Zafar *et al. Genome Biology* (2019) 20:95
https://doi.org/10.1186/s13059-019-1692-5

Genome Biology

CORRESPONDENCE Open Access

Comments on the model parameters in "SiFit: inferring tumor trees from single-cell sequencing data under finite-sites models"

Hamim Zafar[1,2], Anthony Tzen[1], Nicholas Navin[2,3], Ken Chen[2*] and Luay Nakhleh[1*]

Derivation of matrix *Q* in Equation (5) in Methods

Each site in the (diploid) genome can be in one of five states, 0/0, 0/-, 0/1, 1/-, or 1/1, where 0 denotes reference allele, 1 denotes variant allele, and - denotes a deletion.

We introduce two rate parameters:

- λ_d: the rate of recurrent point mutation at a site.
- λ_l: the combined rate of deletion and loss of heterozygosity (LOH).

Using these five states and rate parameters, we have the following instantaneous rate matrix *Q*:

We now abstract the genotypes as follows:

- Genotype 0 corresponds to states 0/- and 0/0.
- Genotype 1 corresponds to state 0/1.
- Genotype 2 corresponds to states 1/- and 1/1.

Under this abstraction and the assumptions detailed in the caption of Table 1, we obtain the matrix *Q* given in Eq. (5) in the main text.

The following clarifications also apply to the main text:

- Given the explanation above, the following statement should be removed: "LOH events can result in the genotype transitions 1 → 0 and 1 → 2 whereas deletions can result in the genotype transitions 1 → 0, 1 → 2 or 2 → 1. To compute the infinitesimal rates for these transitions, we introduce two parameters λ_d and λ_l that account for the effects of deletion and LOH respectively."
- The statement
  "It is important to note that out of the three different types of events that could hint at a deviation from the infinite-sites assumption, SiFit currently models events (deletions, LOH, etc.) that affect the same genomic site more than once and the FP and FN errors in SCS data."
  should be replaced by
  "It is important to note that out of the three different types of events that could hint at a deviation from the infinite-sites assumption, SiFit currently models events (recurrent point mutations, deletions, LOH, etc.) that affect the same genomic site more than once and the FP and FN errors in SCS data."
- The statement
  "These parameters being relative quantities (they denote the rates of deletion and LOH, respectively, relative to the rate of point mutations), we choose a beta distribution as their prior."
  should be replaced by
  "These parameters being relative quantities (they denote the rates of recurrent point mutation and deletion/LOH, respectively, relative to the rate of point mutations), we choosea beta distribution as their prior."

Table 1 Expanded *Q* matrix for ternary data

	0/-	0/0	0/1	1/-	1/1
0/-	0	NA	NA	NA	NA
0/0	λ_l	-1 - λ_l	1	NA	NA
0/1	$\frac{\lambda_l}{2}$	$\frac{\lambda_d}{2}$	$-(\lambda_d + \lambda_l)$	$\frac{\lambda_l}{2}$	$\frac{\lambda_d}{2}$
1/-	NA	NA	NA	0	NA
1/1	NA	NA	λ_d	λ_l	$-\lambda_l - \lambda_d$

$Q(i,j)$ denotes transition from state *i* to state *j*. The transitions for which the entry is 'NA', are not allowed. In particular, we do not allow the transitions 0/- → 1/- or 1/- → 0/- as a reflection of a simplifying assumption that a recurrent point mutation and deletion/LOH occurring at the same site is a very rare event. Furthermore, we do not model copy number gain.

---

*Correspondence: KChen3@mdanderson.org; nakhleh@rice.edu
[2] Department of Bioinformatics and Computational Biology, the University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA
[1] Department of Computer Science, Rice University, Houston, Texas, USA
Full list of author information is available at the end of the article

Authors' contributions
All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Author details
[1]Department of Computer Science, Rice University, Houston, Texas, USA. [2]Department of Bioinformatics and Computational Biology, the University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA. [3]Department of Genetics, the University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA.

Received: 9 April 2019 Accepted: 9 April 2019