Support for a model of neutral evolution in currently available bulk sequencing data

Benjamin Werner[1], Marc J. Williams[2-4], Chris P. Barnes[3,5], Trevor A. Graham[2,*] & Andrea Sottoriva[1,*]

[1] Evolutionary Genomics & Modelling Lab, Centre for Evolution and Cancer, The Institute of Cancer Research, London, UK.
[2] Evolution and Cancer Laboratory, Barts Cancer Institute, Queen Marry University of London, London, UK.
[3] Department of Cell and Developmental Biology, University College London, London, UK.
[4] Centre for Mathematics and Physics in the Life Sciences and Experimental Biology (CoMPLEX), University College London, London, UK.
[5] Department of Genetics, Evolution and Environment, University College London, London, UK.

[*] Correspondence should be addressed to T.A.G. (t.graham@qmul.ac.uk) or A.S. (andrea.sottoriva@icr.ac.uk)

In their letter, McDonald, Chakrabarti & Michor question our assertion that the distribution of mutations in tumour bulk sequencing data suggests an underlying neutral evolutionary process in a proportion of cancers (Williams, Werner *et al.* 2016[1]), and instead propose alternative explanations that incorporate subclonal selection. We agree with the authors' demonstration that it is possible, in principle, to construct models of selection that produce patterns similar to the neutral model. However, the key issue is whether the proposed models of selection are realistic, meaningful, and most importantly more appropriate than the null neutral model. Before examining this issue, we first note that we extensively stressed in the Williams, Werner *et al.* 2016 manuscript that the majority of cases we examined were not consistent with neutral evolution (~70% appeared non-neutral), and we did cite specifically Gerlinger et al. 2012[3] as an example of data dominated by selection[1]. Our finding that the majority of cancers do show evidence of subclonal selection is consistent with previous literature, including the cases highlighted by McDonald, Chakrabarti & Michor[2,3].

Arguably, clonal evolution results from the interplay of three fundamental processes: random alterations (genetic, epigenetic, etc…), random drift, and non-random selection, the third being the most complex to define and model. In the established field of population genetics, extensive effort has been dedicated to model the first two processes without selection, the so-called neutral dynamics[4-6]. This includes the development of entire statistical frameworks based on neutrality, such as coalescent theory[7]. On the contrary, models that include selection, especially in growing populations, have been much harder to derive analytically due to the large number of assumptions in the definition of selection, including whether selection is clone-intrinsic or clone-extrinsic (microenvironmentally defined), and whether the magnitude of selection is constant or fluctuating in response to population dynamics. Importantly, most models of selection describe cancer dynamics in terms of time[8,9] (e.g. time to fixation of a selected mutant) and therefore, although insightful, are hard to apply to cancer genomic data where temporal dynamics are often unobservable.

In the light of this complexity, in our study we asked the simple question: what happens to the mutations in a growing tumour in the case where only the first two processes above, namely random mutations and drift, are operating? This leads to a relatively simple model, which is analytically tractable, wherein subclonal mutations accumulate following a 1/f cumulative distribution[1]. We note that this is the underlying solution of the fully stochastic Luria-Delbrück model, as previously demonstrated[10,11]. Importantly, this model is based on the 'null hypothesis' of molecular evolution in cancer[12-14] and predicts what the *absence* of subclonal selection should look like in a growing tumour. We tested this hypothesis against subclonal mutations from large body of sequencing data and found that in about 30% of cases we could not reject this null hypothesis, at least within the resolution of the currently available data.

In their letter, McDonald, Chakrabarti & Michor and colleagues propose a more complex scenario that includes on-going selection and report that in some cases their model also fits the 1/f cumulative distribution. First, we examine the fit of their proposed model to the data, and highlight that considering the stochastic nature of selection mutants would change the interpretation of their analysis. Second, we discuss the distinction between evaluating the power of a test versus the limitations of the information content in thedata the test is applied to, in this case single-sample bulk sequencing. Third, we analyse the plausibility of the authors' biological assumptions underlying their model.

