frontiers in Physiology

ORIGINAL RESEARCH
published: 23 August 2021
doi: 10.3389/fphys.2021.596194

Computational Reconstruction of Clonal Hierarchies From Bulk Sequencing Data of Acute Myeloid Leukemia Samples

Thomas Stiehl[1,2] *and Anna Marciniak-Czochra*[2]*

[1] *Institute for Computational Biomedicine – Disease Modeling, RWTH Aachen University, Aachen, Germany,* [2] *Institute of Applied Mathematics, Interdisciplinary Center for Scientific Computing and Bioquant Center, Heidelberg University, Heidelberg, Germany*

OPEN ACCESS

Edited by:
Doron Levy, University of Maryland, College Park, United States

Reviewed by:
Alexandra Jilkine, University of Notre Dame, United States
Torbjörn Lundh, Chalmers University of Technology, Sweden

***Correspondence:**
Anna Marciniak-Czochra
anna.marciniak@iwr.uni-heidelberg.de

Specialty section:
This article was submitted to Computational Physiology and Medicine, a section of the journal Frontiers in Physiology

Received: *18 August 2020*
Accepted: *26 July 2021*
Published: *23 August 2021*

Citation:
Stiehl T and Marciniak-Czochra A (2021) Computational Reconstruction of Clonal Hierarchies From Bulk Sequencing Data of Acute Myeloid Leukemia Samples. Front. Physiol. 12:596194. doi: 10.3389/fphys.2021.596194

Acute myeloid leukemia is an aggressive cancer of the blood forming system. The malignant cell population is composed of multiple clones that evolve over time. Clonal data reflect the mechanisms governing treatment response and relapse. Single cell sequencing provides most direct insights into the clonal composition of the leukemic cells, however it is still not routinely available in clinical practice. In this work we develop a computational algorithm that allows identifying all clonal hierarchies that are compatible with bulk variant allele frequencies measured in a patient sample. The clonal hierarchies represent descendance relations between the different clones and reveal the order in which mutations have been acquired. The proposed computational approach is tested using single cell sequencing data that allow comparing the outcome of the algorithm with the true structure of the clonal hierarchy. We investigate which problems occur during reconstruction of clonal hierarchies from bulk sequencing data. Our results suggest that in many cases only a small number of possible hierarchies fits the bulk data. This implies that bulk sequencing data can be used to obtain insights in clonal evolution.

Keywords: computational algorithm, acute myeloid leukemia, clonal evolution, clonal hierarchy, clonal pedigree, phylogenetic tree, bulk sequencing, stem cell

INTRODUCTION

Acute myeloid leukemia (AML) is an aggressive cancer of the blood forming system. It is characterized by expansion of malignant cells and impairment of healthy blood cell formation (Röllig et al., 2011; Döhner et al., 2017; Roloff and Griffiths, 2018). AML originates from a small population of malignant stem-like cells, referred to as leukemic stem cells (LSC) or leukemia initiating cells (LIC). A hallmark of AML is its poor prognosis and the high rate of relapse (Röllig et al., 2011; Döhner et al., 2017; Roloff and Griffiths, 2018).

The main reason for the high risk of relapse is the clonal heterogeneity of the disease. Sequencing studies reveal that the AML cell population is composed of multiple clones. Contributions of the individual clones to the total malignant cell burden vary over time

(Ding et al., 2012; Cancer Genome Atlas Research Network, 2013; Greif et al., 2018; Cocciardi et al., 2019; Ediriwickrema et al., 2020). Due to the high number of different clones, the probability is high that a subset of clones has a low sensitivity to chemotherapy, survives treatment and initiates relapse (Ding et al., 2012; Cancer Genome Atlas Research Network, 2013; Stiehl et al., 2014; Greif et al., 2018; Cocciardi et al., 2019; Ediriwickrema et al., 2020).

The clinical course of the disease shows a significant among-patient variability which can only be partially predicted based on currently existing risk-stratifications (Stiehl et al., 2014, 2015, 2020; Döhner et al., 2017; Wang et al., 2017; Roloff and Griffiths, 2018). To better understand the mechanism of relapse and to identify patients at risk, a quantitative understanding of clonal dynamics is required (Ding et al., 2012; Cancer Genome Atlas Research Network, 2013; Stiehl et al., 2014; Greif et al., 2018; Banck and Görlich, 2019; Cocciardi et al., 2019; Lorenzi et al., 2019; Ediriwickrema et al., 2020).

Next-generation sequencing studies have revealed a high number of genetic hits involved in AML pathogenesis. Genetic variability among different patients is considerable and new mutations are acquired during disease evolution (Ding et al., 2012; Cancer Genome Atlas Research Network, 2013; Greif et al., 2018; Cocciardi et al., 2019; Ediriwickrema et al., 2020). Correlation of mutations with clinical outcome has resulted in a genetics-based risk-stratification (Grimwade et al., 1998; Röllig et al., 2011; Döhner et al., 2017). However, the effect of many mutations on cell dynamics remains unclear (Bacher et al., 2008; Ding et al., 2012).

Relating genetic data to patient prognosis and malignant cell properties is challenging, since different genetic hits may enhance or inhibit each other (Grimwade et al., 1998; Bacher et al., 2008; Cancer Genome Atlas Research Network, 2013; Stiehl et al., 2014; Greif et al., 2018; Roloff and Griffiths, 2018). Furthermore, potentially unknown or undetected hits may impact the aberrations that are observed in clinical routine. Mathematical and computational models are important to link genetic data to functional cell properties such as proliferation and self-renewal of leukemic stem cells, both of which are of prognostic relevance (Stiehl et al., 2014, 2015, 2016, 2018, 2020; Banck and Görlich, 2019; Lorenzi et al., 2019).

Such models allow to estimate which leukemic cell properties correspond to the clinical course of an individual patient and to link the estimates to mutation data (Stiehl et al., 2014, 2015, 2020). This provides insights into the impact of different mutations and leads to new hypotheses about the underlying biological mechanisms and genotype-phenotype correlation.

Leukemic stem cell dynamics are governed by two key properties: proliferation rate and fraction of self-renewal. The proliferation rate describes how often LSC divide per unit of time. Upon division a LSC gives rise to two progeny which can either be LSC or of a more differentiated progenitor type. The fraction of self-renewal corresponds to the fraction of LSC among the progeny (Lutz et al., 2013; Stiehl and Marciniak-Czochra, 2017). Mathematical and computational models suggest that stem cell properties at diagnosis differ from those at relapse. Particularly, LSC at diagnosis are characterized by an increased self-renewal fraction and a higher proliferation rate compared to healthy cells. LSC at relapse are characterized by a slow proliferation rate and a further increase of the self-renewal fraction (Stiehl et al., 2014, 2016). Computer simulations and model analysis indicate that increased self-renewal leads to a competitive advantage of the respective clones and that clones appearing later in the course of the disease have a higher self-renewal compared to clones emerging earlier (Stiehl et al., 2014, 2016; Busse et al., 2016; Banck and Görlich, 2019; Lorenzi et al., 2019).

Single cell sequencing technology allows to detect mutations that are present in a single cell. Sequencing of a sufficiently large number of single cells allows to reconstruct the order of mutation acquisition and to visualize it as a so-called clonal hierarchy, clonal pedigree or phylogenetic tree (Kuipers et al., 2017; Ediriwickrema et al., 2020). Computational models have led to the hypothesis that the position of a clone in the phylogenetic tree correlates with its fraction of self-renewal (Stiehl et al., 2016). Therefore, phylogenetic trees may contain important information about cell properties that could be used to decipher the impact of mutations on the malignant cell kinetics.

