer invasion and metastasis. Cancer Res. *38*, 2651–2660.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Huerta-Cepas, J., Dopazo, J., and Gabaldón, T. (2010). ETE: a python environment for tree exploration. BMC Bioinformatics *11*, 24.

Jamal-Hanjani, M., Wilson, G.A., McGranahan, N., Birkbak, N.J., Watkins, T.B.K., Veeriah, S., Shafi, S., Johnson, D.H., Mitter, R., Rosenthal, R., et al. (2017). Tracking the evolution of non–small-cell lung cancer. N. Engl. J. Med. *376*, 2109–2121.

Jiang, Y., Qiu, Y., Minn, A.J., and Zhang, N.R. (2016). Assessing intratumor heterogeneity and tracking longitudinal and spatial clonal evolutionary history by next-generation sequencing. Proc. Natl. Acad. Sci. U S A *113*, E5528–E5537.

Knudson, A.G. (1971). Mutation and cancer: statistical study of retinoblastoma. Proc. Natl. Acad. Sci. U S A *68*, 820–823.

Komarova, N.L. (2007). Loss- and gain-of-function mutations in cancer: mass-action, spatial and hierarchical models. J. Stat. Phys. *128*, 413–446.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and

the search for new cancer-associated genes. Nature *499*, 214–218.

Levy, S.F., Blundell, J.R., Venkataram, S., Petrov, D.A., Fisher, D.S., and Sherlock, G. (2015). Quantitative evolutionary dynamics using high-resolution lineage tracking. Nature *519*, 181.

McFarland, C.D., Korolev, K.S., Kryukov, G.V., Sunyaev, S.R., and Mirny, L.A. (2013). Impact of deleterious passenger mutations on cancer progression. Proc. Natl. Acad. Sci. U S A *110*, 2910–2915.

McFarland, C.D., Mirny, L.A., and Korolev, K.S. (2014). Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. Proc. Natl. Acad. Sci. U S A *111*, 15138–15143.

McGranahan, N., and Swanton, C. (2017). Clonal heterogeneity and tumor evolution: past, present, and the future. Cell *168*, 613–628.

Michor, F., Iwasa, Y., and Nowak, M.A. (2004). Dynamics of cancer progression. Nat. Rev. Cancer *4*, 197.

Michor, F., Hughes, T.P., Iwasa, Y., Branford, S., Shah, N.P., Sawyers, C.L., and Nowak, M.A. (2005). Dynamics of chronic myeloid leukaemia. Nature *435*, 1267.

Michor, J.F., Leder, K., and Michor, F. (2011). Stochastic dynamics of cancer initiation. Phys. Biol. *8*, 15002.

Mina, M., Raynaud, F., Tavernari, D., Battistello, E., Sungalee, S., Saghafinia, S., Laessle, T., Sanchez-Vega, F., Schultz, N., Oricchio, E., et al. (2017). Conditional selection of genomic alterations dictates cancer evolution and oncogenic dependencies. Cancer Cell *32*, 155–168.e6.

Nowell, P.C. (1976). The clonal evolution of tumor cell populations. Science *194*, 23–28.

Raynaud, F., Mina, M., Tavernari, D., and Ciriello, G. (2018). Pan-cancer inference of intra-tumor heterogeneity reveals associations with different forms of genomic instability. PLoS Genet. *14*, e1007669.

Roth, A., Khattra, J., Yap, D., Wan, A., Laks, E., Biele, J., Ha, G., Aparicio, S., Bouchard-Côté, A., and Shah, S.P. (2014). PyClone: statistical inference of clonal population structure in cancer. Nat. Methods *11*, 396.

Siravegna, G., Mussolin, B., Buscarino, M., Corti, G., Cassingena, A., Crisafulli, G., Ponzetti, A., Cremolini, C., Amatu, A., Lauricella, C., et al. (2015). Monitoring clonal evolution and resistance to EGFR blockade in the blood of metastatic colorectal cancer patients. Nat. Med. *21*, 795–801.

Vermeulen, L., Morrissey, E., van der Heijden, M., Nicholson, A.M., Sottoriva, A., Buczacki, S., Kemp, R., Tavaré, S., and Winton, D.J. (2013). Defining stem cell dynamics in models of intestinal tumor initiation. Science *342*, 995–998.

Yates, L.R., Knappskog, S., Wedge, D., Farmery, J.H.R., Gonzalez, S., Martincorena, I., Alexandrov, L.B., Van Loo, P., Haugland, H.K., and Lilleng, P.K. (2017). Genomic evolution of breast cancer metastasis and relapse. Cancer Cell *32*, 169–184.e7.