McDonald, Chakrabarti & Michor's letter shows that in a considerable proportion of simulations with subclonal selection, neutrality was correctly rejected (R^2<0.98; their Figure 1b). The exact proportion of cases incorrectly classified as neutral is not reported, but a few specific examples are shown in their Figures 1c-f. Importantly, in those cases, the mutant proportion at the time of sampling is not reported, nor the time when the mutant was introduced. Both are key factors in judging the strength of the selection signal, for two reasons: (1) in the case of strong and early selection, wherein a selected mutant sweeps to fixation, the evolutionary dynamics revert to neutral, and hence accepting the null for the final tumour is correct (as all cells in the tumour bear the selected mutation, so there is no subclonal selection). (2) due to the inherent stochasticity of the evolutionary process, selected mutants can occur either too late to grow to a detectable size or are weakly selected so that the clonal population of the tumour remains virtually unchanged with respect to the neutral expectation. Judging from the authors' Figure 2A this seems to be what happens often: most mutants have fitness slightly higher than one (1=neutral), and many even lower than one (should be negatively selected), but all persist in the population. In such model it is clear that selection is not sculpting the population by removing unfit clones and benefitting fitter ones, since any mutant fit or unfit, seems to survive. Thus, the dynamics described in the models of McDonald, Chakrabarti & Michor are "effectively neutral", and relatedly it is not surprising that deviations from neutrality are hard to detect.

We highlight that it is fundamentally important to consider the size of differentially selected subclones when considering whether or not a tumour can be classified as neutrally evolving or not. In the authors' second simulation model (their Figure 2), many clones arise very late and are therefore undetectable in the data (high frequency of red dots representing clone size of one cell in their Figure 2A). We argue that no test will ever be able to detect a subclone made of a single cell in a whole malignancy – and indeed it is debatable whether a clone of size 1 can even be considered to have been selected. We discuss the detection limits imposed by current data in our original manuscript[1] (Figure 5), as well as in subsequent work[15,16].

To demonstrate the impact of subclone size in determining whether a tumour is classified as (effectively) neutral or not we performed a more thorough analysis of our previous model of a stochastic branching process under selection (Figure 1 in this article). These simulations show that, in the presence of a subclone of detectable size in the data (e.g. not too small to be out of the frequency distribution, and which has not swept through the whole tumour), the 1/f test is powered to rejected neutrality (1/f test calculated over the frequency range [0.05-0.5] of subclonal mutations from simulated diploid tumours – Figure 1).

McDonald, Chakrabarti & Michor also suggest that improved fits to a 1/f distribution are found in larger populations, irrespective of the underlying model. This assertion is based on the three individual examples presented in figure 2B-C, but appears to be contradicted by their figure 2D, which summarises the results of 25 simulations across different selection regimes and no difference in the distribution of goodness-of-fit values is evident between small or large tumours (no tests of significance were reported). Aside, the frequency interval for inferences is also changed 20-fold between realisations of their models (Figure 2B, C, E), making comparisons difficult.

In general, we agree with the authors' suggestion that R^2 values are not the optimal measurement of fit of a cumulative distribution. Moreover, we note that a limitation of the 1/f statistical test is the

sensitivity to the choice of integration range in the variant allele frequency (VAF) distribution. To be optimal, the 1/f test should be applied to subclonal mutations only, and the whole detectable frequency spectrum should be used (e.g. for ~100x depth sequencing, from a minimum of 5% to a maximum of 50% VAF in a diploid tumour – Figure 1). We note that in follow up work we have developed more sensitive tests, as well as a Bayesian model selection method that uses the whole VAF distribution to directly compare neutrality vs selection[16]. Our original threshold of R^2<0.98 to reject neutral evolution as an explanation for the data was *ad hoc* but based on an estimation of the variability in the cumulative VAF due to stochasticity into the neutral evolution process together with the stochasticity caused by library preparation and moderate depth sequencing.