In contrast to the single cell sequencing approach, bulk sequencing analyses a mixture of DNA of multiple cells, to which each cell contributes its specific (either mutated or non-mutated) alleles. Since in most cases each cell carries two versions of each allele, the bulk sample from n cells is a mixture of $2n$ allele versions. The so-called variant allele frequency (VAF) is the percentage of allele versions that is mutated. Bulk sequencing quantifies the frequency of a mutated allele in a cell population however does not determine how the detected mutations are distributed among the different clones (Roth et al., 2014; Kuipers et al., 2017; Brierley and Mead, 2020).

Single cell sequencing is a relatively new and costly technology that so far is not used in clinical routine (Brierley and Mead, 2020). To deduce clinically relevant knowledge from genetic data large patient groups have to be studied due to the high inter-individual heterogeneity of the detected mutations and their unknown interaction. For this reason, it is a relevant question whether clonal hierarchies can be deduced from bulk sequencing data which are routinely obtained after initial diagnosis of AML (Roth et al., 2014; Brierley and Mead, 2020), although most of the diagnostic sequencing is targeted on limited panels of "typical" driver mutations.

In this work we propose an algorithm that systematically constructs all phylogenetic trees that are in agreement with bulk sequencing data of an individual patient. This algorithm provides a tool to better understand the ambiguity of such reconstructions and their sensitivity to measurement errors.

To test our approach, we choose a recently published set of single cell sequencing data as a gold standard (ground truth) (Ediriwickrema et al., 2020). Based on the single cell sequencing data we calculate the variant allele frequency of the different mutations in a bulk sequencing sample and test whether the "real" clonal pedigree, i.e., the pedigree deduced from single cell sequencing data, can be reconstructed from it. We investigate how the correctness and uniqueness of the reconstruction depend on sampling and measurement errors.

Different approaches have been developed to track the order of mutations in AML. They include population based cross sectional studies (Delhommeau et al., 2009; Abdel-Wahab et al., 2010; Papaemmanuil et al., 2016), targeted and deep sequencing of paired samples taken at different time points (Bachas et al., 2012; Ding et al., 2012), single cell sequencing (Ediriwickrema et al., 2020) and others such as fluorescence-*in situ*-hybridization, xenografting, cell cultures or IPS technology (Anderson et al., 2011; Ran et al., 2012; Jonas, 2017; Nobile et al., 2019; Herudkova et al., 2020; Sandén et al., 2020). From the computational side, a range of tools have been developed to fit models and extract quantitative information from data in the context of AML (Attolini et al., 2010; Nobile et al., 2019) and other cancers, see e.g., (Roth et al., 2014; Caravagna et al., 2020) for statistical approaches using variant allele frequencies, (Attolini et al., 2010) for a population-based model and (Nobile et al., 2019) for xenotransplant data. These approaches are complemented by process-based models (Stiehl et al., 2014, 2016; Rahman et al., 2018; Banck and Görlich, 2019; Dinh et al., 2019, 2020; Salichos et al., 2020).

MATERIALS AND METHODS

Aim

We use variant allele frequencies from bulk sequencing as input data. The output we want to obtain are all clonal hierarchies that are compatible with the input data.

Assumptions

We assume that each mutation is only acquired once. Variant alleles cannot mutate back to wild type alleles. We only consider heterozygous mutations. We rescale the measurements such that the variant allele frequency of the most abundant mutation is equal to 100%.

Computational Methods

The method is summarized in **Figure 1**. Assuming that each mutation is irreversible and only acquired once, clonal pedigrees have the structure of labeled rooted trees. An (unrooted) tree is an undirected acyclic connected graph (Diestel, 2017). If one node of the tree is designated as root, a rooted tree is obtained. In a rooted tree we naturally assign directions to the edges pointing from the root towards the leaves. If a unique label is assigned to each node, the tree is referred to as a labeled tree (Diestel, 2017). The root of the tree corresponds to a genetic trait that is present in all clones. If the disease originates from a single founding mutation that is present in all malignant cells, the root can be identified with the founding mutation. This configuration applies to most leukemic patients. If there exist multiple founding mutations the root of the tree corresponds to the healthy phenotype. Each node in the tree corresponds to one clone. The label assigned to a node indicates the mutational events that gave rise to the clone. The edge pointing towards the node indicates which ancestor clone acquired the mutational event indicated by the label.

The tree structures can be mapped to matrices. We consider a tree with n nodes, corresponding to n clones denoted by *clone* 1 to *clone n*. Since each clone differs from its ancestor by exactly one new mutation, there exist n different mutations, which we number from 1 to n. Denote by $A_{1=i,j=n}$ a matrix. We set $a_{ij} = 1$ if clone j carries mutation i, otherwise we set $a_{ij} = 0$. We number the clones starting from the root (= *clone* 1) and proceed with increasing depth, i.e., if the depth of *clone i* is higher than the depth of *clone j*, then $j < i$. We denote the founding mutation as *mutation* 1 and the mutation that is present in *clone j* but not in its direct ancestor as *mutation j*. Then $A_{1=i,j=n}$ is an upper triangular matrix, with $a_{ii} = 1$, and a_{ij} from the set {0,1}.

We aim to solve the linear system of equations $Ax = b$, where b_i is the measured frequency of *mutation i* in the bulk sample and x_i is the abundance of *clone i* in the sample. We note that A has determinant 1 and therefore this system of equations has a unique solution. The solution is biologically feasible if all x_i are non-negative. The existence of a non-negative solution can be easily checked since the solutions of $Ax = b$ are given by $x_n = b_n$, $x_j = b_j - a_{jj+1}\, x_{j+1} - \ldots - a_{jn} x_n$. We say that the dataset b is compatible with the clonal hierarchy represented by matrix A if $Ax = b$ has a non-negative solution.

The founder mutation is denoted as *mutation* 1. It is present in all clones and, therefore, is the most abundant mutation in the bulk sample. This implies that $a_{1j} = 1$ for $1 \leq j \leq n$. Since we normalized the frequency of the most abundant mutation to

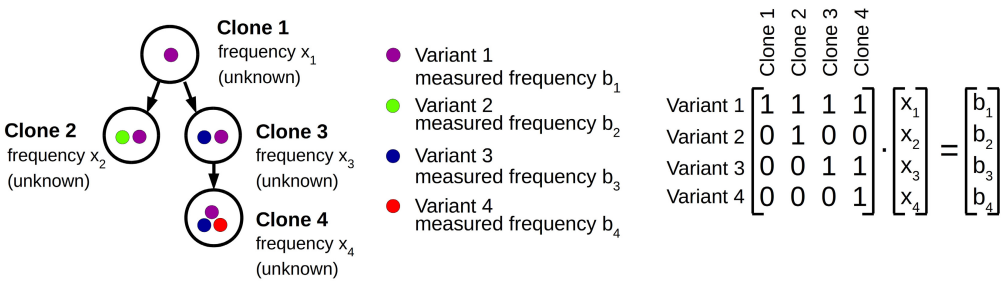

FIGURE 1 | Computational approach. Clonal hierarchies are rooted trees. The root of the tree either corresponds to wild type cells or to the AML founder clone. The colored dots represent different alleles or mutations. From bulk sequencing the allele frequencies b_i in the sample are known. The frequencies x_i of the different clones are unknown. The tree structure is represented by a triangular matrix. The measured data is compatible with the tree structure if the system $Ax = b$ has a non-negative solution.

100% it holds $b_1 = 100$. This implies that the sum over the x_i is equal to 100.

To systematically generate all possible trees, we use Pruefer sequences, a classical concept to bijectively map unrooted trees with n nodes to sequences of length n-2 (Prüfer, 1918). Each unrooted labeled tree with n nodes then corresponds to a sequence of length n-2 with elements from {1, ..., n}. This implies that there exist n^{n-2} unrooted labeled trees. Since each of the n nodes can be designated as root, there exist n^{n-1} labeled rooted trees.