ISCI, Volume *21*

Supplemental Information

Dynamic Emergence of Observed and Hidden Intra-tumor Heterogeneity

Franck Raynaud, Marco Mina, and Giovanni Ciriello

Supplemental Information

Supplemental Data items

Figure S1; related to Figure 1

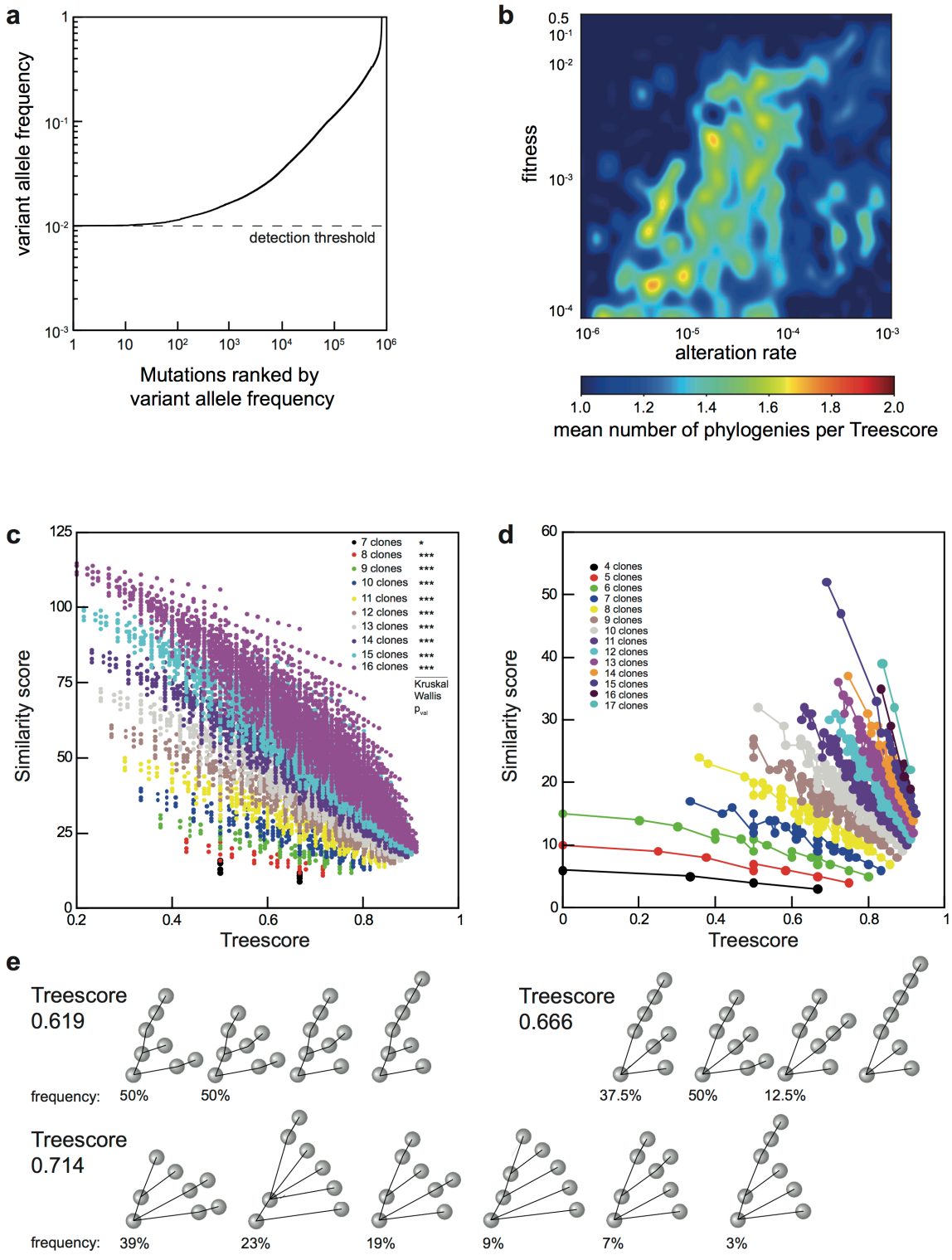


Figure S1. Association between Tree score and topology of phylogenies, Related to Figure 1

a) Estimation of the detection threshold. Rank plot of the variant allele frequencies (VAF) of point mutation in TCGA dataset. No mutations are observed with a VAF lower than 1%, thus defining the detection threshold for the simulations.

b) Mean number of different phylogenies per Tree score obtained by simulated tumors as a function of their alteration rate μ (X-axis) and fitness *s* (Y-axis). For each pair of coordinates, μ and *s*, the mean number of different phylogenies observed in simulations corresponding to those coordinates is color coded (cold colors for low numbers, warm colors for high numbers).

c) Similarity score (sum of the distances of each node to the root) as function of the Tree score for all possible phylogenies with a number of nodes up to 16 (color coded by their respective number of nodes). Kruskal-Wallis test was performed for Tree scores having at least three different phylogenies.

d) Similarity score (sum of the distances of each node to the root) as function of the Tree score for simulated tumors with a number of nodes between 4 and 17 (color coded by their respective number of nodes).

e) Tree representation for simulated tumors with 8 clones with different Tree score and different phylogenies of each Tree score along with their frequencies in the simulated data.