Next, we consider key biological assumptions of McDonald, Chakrabarti & Michor's proposed models.

Both models include a strong assumption that almost all mutations affect fitness**.** In the first scenario (their Figure 1A) all mutations increase fitness in an additive manner, implying that all mutations in a tumour would be categorised as 'drivers'. This would include non-coding mutations (>98% of all mutations in a cancer genome) and even synonymous variants (~25% of all exonic mutations). This assumption is at odds with what we currently know about the human genome and with a whole body of evidence from large scale cross-sectional genomic studies that have identified a relatively small number of driver alterations[17] amongst a large number of passengers[18]. Consequently, we think this is an implausible assumption. Furthermore, recognizing that most point mutations in cancer genomes are passengers is key to performing subclonal analysis from bulk samples[19,20] as it is the increase in frequency of passenger mutations hitchhiking within a selected mutant that allows to detect subclones.

A second questionable assumption of the presented subclonal selection models is that of *infinite improvement*. In this model, populations evolve by climbing a fitness slope and leading to linear evolution. The "fitness landscape" interpretation has now replaced this model[21], wherein the fitness is defined by a complex landscape of peaks and valleys corresponding to distinct phenotypes in the population, which is a model that is consistent with clonal evolution as well[22]. We note that some of the authors have recently rejected an infinite improvement model as unsupported by their data in their interesting recent publication[23].

In summary, we argue that if the data are equally consistent with the simple mechanistic null model incorporating cell division, cell death, mutation and absence of selection, and also a complex model of selection, then by Occam's razor, we should not reject the null in favour of the complex alternative. Thus, when faced with only a 1/f distribution as evidence, we think it would be a fallacy to assume the complex model rather than the neutral model produced it.

We do agree however that it is paramount to recognise the limitations of currently available data. Specifically, single-sample bulk sequencing data at moderate depth are intrinsically limited in capturing the subclonal evolutionary dynamics and intra-tumour heterogeneity of a tumour. To better measure evolutionary dynamics in cancer, we advocate the need for better data, such as extremely deep sequencing, multi-region sampling and ultimately, single-cell point mutation sequencing.

Figure 1. Sensitivity of the 1/f test to subclone cancer cell fraction. (A) We simulated tumour evolution using a stochastic birth-death process as in Williams, Werner *et al.* 2016, where we could either model a neutral evolutionary process or a process whereby a selected subclone reaches a certain proportion, or cancer cell fraction (CCF) in the tumour. We chose parameters in our model that matched the characteristics of our TCGA colon cancer cohort, namely that we observe around 300 mutations per exome per sample. Simulation parameters: mutation rate=8 per division, cellularity=1, ploidy=2, read depth=100X, birth rate=ln(2), death rate=ln(2)/2. To implement selection, a cell at a random time t_{event} is given a random selection coefficient s, which either decreases the death rate or increases the birth rate. We ran 1000 neutral simulations and for simulations with selected subclones we ran 1000 simulations for each of the following resultant subclone cancer cell fraction, CCF: i) 0.1<CCF<0.9, ii) 0.2<CCF<0.8, iii) 0.3<CCF<0.7. The

distribution of R^2 values (1/f integrated over VAF range [0.05,0.5] of tumour subclonal mutations only) is significantly different between neutral and non-neutral tumours (green line shows 0.98 cutoff). Box plots show the median and IQR; the upper whisker is the 3rd quantile + 1.5 × IQR; and the lower whisker is the 1st quantile – 1.5 × IQR. **(B)** A ROC analysis also shows that the R^2 metric has more discriminatory power for subclones that are in the centre of the VAF distribution, as one would expect since their subclonal cluster becomes more evident.

Author contributions

B.W. performed mathematical analysis. M.W. performed simulation analysis. All authors participated in the discussion, conceived and designed the response, and wrote the manuscript.