Interpretation

If a biologically feasible solution of the system $Ax = b$ exists, the measured bulk allele frequencies b can be explained by the tree structure that corresponds to the matrix A. This means that the bulk allele frequencies b are obtained by mixing the different clones from the tree in appropriate proportions (the abundance of *clone i* has to equal x_i). For each pair A, b a biologically feasible solution can exist or cannot exist. For example, a tree with founder mutation X (i.e., each clone carries mutation X) cannot match to samples where the abundance of X in non-maximal.

Measurement Errors

If the measured data b are exact, non-existence of a biologically feasible solution indicates a mismatch of the tree structure and the allele frequencies. In case of experimental data, the non-existence of a biologically feasible solution can alternatively arise from measurement errors. For this reason it may be necessary to also consider solutions fulfilling $|| Ax\text{-}b|| < \varepsilon$ for an appropriate ε, where $||.||$ denotes e.g., the Euclidean norm.

To find such solutions, especially in the case where no biologically feasible (i.e., exact non-negative) solutions exist, we use an optimization approach to obtain a non-negative solution that reproduces the data as good as possible. For each matrix A that corresponds to a tree structure we minimize $||$ Ax-b$||$ under the constraints $x_i \geq 0$ (i = 1,...n), $x_1 + ... + x_n = 100$. If the measured VAF have different confidence intervals, we minimize the weighted error function $|| W(Ax\text{-}b)||$, where W is a diagonal matrix with entries related to the confidence intervals.

Solving the minimization problem for each possible tree structure allows to rank the tree structures based on the mismatch $|| Ax\text{-}b||$ and to identify which tree optimally fits to the data. A solution is referred to as exact if $|| Ax\text{-}b|| <$ 10E-16. We say that the tree structure corresponding to matrix $\tilde{A}$ is optimal if it holds $|| \tilde{A}x\text{-}b|| \leq || Ax\text{-}b||$ for all matrices A that represent a suitable tree and vectors x fulfilling $x_i \geq 0$ (i = 1,...n), $x_1 + ... + x_n = 100$. The optimization was carried out using the python cvxopt package (Andersen et al., 2018).

The impact of measurement errors in b on the reconstructed clonal frequencies x can be calculated based on Cramer's rule. For two vectors b, b' and the corresponding solutions x, x' we obtain $A(x\text{-}x') = b\text{-}b'$. Since the determinant of A is equal to one, Cramer's rule implies $x_i\text{-}x_i' = det(A_i)$, where A_i denotes the matrix A with the ith column replaced by $b\text{-}b'$. Consequently, $x_n\text{-}x_n' = b_n\text{-}b_n'$ and $|x_{n-1}\text{-}x_{n-1}'| \leq |b_n\text{-}b_n'| + |b_{n-1}\text{-}b_{n-1}'|$. Analogous formulas can be derived for $i < n\text{-}1$. However, depending on the structure of A, they can be lengthy. Therefore, the use of the condition number of A seems to be more convenient to estimate the errors. It quantifies how perturbations in b impact on the changes of x. For all considered tree structures, the condition numbers of the related matrices computed in the l^2 norm are provided in Section 2 of the Supplement.

Data

We plan to investigate if it is possible to reconstruct clonal hierarchies from bulk sequencing samples. This requires that the "true" clonal hierarchy is known, so that we can compare the result of our algorithm with reality. To know the "true" hierarchy we use single cell sequencing data from ref. (Ediriwickrema et al., 2020). We understand the clonal hierarchy and the clonal frequencies obtained from the single cell sequencing as ground truth. Since for the samples analyzed in Ediriwickrema et al. (2020) no bulk data are available, we calculate the bulk allele frequencies based on the single cell data. For simplicity we assume that the considered sample only contains leukemic cells and we exclude all sequenced wild type cells from the data. We calculate the bulk VAF of variant allele i as $a_{i1}f_1 + ... + a_{in}f_n$, where f_i is the frequency of clone i in the single cell data set and $a_{ij} = 1$ if clone j carries variant allele i and 0 otherwise. Since we consider a purely leukemic sample, the calculated VAF are normalized such that the frequency of the most abundant variant allele is 100%. We consider all patients from Ediriwickrema et al. (2020) that carry only heterozygous mutations and for whom data at diagnosis and relapse is available.

RESULTS

Exact Input Data Often Result in Unique Clonal Hierarchies

As gold standard we use the single cell sequencing data from Ediriwickrema et al. (2020), which provide the true clonal hierarchy and hence can be used to test the proposed algorithm. Based on the single cell data we calculate the variant allele frequencies in the bulk sample. The first question we ask is how many clonal hierarchies are compatible with the bulk variant allele frequencies of a given patient. **Figure 2** shows for each patient which hierarchies exactly fit to the data at diagnosis. We observe that, for 5 out of 6 patients, only one hierarchy exactly fits the bulk data. For one patient 6 hierarchies are consistent with the bulk data.

Similar observations hold for the relapse samples of the considered patients, **Figure 3**. Here all samples, except one (Patient 5) lead to unique tree configurations. In case of patient five all sequenced cells belong to the same clone, which makes it impossible to infer the order of mutations. In the next step we combine the diagnosis and relapse sample of each patient. For each patient **Figure 4** shows the tree configurations that are compatible with the data at both time points. We have uniqueness in all except one case.

To provide insights into the question whether the structure of the true hierarchy (e.g., linear vs. branched) determines how many tree configurations fit to the bulk dataset, we perform a computational experiment. We consider all tree structures for n = 4. For each of them, we generate 10000 bulk data sets by a random distribution among the different clones. Then, for