Figure S2; related to Figure 2

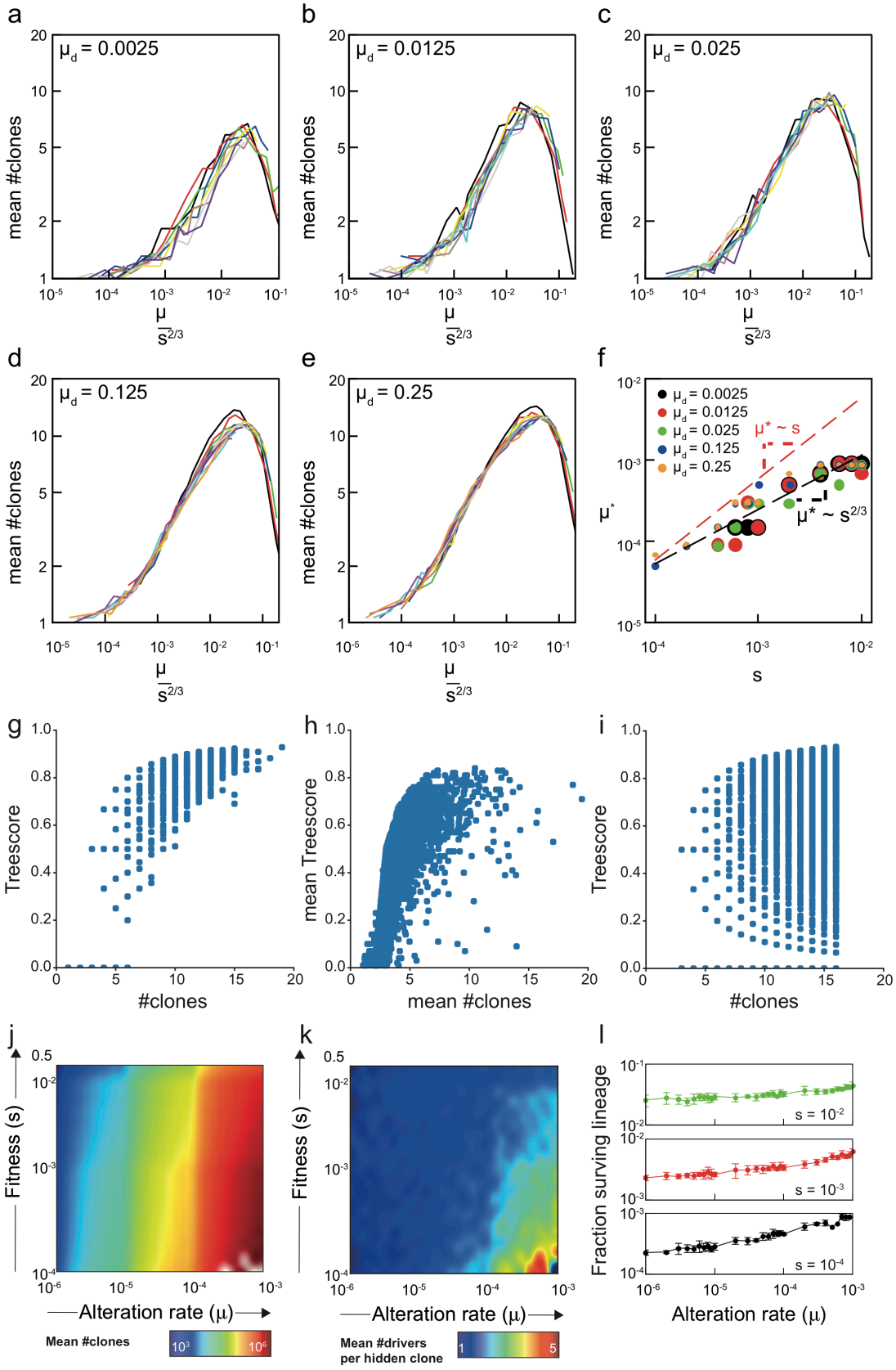


Figure S2. Comparison of different models of cancer evolution and association between number of clones and Tree score, Related to Figure 2

a-e) Mean number of observable clones for different fitness values s as function of the rescaled mutation rate $\frac{\mu}{s^{2/3}}$ for different driver mutation rates 0.0025 (a); 0.0125(b); 0.025(c); 0.125(d) and 0.25 (e)

f) Transition mutation rate μ_t as function of the fitness s for different driver mutations rates. The red dashed line represents the mutation-selection balance $\mu \sim s$; the black dashed line represents the transition line $\mu \sim s^{2/3}$.

g-i) Association between the Tree score and the number of clones for simulated tumors (g), human tumors (h) and all possible phylogenies (i).

j) Mean number of clones obtained by simulated tumors as a function of their alteration rate μ (X-axis) and fitness *s* (Y-axis). For each pair of coordinates, μ and *s*, the mean number of clones observed in simulations corresponding to those coordinates is color coded (cold colors for low numbers, warm colors for high numbers).

k) Mean number of drivers per hidden clone as a function of the alteration rate μ (X-axis) and fitness *s* (Y-axis). Cold colors indicate low number of drivers, warm colors indicate high number of drivers.

l) Fraction of the hidden population with a surviving lineage for different fitness values: $s=10^{-2}$ (top panel), $s=10^{-3}$ (middle panel) and $s=10^{-4}$ (low panel). Data are presented as mean ± SD.