Acknowledgments

A.S. is supported by The Chris Rokos Fellowship in Evolution and Cancer and by Cancer Research UK (A22909). T.A.G. is supported by Cancer Research UK (A19771). C.P.B. is supported by the Wellcome Trust (097319/Z/11/Z). B.W. is supported by the Geoffrey W. Lewis Post-Doctoral Training fellowship. A.S. and T.A.G. are jointly supported by the Wellcome Trust (202778/B/16/Z and 202778/Z/16/Z respectively). This work was also supported by Wellcome Trust funding to the Centre for Evolution and Cancer (105104/Z/14/Z).

Data Availability Statement

Simulation code is available in https://github.com/marcjwilliams1/neutral-tumour-evolution-werner-2018.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Williams, M. J., Werner, B., Barnes, C. P., Graham, T. A. & Sottoriva, A. Identification of neutral tumor evolution across cancer types. *Nature Genetics* **48,** 238–244 (2016).
2. Ding, L. *et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481,** 506–510 (2012).
3. Gerlinger, M. *et al.* Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *New England Journal of Medicine* **366,** 883–892 (2012).
4. P Donnelly, A. & Tavaré, S. Coalescents and Genealogical Structure Under Neutrality. *Annual Review of Genetics* **29,** 401–421 (2003).
5. Griffiths, R. C. & Tavaré, S. The age of a mutation in a general coalescent. *Communications in Statistics. Part C: Stochastic Models* **14,** 273–295 (1998).
6. Durrett, R. Population genetics of neutral mutations in exponentially growing cancer cell populations. *The Annals of Applied Probability* **23,** 230–250 (2013).
7. Griffiths, R. C. & Tavaré, S. The age of a mutation in a general coalescent tree. *http://dx.doi.org/10.1146/annurev-ecolsys-102209-144621* **14,** 273–295 (2007).
8. Beerenwinkel, N. *et al.* Genetic progression and the waiting time to cancer. *PLoS Comput. Biol.* **3,** e225 (2007).
9. Bozic, I. *et al.* Accumulation of driver and passenger mutations during tumor progression. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 18545–18550 (2010).
10. Kessler, D. A. & Levine, H. Large population solution of the stochastic Luria-Delbruck evolution model. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 11682–11687 (2013).
11. Kessler, D. A. & Levine, H. Scaling solution in the large population limit of the general asymmetric stochastic Luria-Delbrück evolution process. *J Stat Phys* **158,** 783–805 (2015).
12. Wu, C.-I., Wang, H.-Y., Ling, S. & Lu, X. The Ecology and Evolution of Cancer: The Ultra-Microevolutionary Process. *Annual Review of Genetics* (2016). doi:10.1146/annurev-genet-112414-054842
13. Niida, A., Iwasaki, W. M., Innan, H. & Kumar, S. Neutral Theory in Cancer Cell Population Genetics. *Mol. Biol. Evol.* **76,** 5605 (2018).
14. Cannataro, V. L. & Townsend, J. P. Neutral theory and the somatic evolution of cancer. - PubMed - NCBI. *Mol. Biol. Evol.* **47,** 1402 (2018).

15. Sun, R. *et al.* Between-region genetic divergence reflects the mode and tempo of tumor evolution. *Nature Genetics* **49,** 1015–1024 (2017).
16. Williams, M. J. *et al.* Quantification of subclonal selection in cancer from bulk sequencing data. *Nature Genetics* (2018). In press.
17. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339,** 1546–1558 (2013).
18. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458,** 719–724 (2009).
19. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149,** 994–1007 (2012).
20. Griffith, M. *et al.* Optimizing Cancer Genome Sequencing and Analysis. *Cell Systems* **1,** 210–223 (2015).
21. Orr, H. A. Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.* **10,** 531 (2009).
22. Lipinski, K. A. *et al.* Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. *Trends in Cancer* **2,** 49–63 (2016).
23. Gao, R. *et al.* Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nature Genetics* (2016). doi:10.1038/ng.3641