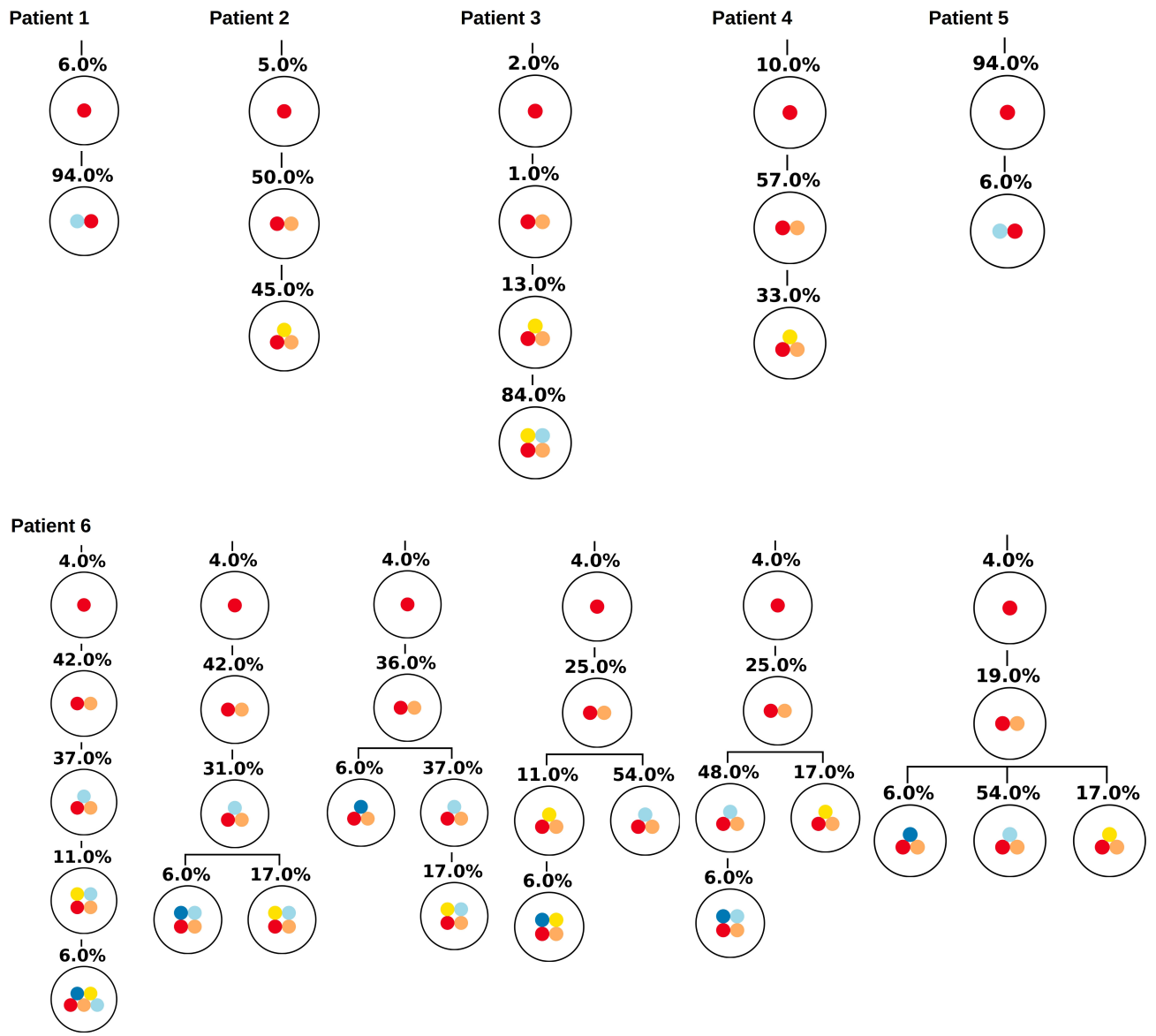

FIGURE 2 | Clonal hierarchies compatible with the bulk allele frequencies measured at diagnosis. For each considered patient all clonal hierarchies are depicted that are compatible with the bulk variant allele frequencies measured at diagnosis. The root of the tree corresponds to the founder mutation that is present in all leukemic cells. The percentages indicate the frequencies of the respective clones that have to be mixed to obtain the measured bulk VAF. We observe that in most cases the hierarchies are unique.

each randomly generated dataset, we check how many other tree structures reproduce the data without an error. The results are shown in the Supplement (**Supplementary Figure 1**). They suggest that linear hierarchies or hierarchies where branches appear only at nodes with a high depth exhibit uniqueness in many cases. The structures with branches near the root often admit multiple reconstructions. However, if data at two time-points are available, e.g., at diagnosis and relapse, in 50–70% of the cases only one or two configurations exactly fit to the bulk data.

Sampling Error Has Little Impact on the Uniqueness of Clonal Hierarchies

If the frequency of different clones in a large population is estimated based on a small sample, sampling errors can occur. To study the impact of sampling errors on the reconstructed clonal hierarchies we again use the single cell sequencing data from Ediriwickrema et al. (2020). We assume that the single cell data reflect the true frequencies of the clones in the malignant cell bulk of the respective patient. For an arbitrary patient k we

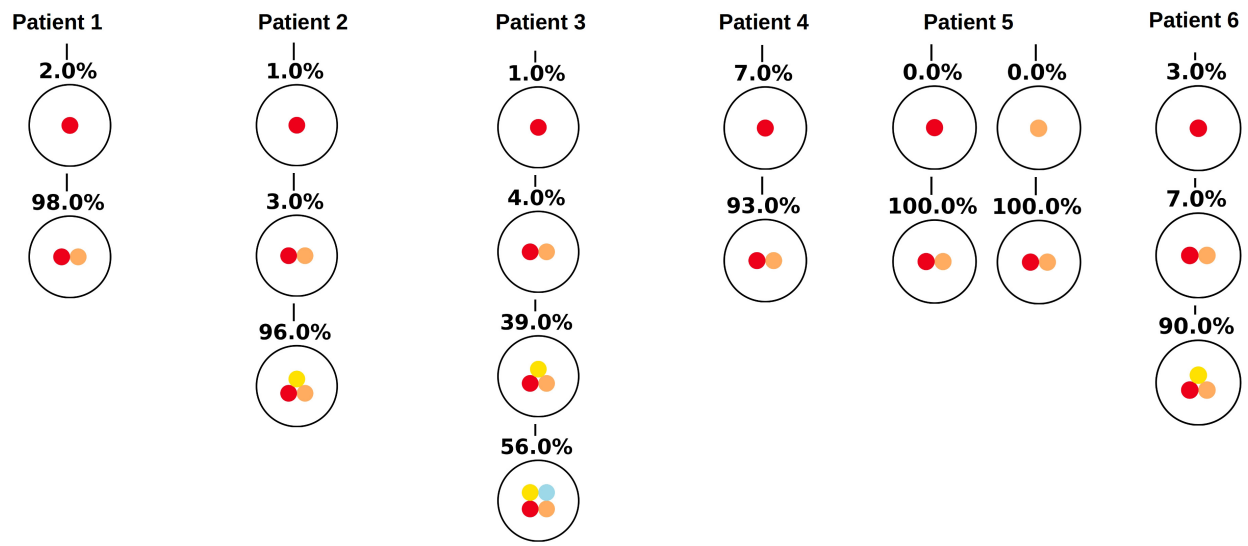

FIGURE 3 | Clonal hierarchies compatible with the bulk allele frequencies measured at relapse. For each considered patient, the figure shows which clonal hierarchies are compatible with the bulk variant allele frequencies measured at relapse. The root of the tree corresponds to the founder mutation. We observe that in most cases the hierarchies are unique.

know the total number n_k of sequenced leukemic single cells. Furthermore we know the frequencies $f_{i,k}$ of each clone that has been detected (here $f_{i,k}$ denotes the frequency of clone i in the sample of patient k). To study the impact of sampling on the bulk variant allele frequencies and on the reconstructed hierarchies for patient k, we draw 1,000 random samples of size n_k from a multinomial distribution with probabilities $p_i = f_{i,k}$. This approach is referred to as resampling (Gigli, 1996). For each of these 1,000 random samples we calculate the bulk variant allele frequencies and apply our algorithm to reconstruct the clonal hierarchies. The results are shown in **Figure 5**. In all cases except one the hierarchies fitting exactly to the data remain unique and are identical to the hierarchies obtained based on the exact data. For one patient in some of the resampled datasets the number of hierarchies matching the data increases by one. These results imply that the sampling error has a negligible impact on the clonal pedigrees that fit to the data. The sampling error also affects the clonal frequencies x_i obtained from the reconstruction. The reconstruction is based on linear equations. Therefore, if many samples are drawn from the same patient, the mean over the reconstructed frequencies approximates the true frequencies of the respective clones. We have assessed the standard deviations of the reconstructed clonal frequencies numerically based on 1,000 re-samplings. In all cases they were less than 4%, in patients 1 to 5 they were less than 1.5%.