Figure S3; related to Figure 3

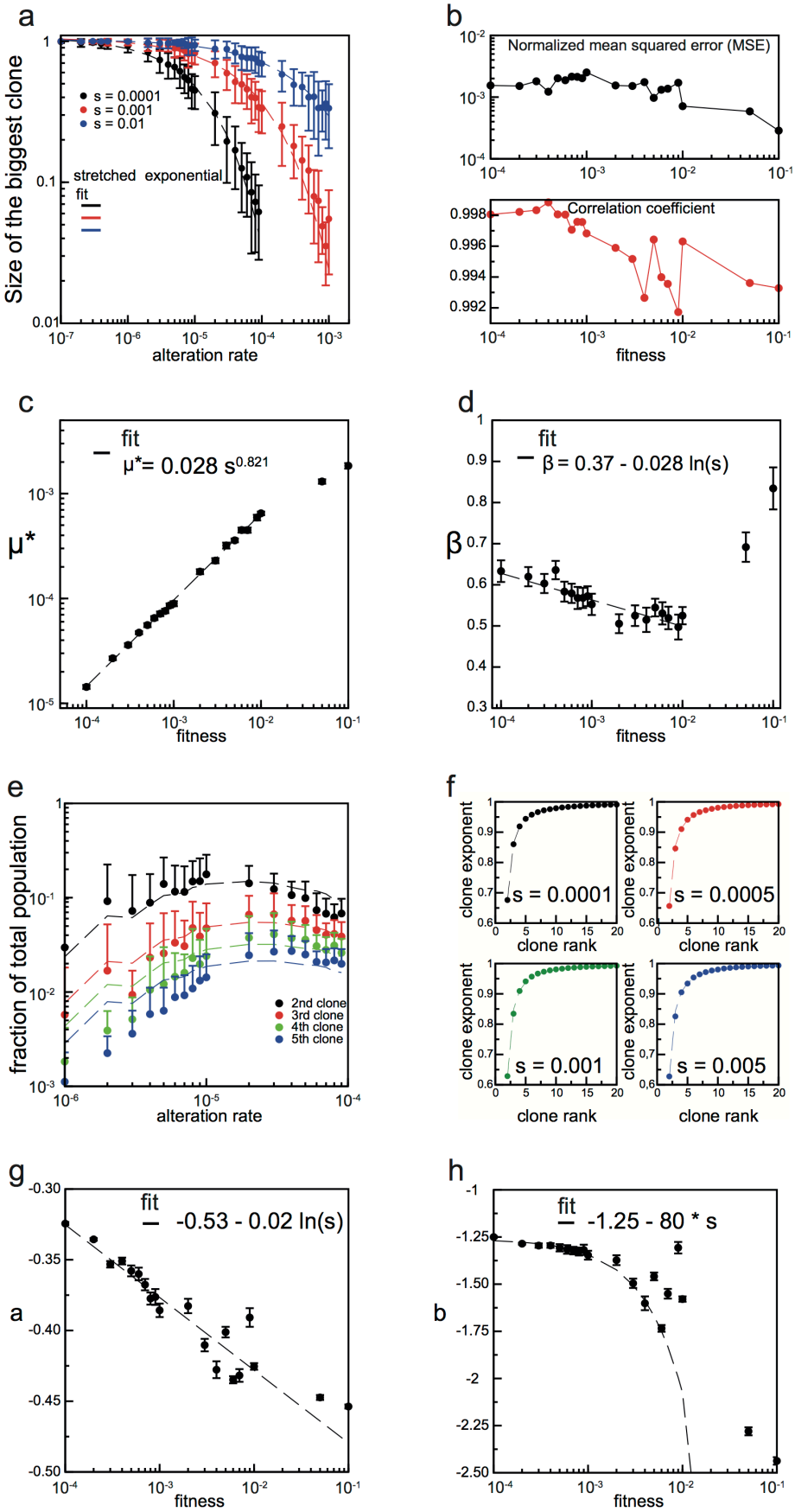


Figure S3. Characterisation of the clone size distribution; Related to Figure 3

a) The normalized size of the biggest clone decreases as a stretched exponential with the alteration rate.

b) Statistics of the stretched exponential fit. Upper panel normalized MSE (in units of clone size), lower panel correlation coefficient R^2 between the fitted equation and the numerical data.

c) The characteristic alteration rate μ^* as function of the fitness.

d) The stretching exponent β as function of the fitness

e) The sizes of the subsequent clones (2th, 3rd, 4th and 5th) are non-monotonic and vary as $C_i = C_1 - C_1^{\alpha}$ where i is the rank and α a function of the rank and the fitness. For clarity, errors are only displayed for the upper bound.

f) The clone exponent α_i converges to 1 as $\alpha_i = 1 + a(s) * i^{b(s)}$.

g) Proportionality constant a(s) as function of the fitness

h) The exponent b(s) as function of the fitness. The fits for μ^*, β, a(s) and b(s) are estimated for s between 10^{-4} and 10^{-2}.