Impact of Measurement Errors on Reconstruction of Clonal Hierarchies

Inaccuracies in sequencing are another possible source of error. To study their impact on the reconstructed clonal hierarchies, we add a normally distributed error to the bulk frequency of each allele. Such errors can have different impacts on the reconstructed clonal hierarchies. For each patient we considered 1,000 randomly perturbed versions of the original data. If the standard deviation of the error distribution is 0.5% (i.e., in 68% of cases the error is less or equal 0.5%, in 95% of cases the error is less or equal to 1%) the reconstruction algorithm works reliably in the sense that the true configuration is an optimal configuration, see **Figure 6**. In 5 out of 6 considered patients the optimum is unique. We repeated the simulation for a normally distributed error with a standard deviation of 5%, i.e., in 95% of cases the error is less than 10%, see **Figure 6**. For an error of this magnitude the true configuration not always remains an optimal configuration. Examples illustrating this observation are provided in the Supplement (**Supplementary Figure 2**). This especially applies to patients in whom the frequency of the founding clone is small (i.e., patients 2, 3 and 6). If the error is larger than the frequency of the founder clone it becomes impossible to reliably detect which hit occurs first. However, also in a single cell sequencing approach, rare clones can remain undetected due to sampling or sequencing errors, implying that the first hit remains unknown. In terms of variant allele frequencies this implies that trees cannot be reliably reconstructed if the difference between the two most abundant allele frequencies is of the order of magnitude of the sequencing error. In patients with many clones, our algorithm can often rule out most of the possible hierarchies and identify a small number of configurations fitting the data. In case of Patient 5 the true configuration is always among the upper 12% of the best fitting configurations (i.e., the best or second best), and in patient 6 among the upper 3.3% of the best fitting configurations (i.e., 4 out of 125). In case of small clone numbers such as for patients 2 or 4, the true configuration is always among the two best fitting hierarchies.

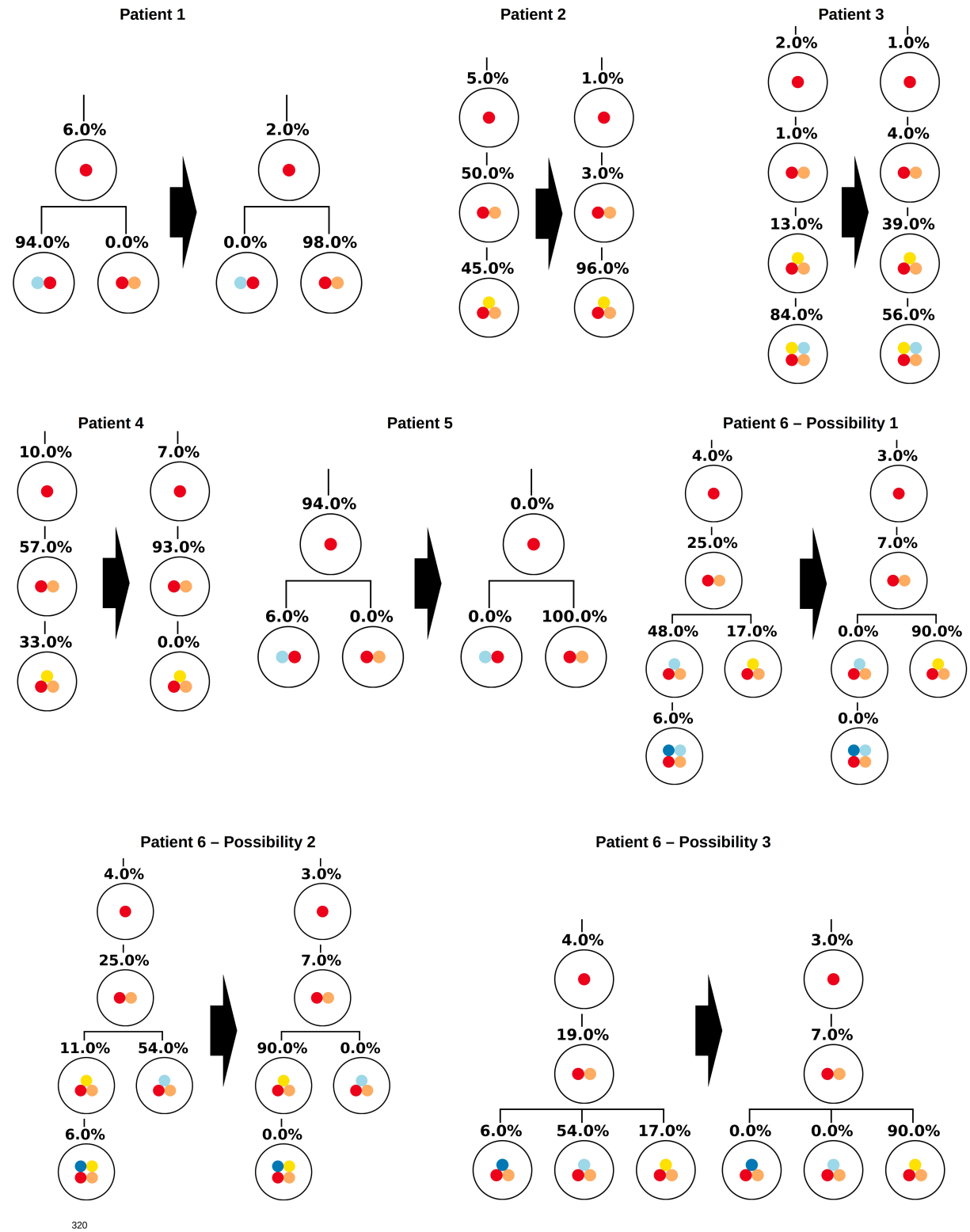

FIGURE 4 | Clonal hierarchies compatible with the bulk allele frequencies measured at diagnosis and relapse. For each patient, the figure shows all clonal hierarchies that are compatible with the bulk VAFs measured at diagnosis and relapse. The root of the tree corresponds to the founder mutation. We observe that in case of patient 6, the number of hierarchies compatible with the data is reduced compared to **Figure 2**. For patients 1–5, the reconstructed hierarchies coincide with the result from single cell sequencing. For patient 6, Possibility 3 corresponds to the true configuration.