Data are presented as mean ± SD.

Figure S4; related to Figure 4

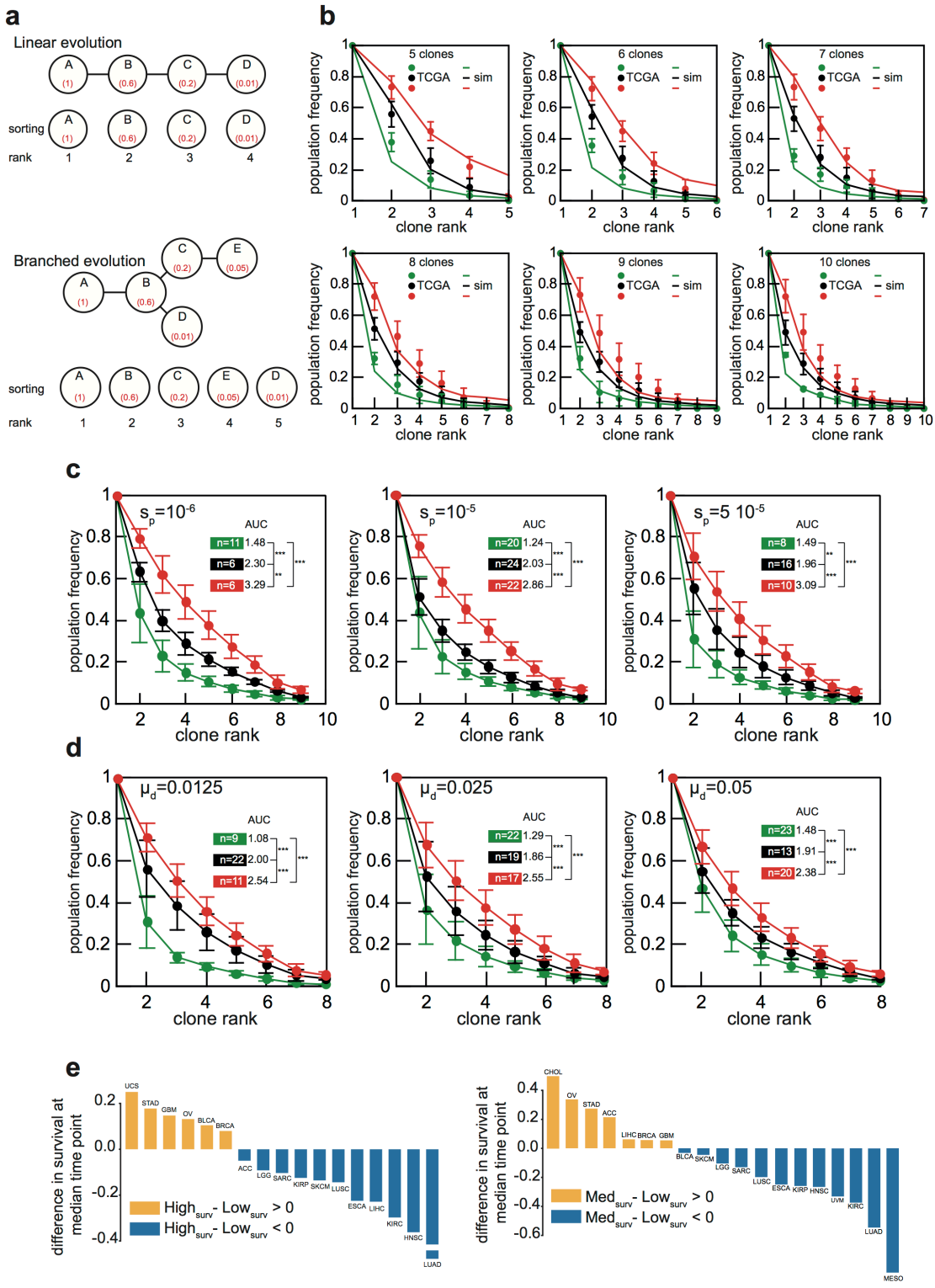


Figure S4. Distribution of population frequency for different models of clonal evolution, related to Figure 4

a) Sorting the clones by population frequency (in red) to estimate the average ranked population frequency for linear evolution and branched evolution. The population frequencies are indicative, not representative of real phylogeny.

b) Comparison between predicted (lines) and observed (points) population frequencies for clonal evolutions with 5 to 10 clones in regions of low hidden heterogeneity (green), intermediate hidden heterogeneity (black) and high hidden heterogeneity (red). Error bars represent the standard deviation. Data are presented as mean ± SD.

c) Distribution of population frequencies in a model of clonal evolution with moderate deleterious passenger mutations with deleterious fitness $s_p = 10^{-6}$; 10^{-5} and 5.10^{-5}. All simulations were done with a driver fitness $s = 10^{-4}$ and mutation rates $\mu = 2.10^{-5}$ for low hidden heterogeneity (green), $\mu = 6.10^{-5}$ for intermediate hidden heterogeneity (black) and $\mu = 1.5\ 10^{-4}$ for high hidden heterogeneity (red). AUC were compared using a two-tailed t-test. P-values less than 0.05 are represented with one asterisk; less than 0.01 with two asterisks; less than 0.001 with three asterisks. Data are presented as mean ± SD.

d) Distribution of population frequencies for the model of clonal evolution with different driver mutation rates $\mu_d = 0.0125$, 0.025 and 0.05. All simulations were done with a fitness $s = 4.10^{-4}$ and mutation rates $\mu = 4.10^{-5}$ for low hidden heterogeneity (green), $\mu = 1.75\ 10^{-4}$ for intermediate heterogeneity (black) and $\mu = 3.10^{-4}$ for high hidden heterogeneity (red). AUC were compared using a two-tailed t-test. P-values less than 0.05 are represented with one asterisk; less than 0.01 with two asterisks; less than 0.001 with three asterisks. Data are presented as mean ± SD.

e) For each tumor type, we compared the fraction of surviving patients at the median time point (median follow-up of the cohort) for patients with high and low hidden heterogeneity (left panel) and intermediate and low hidden heterogeneity (right panel). Each bar is the difference between these two values, negative values are in blue (higher survival in low hidden heterogeneity group).

Subsampling

	5	6	7	8	9	10
μ_1^{SS}	1.09	1.11	1.14	1.13	1.02	1.21
μ_2^{SS}	1.47	1.56	1.65	1.73	1.66	1.86
μ_3^{SS}	1.88	1.99	2,18	2.43	2.49	2.56

	5	6	7	8	9	10
P- μ_1^{SS}, μ_2^{SS}	$2.18\ 10^{-36}$	$3.6\ 10^{-45}$	$7.08\ 10^{-41}$	$1.3\ 10^{-44}$	$3.4\ 10^{-33}$	$3.7\ 10^{-25}$
P- μ_2^{SS}, μ_3^{SS}	$2.77\ 10^{-39}$	$7.5\ 10^{-34}$	$1.19\ 10^{-34}$	$4.8\ 10^{-24}$	$1.3\ 10^{-28}$	$2\ 10^{-14}$
P- μ_1^{SS}, μ_3^{SS}	9.210^{-103}	$1.1\ 10^{-71}$	$8.7\ 10^{-75}$	$1.04\ 10^{-46}$	$4.1\ 10^{-51}$	$3.5\ 10^{-36}$

Bootstrap

	5	6	7	8	9	10
μ_1^{BS}	1.08	1.1	1.15	112	1.09	1.21
μ_2^{BS}	1.41	1.53	1.67	1.77	1.78	1.94
μ_3^{BS}	1.86	1.95	2.16	2.36	2.6	2.74

	5	6	7	8	9	10
P- μ_1^{BS}, μ_2^{BS}	$6.3\ 10^{-55}$	$3.3\ 10^{-74}$	$5.6\ 10^{-64}$	$1.2\ 10^{-70}$	$2.4\ 10^{-44}$	$2.1\ 10^{-58}$
P- μ_2^{BS}, μ_3^{BS}	$3.3\ 10^{-67}$	$3.1\ 10^{-71}$	$7\ 10^{-58}$	$8.2\ 10^{-45}$	$9.3\ 10^{-41}$	$1.1\ 10^{-18}$
P- μ_1^{BS}, μ_3^{BS}	$7.5\ 10^{-129}$	$4.6\ 10^{-140}$	$5.8\ 10^{-114}$	$4.3\ 10^{-76}$	$3.2\ 10^{-66}$	$3.7\ 10^{-39}$

Table S1, Related to Figure 4. Statistical analysis of the distribution of PF-AUC for subsampled data (SS, upper panel) and bootstrapped data (BS, lower panel). Means of the three components of the Gaussian mixtures and pairwise comparisons between the components. We kept for the analysis only human tumor samples with 5 to 10 inferred clones.

Subsampling

	5	6	7	8	9	10
Uniform	1.17	1.14	1.03	0.77	0.87	1.33
GM_3	0.05	0.11	0.12	0.06	0.42	0.23
Normal	0.21	0.30	0.34	0.16	0.46	0.69

Boostrap

	5	6	7	8	9	10
Uniform	1.15	1.14	0.96	0.77	0.95	1.06
GM_3	0.05	0.16	0.15	0.09	0.45	0.28
Normal	0.21	0.29	0.29	0.16	0.52	0.59

Table S2, Related to Figure 4. Kullback-Liebler divergence of the distributions of PF-AUC in human data compared with uniform, three components Gaussian mixtures and normal distribution

Data S1. Summary of the simulations, related to Figures 1, 2, 3 and 4
List of all simulations with label of the simulation, mutation rate, fitness, number of nodes, number of leaves and Tree score.