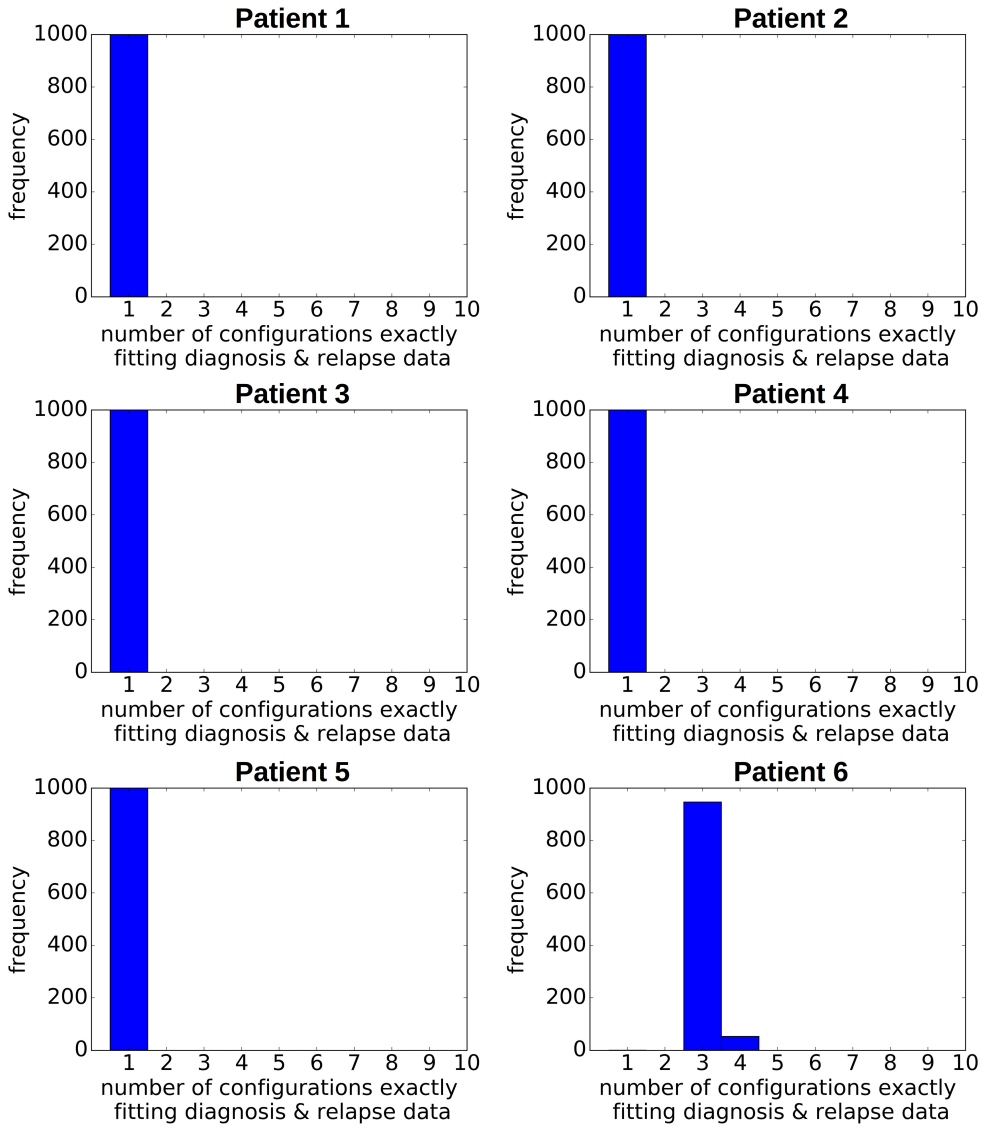

FIGURE 5 | Impact of sampling error on the reconstructed hierarchies. For each patient, we generate 1,000 random pairs of diagnosis and relapse samples from a multivariate distribution. The probabilities of the multivariate distribution equal the clonal frequencies in the single cell data. The size of the samples equals the number of sequenced leukemic cells. We recorded for each pair of randomly generated diagnosis/relapse samples the number of clonal hierarchies compatible with the resampled diagnosis and relapse data. The vertical axis shows how many of the 1,000 samples were compatible with 1, 2, 3, . . . hierarchies, respectively.

An Example of a Patient With Two Founder Clones

We now consider an example of a patient with two different founder clones. This scenario either corresponds to the rare case where the AML cell population originates from clones with different initial mutations or it corresponds to the case where the common founding mutation has not been detected. The latter may especially occur in the setting of targeted sequencing, where only a predefined subset of mutations is considered. Such a scenario occurs if in a purified AML sample (i.e., in a sample without healthy cells) all bulk VAF are significantly different from the expected maximum of 50% (for heterozygous mutations) or 100% (for homozygous mutations).

The proposed algorithm can take this scenario into account by considering healthy cells as the root of the tree. This means a healthy reference allele that is present in all cells is added to the list of variant allele frequencies, to obtain a single tree with a unique root. **Figure 7** shows all tree structures that are compatible with the measured data. The tree structures can be divided into two classes. In the first class of solutions, the frequency of healthy cells is zero at diagnosis and relapse (possibilities 1–2), in the second class the frequency of the healthy cells is positive (here 15%) at least one time point (possibilities 3–7). Solutions of the first class imply that there exist two founding clones (or an undetected unique founder mutation), solutions of the second class may imply that the sample contains a mixture of healthy and

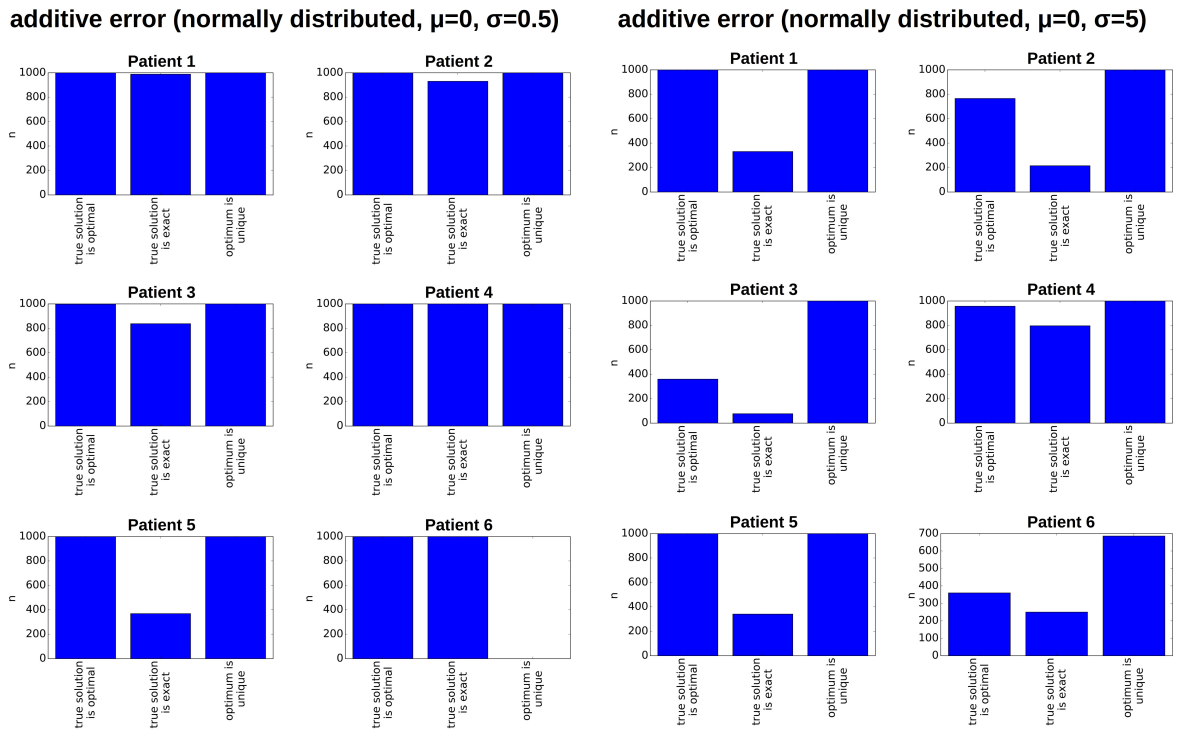

FIGURE 6 | Impact of measurement errors on the reconstructed hierarchies. For each patient, we generate 1,000 randomly perturbed pairs of diagnosis and relapse samples. The additive random perturbations were drawn from a normal distribution with mean zero and standard deviation 0.5% (left) or 5% (right), respectively. Perturbations leading to a VAF of less than zero or more than 100% were excluded. For each of the perturbed diagnosis/relapse pairs, we reconstructed all compatible clonal hierarchies. The figure indicates for how many of the 1,000 perturbed samples the true hierarchy optimally fits the perturbed data (compared to all other existing hierarchies), whether the true hierarchy can exactly reproduce the perturbed data and whether the optimal configuration is unique.

leukemic cells. If we can be sure that the experimental procedures prevent healthy cells from being sequenced (e.g., by FACS sorting for a leukemia specific surface marker before sequencing), only two possible tree structures remain.

As for the other patients the sampling error only leads to small changes in the numbers of clonal hierarchies that fit the data, **Figure 8A**. However, already small errors added to the bulk VAFs (normally distributed with a standard deviation of 0.5%) imply that in a majority of cases the true solution is no longer optimal, **Figures 8B,C**. The reason for this observation is as follows (see **Figure 9**). In the exact scenario there exist two founder mutations. The frequencies of both founder mutations add up to 100%. In presence of errors, it can happen that the frequencies of both founder mutations do not add up to exactly 100%. If their sum is slightly less than 100% the true hierarchy still leads to an exact solution (to compensate for the error the exact solution contains a small number of healthy cells). If due to the random error the sum over both sub-trees is slightly more than 100%, an exact solution is no longer possible. To circumvent this, we can relax the dataset by artificially adding a small number of healthy cells, e.g., x% to the dataset. In this case, for measurements where the frequencies of both founding clones add up to less than 100% + x% the true configuration still is an exact solution. We see in **Figure 10** that this relaxation increases the number of cases where the true solution is an optimal solution.

DISCUSSION

The aim of this study is to investigate the ambiguity of clonal hierarchies that are reconstructed from bulk sequencing data. For this purpose, we develop an algorithm that systematically tests which subset of all clonal hierarchies optimally fits a given dataset. We test this algorithm using bulk VAFs that have been calculated based on cell sequencing data sets. Since single cell sequencing reveals the true clonal hierarchy, this approach enables us to compare the output of our algorithm to the real configuration (Kuipers et al., 2017; Brierley and Mead, 2020).

First, we assume that the input data is exact, i.e., neither sampling nor measurement errors occur. Then for most of the considered patient samples exactly one clonal hierarchy optimally fits the bulk VAF. This clonal hierarchy is identical to the hierarchy obtained from single cell sequencing. In two of the considered patients, even for exact input data more than one clonal hierarchy is compatible with the bulk allele frequencies. The true hierarchy obtained from single cell sequencing is among them. This finding implies that even in absence of measurement error, the clonal hierarchy may not be uniquely defined by the bulk VAF.

When drawing multiple samples from the same malignant cell population the variant allele frequencies may differ from one sample to another. This may be caused by sampling error, or it

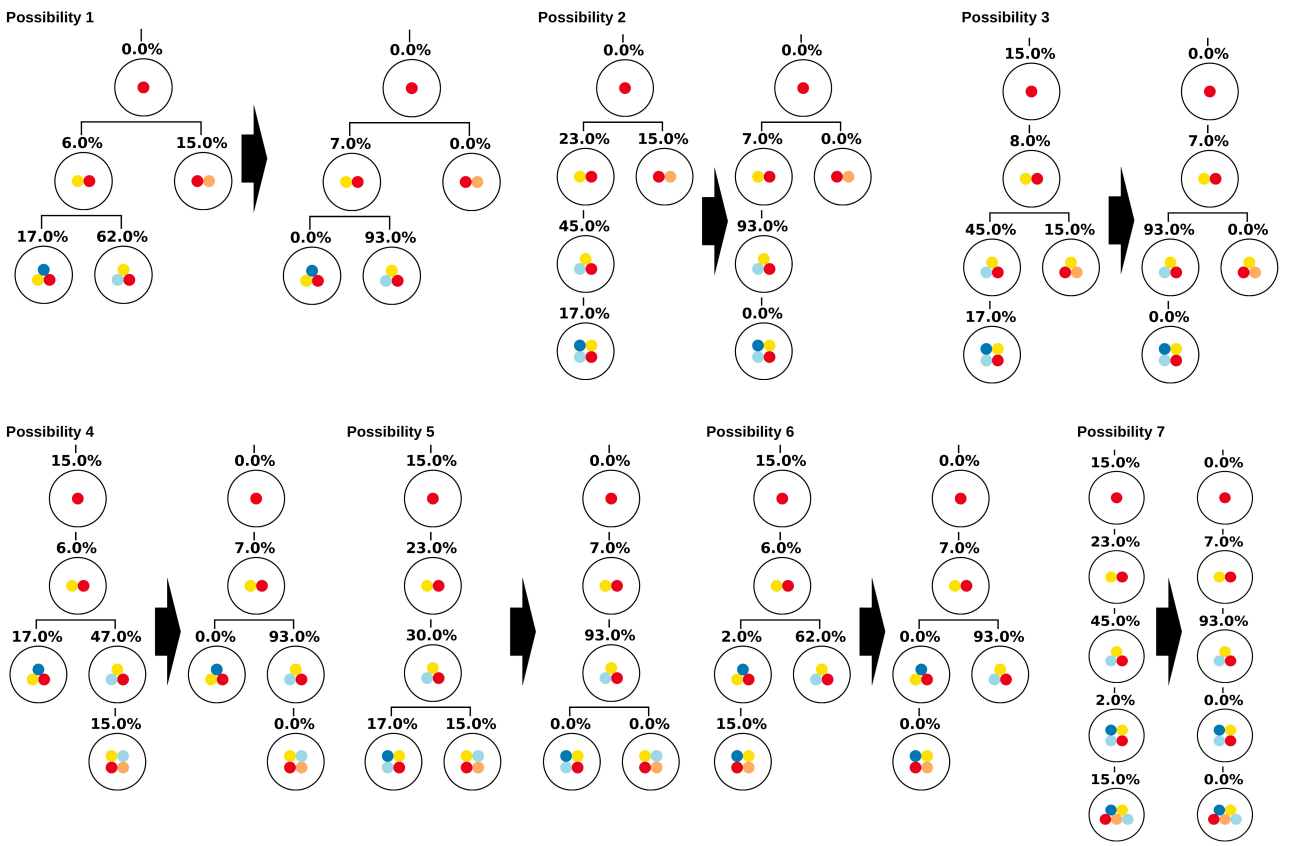

FIGURE 7 | Example of a patient with multiple founder clones. In this figure the root corresponds to wild type cells. The two founding events are indicated by yellow and orange circles. The Figure shows all hierarchies that are compatible with VAFs at diagnosis and relapse. Possibility 1 coincides with the hierarchy obtained from single cell sequencing.

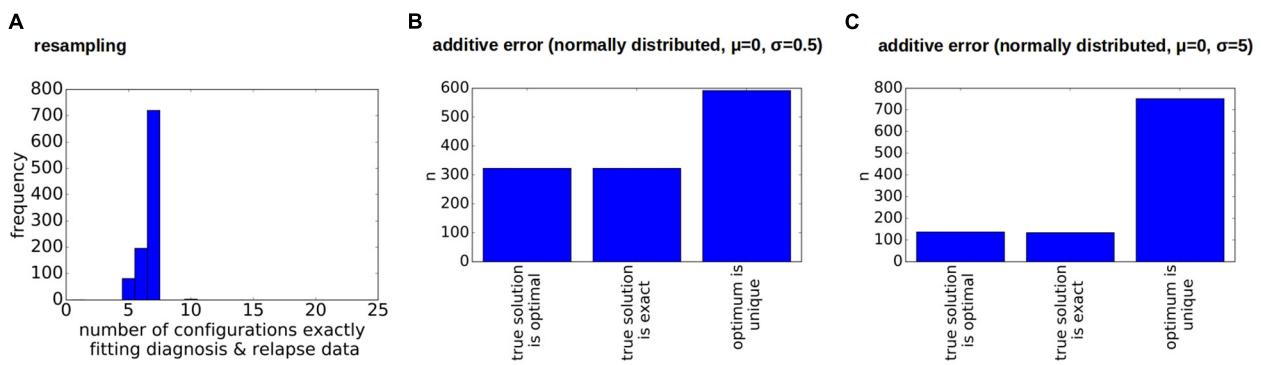

FIGURE 8 | Impact of errors on the reconstructions for a sample with two founders. **(A)** Impact of sampling error on the number of clonal hierarchies compatible with diagnosis and relapse data. The vertical axis shows how many of the 1,000 multinomial respamplings were compatible with 1, 2, 3, . . . hierarchies respectively. **(B,C)** Impact of measurement errors on the reconstructed hierarchies. We considered 1,000 perturbed versions of the original data. Additive perturbations of the VAF were drawn from a normal distribution with mean zero and standard deviation 0.5% **(B)** or 5% **(C)**. We observe that in the majority of cases the true configuration is not optimal.

may reflect inhomogeneity of the tumor. Assuming the tumor to be homogeneous, we aim to quantify the impact of sampling error on the reconstructed hierarchies. For each patient, the number of sequenced leukemic single cells n and the frequencies f_i of the different clones are known. To simulate the sampling error, for each patient we generate 1,000 random samples of size n