Data S2. Population frequencies, related to Figure 4
List of population frequencies with: mutation rate, fitness and normalized population frequency of the clones

Transparent Methods

Model of cancer evolution with deleterious passenger mutations

To account for deleterious passenger mutations, we modified equation XXX such as a cell with kd driver mutations and kp passenger mutations the probability to die $d_{kd,kp}$ is given by:

$$d_{kp,kd} = 0.5\ \frac{(1-s_d)^{kd}}{(1-s_p)^{kp}}$$

Where s_d is the driver fitness parameter and s_p is the deleterious fitness parameters.

Determination of the size of the clones

The average size of the biggest clone (normalized by the total population size) varies as a stretched exponential with the alteration rate μ (Figure S3a,b):

$$C_1(s,\mu) = e^{-(\frac{\mu}{\mu^*})^{\beta}}$$

Figure S3c indicates that the characteristic alteration rate μ^* increases as a power law with s, $\mu^* \sim s^{0.821\pm0.007}$ while the stretching exponent β decreases logarithmically with s as $\beta = -0.37 - 0.028\log(s)$ (Figures S3d). The intervals for the two parameters of the fit are 0.37 ± 0.025 and 0.028 ± 0.003. The fits of μ^* and β tend to deviate at large fitness for $s > 0.01$ from the numerical results, however this regime of very high fitness is not the most relevant for our study since it results in monoclonal tumor evolution.

Figure S3e suggests defining the size of clone C_i as:

$$C_i = C_1 - C_1^{\alpha_i}$$

We estimated the exponent α_i by minimizing the squared error $\left(\left(C_1 - C_1^{\alpha_i}\right) - C_i\right)^2$.

It turned out that α_i can be written as function of the fitness and the rank i (Figure S3f):

$$\alpha_i = 1 + a(s) * i^{b(s)}$$

Finally, we found $a(s) = -0.53 - 0.022\log(s)$, with 0.53 ± 0.01 and 0.022 ± 0.0002, and $b(s) = -1.25 - 80s$, with 1.25 ± 0.007 and 80 ± 4.5.

Further investigations are required to understand the values of the different constants obtained from the numerical results, but interestingly we observed that several of them are close to the ratio between the probabilities of driver and passenger alterations $\frac{\mu_d}{\mu_p} \sim 0.025$.

Given these expressions for clone sizes we can analytically estimate the number of clones from the values of μ and s. Given μ and s sampled within the space of parameters used in our simulations, we calculated the size of the biggest and subsequent clones using equations above. The parameters in the equations are chosen from normal distributions with mean and variance corresponding to the values obtained from the fits. Only clones that satisfied the constraint $C_i > 0.01$ were counted. We repeated this procedure as many times as the total number of simulations and estimate the mean number of clones for each couple of values μ, s.

Inference of tumor phylogenies: PhyloWGS, numerical procedure and scoring

PhyloWGS provides inference of evolutionary relationships between clonal subpopulations using both variant allele frequencies of point mutations and copy number alterations at the mutated loci. In particular, PhyloWGS does not compute a unique tree representing the phylogenetic evolution of the tumor, but a series of trees each scored by its complete-data log likelihood (1). For each sample in our dataset, we made 10 runs with different seeds and kept the 50 trees with the best complete-data log likelihood for each run for a total of 500 phylogenies. We then ordered the trees and retained the top 10% (50 trees) to assign a score S_{50}^{i} to each tree i according to:

$$S_{50}^{i} = \frac{CDLL_{50}^{i} - \min(CDLL_{50})}{\max(CDLL_{50}) - \min(CDLL_{50})}$$

where $CDLL^{i}_{50}$ is the complete-data log likelihood of the tree i and $\min(CDLL_{50})$ (resp. $\max(CDLL_{50})$) is the minimum (resp. maximum) complete-data log likelihood value within the reduced set of trees. We then use the aforementioned score to weight the population frequency of each in each sample.

While PhyloWGS was originally designed for whole genome sequencing, the authors demonstrated using simulated data that it did not require the typical amount of mutations from whole genome sequencing. Instead, they show that PhyloWGS is able to confidently reconstruct phylogenies provided that the number of mutations and the read depth are high enough. In most of the cases, our TCGA dataset fulfills the both conditions.

Similarity of tumor phylogenies with identical Tree score

The Tree score defined in the main text (equation 3) assesses whether a phylogeny is branched or linear, but does not identify uniquely a phylogeny. In other words, different phylogenies can receive the same Tree score. While there is a degeneracy of the Tree score, this degeneracy remains moderate (Figure S1b). To investigate the similarity between phylogenies with the same Tree score, we separated all the possible trees according to their number of nodes and measured their variance of the root to leaves distances (VRLD). At fixed number of nodes, the pair of values {Tree score, VRLD} allowed identifying each phylogeny. We manually checked (up to 9 clones) that all phylogenies with the same Tree score received different pair of values {Tree score, VRLD}. We used as a representative quantity of the tree structure the sum of the distances of each node from the root. The underlying idea is that trees receiving close values of the sum of the distances have similar tree structures. Considering that nodes do not have labels or features it was the most straightforward measure of similarity. We then compared the values of the sum of the distances for all the possible Tree scores using a Kruskal-Wallis test. For this test, a group is composed of all different phylogenies having the same Tree score and for efficient statistical analysis we retained Tree scores with at least three different phylogenies.

Our statistical analysis showed that at fixed number of clones, the sums of the distances were significantly different between groups (Figure S1c). Furthermore while aggregating the values for all number of clones (and comparing the mean of nodes to root distances) the difference between Tree scores remained significant. Thus phylogenies with the same Tree score are more similar than phylogenies with a different Tree score. The Tree score appear to be representative of the tree structure.

Average ranked population frequency (*PF)* in human samples

To estimate the average ranked population frequency PF, we first sorted the clones inferred by PhyloWGS by their population frequency. For a linear evolution the ranks of the sorted clones are the same as the ranks of the clones from the root to the leaf. However, for phylogenies with multiple branches (or various phylogenies inferred by PhyloWGS with different topologies and/or number of clones) the rank of the clones from the root to the leaves is ambiguous. For this reason, we sorted the clones by their population frequency independently of the underlying topology of the phylogeny (Figures S4a,b).
After inferring the reduced set of best trees for each human sample and sorting the clones of each tree by their population frequencies, we calculated the average ranked population frequency as:

$$PF(i) = \{PF_i\}_{i=1,2,3\ldots} = \left\langle \frac{1}{\sum_{i=1}^{50} S_{50}^{i}} \sum_{j=1}^{50} S_{50}^{j}\, \widetilde{PF_i^J} \right\rangle$$

where i is the rank of the clone, $\langle\ \rangle$ denotes the ensemble average over all samples, S_{50}^{j} is our PhyloWGS score and $\widetilde{PF_i^J}$ is the population frequency of the clone i of the *jth* inferred phylogeny ($\sim$ indicates that the clones have been sorted by their population frequency in the phylogeny j). We should note that to compare the ranked population frequency in simulations and in human samples at a given number of clones (Figures S4c-h), we considered for the human samples the maximum number of clones inferred by PhyloWGS and not the average number of clones. Furthermore, human samples with a maximum PF lower than 0.75 were removed from this analysis, as these tumors are likely characterized low purity (for 100% tumor content, the max PF is equal to 1).

Bootstrap and subsampling of Population Frequency Area Under the Curve (PF-AUC)

We aim to test whether the distribution of PF-AUC values in human samples is consistent with a trimodal distribution and differs from uniform and normal distribution.

To do so, we first split the samples by their number of clones and then assigned each sample to a group (A, B and C) based on its extent of hidden heterogeneity by matching its distribution with the closest distribution from simulated tumors. We obtained N_A^i samples in the group A, N_B^i in B and N_C^i in C, where i is the number of clones.

We selected randomly in each group N_{SS} (resp. N_{BS}) samples for the subsampling (resp. bootstrap), with $N_{SS} = \min(N_A^i, N_B^i, N_C^i)$ and $N_{BS} = \max(N_A^i, N_B^i, N_C^i)$.
For both the subsampling and the bootstrap, we fit the distributions with a Gaussian mixture of three components using Scikit-learn machine learning library (2).

We repeated the procedures of subsampling and bootstrapping 50 times and then compared among each procedure the means of the component of the mixtures $\mu^{SS}_{\{1,2,3\}}$, $\mu^{BS}_{\{1,2,3\}}$ using a two tailed t-test (Table S1).

Finally, we compared the resulting bootstrap and subsampling distributions with the three components Gaussian mixtures, normal and uniform distribution using the Kulback-Liebler divergence D_{KL} (3):

$$D_{KL}^{SS,BS}(p^{SS,BS}||q^{SS,BS}) = \sum_i p^{SS,BS}(i) \log\left(\frac{p^{SS,BS}(i)}{q^{SS,BS}(i)}\right)$$

where $p^{SS,BS}$ correspond to the subsampling and bootstrap distribution of PF-AUC of human samples and $q^{SS,BS}$ is either the Gaussian mixtures, uniform or normal (with mean and variance estimated from human samples) distributions for both subsampling and bootstrap. The smaller D_{KL} is, the closer informationally p and q are (Table S2).

TCGA human dataset

Data for the tumor types analyzed in this study had been collected from:
Raynaud F et. al (2018) Pan-cancer inference of intra-tumor heterogeneity reveals associations with different forms of genomic instability. PLOS Genetics 14(9): e1007669

Model of cancer evolution

https://github.com/CSOgroup/Model_Cancer_Evolution