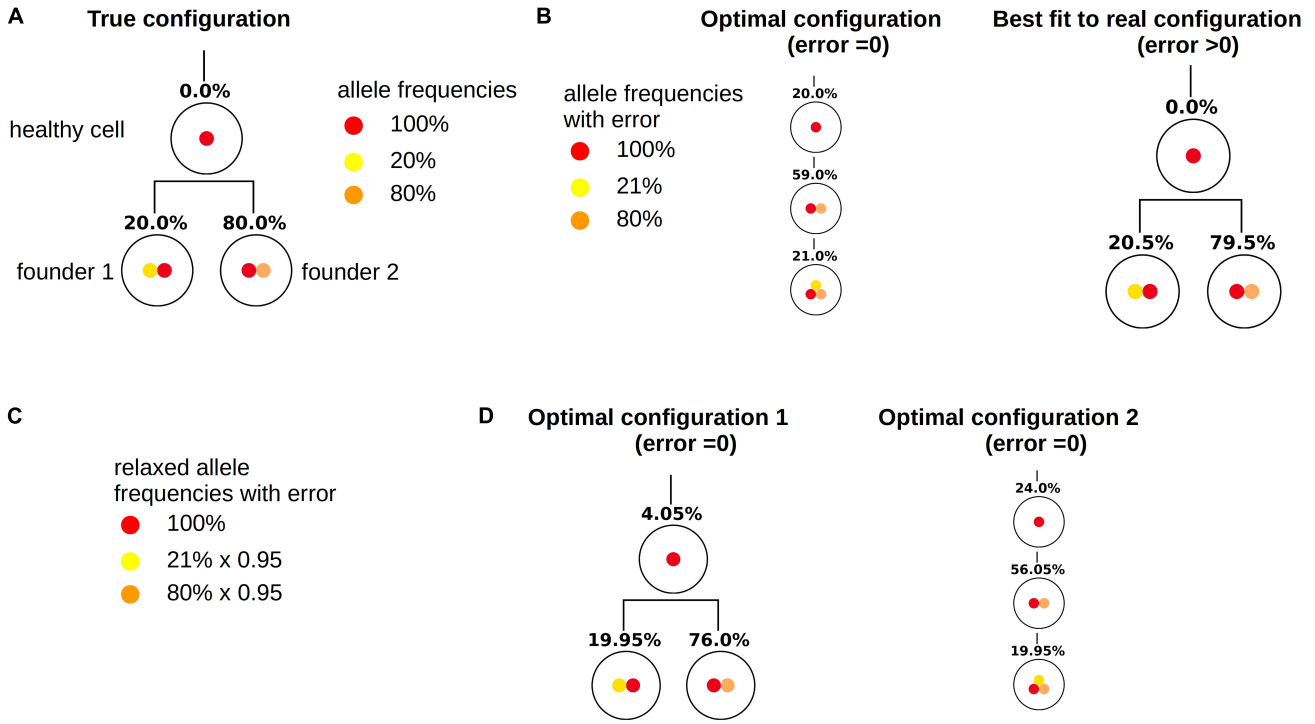

FIGURE 9 | Relaxation of measurements. This figure demonstrates how relaxation of the VAF can help to make the true configuration an optimal configuration. **(A)** True configuration with two founder clones. Wild type cells are added to the tree to have a unique root. Since the sample only contains leukemic cells the frequency of wild type cells equals zero. The frequencies of the two founder clones add up to 100%. **(B)** Variant allele frequencies are perturbed by a measurement error. The frequencies of the two founding mutations no longer add up to 100%. Therefore the true hierarchy is no longer an optimal hierarchy. **(C)** By multiplying the frequencies of mutated alleles with 0.95 we artificially add 5% of healthy cells to the sample. **(D)** If we reconstruct the clonal hierarchies for the modified data the true hierarchy is among the optimal hierarchies.

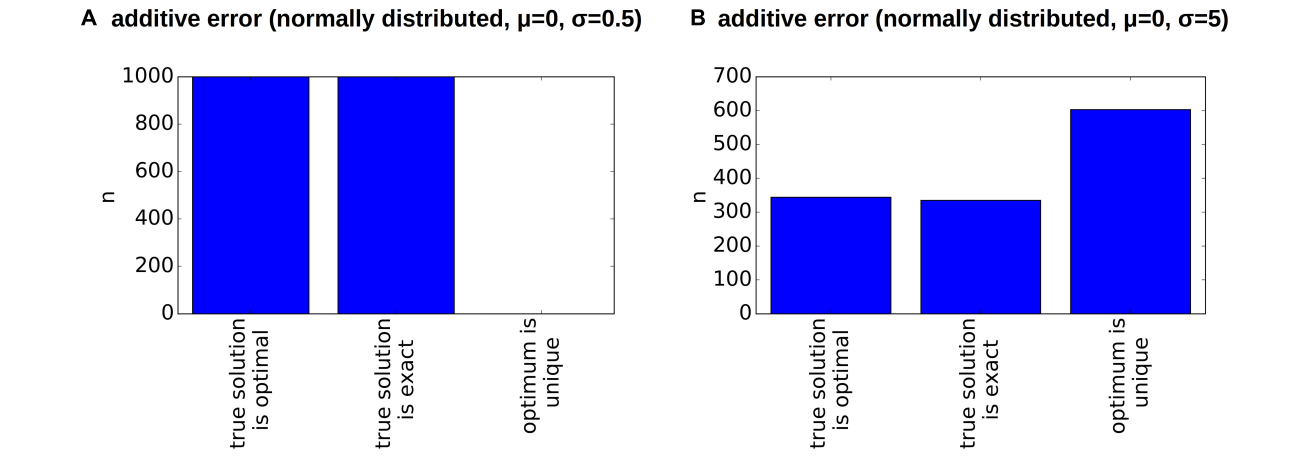

FIGURE 10 | Clonal hierarchies fitting the relaxed dataset. Here, we consider the relaxed version of the dataset from Patient 7. The relaxation makes the reconstruction more robust to errors. Additive perturbations of the VAF were drawn from a normal distribution with mean zero and standard deviation 0.5% **(A)** or 5% **(B)**. In comparison to **Figures 8B,C** the true hierarchy is in many cases an optimal hierarchy.

drawn from a multinomial distribution with probabilities $p_i = f_i$. For each of these random samples we calculate the bulk allele frequencies and construct all clonal hierarchies compatible with them. Based on results of this exercise we conclude that the sampling error has a negligible impact on the obtained clonal hierarchies, at least for the data at our disposal.

We test the robustness of the reconstruction by adding normally distributed errors of different amplitude to the bulk VAFs calculated from the single cell sequencing data. This takes into account potential misreads during the sequencing, amplification errors or impurities of the sample. We observe that for errors of about 5–10% the true hierarchy not necessarily

remains optimal. This especially applies to data sets where the frequency of the founding clone is of the order of magnitude of the error. However, even in this case, the true clonal structure is among the upper 3–15% of the best fitting hierarchies. This implies that also in the presence of relevant errors, our algorithm allows to rule out most tree configurations and results in a small subset of possible clonal hierarchies fitting to a data sample.

Mathematical models indicate that tree characteristics, e.g., the depth of the tree, correlate with clonal properties such as self-renewal and proliferation rate (Stiehl et al., 2016). In this context it can be sufficient to have an estimate of the depth of the true clonal hierarchy to draw conclusions about the effect of a mutation on cell kinetics or patient prognosis. This implies that in the case of non-unique clonal hierarchies, biological conclusions can be drawn if the potential hierarchies are sufficiently similar to each other.

Having measurements of bulk VAFs provided, our computational approach can be used to rank all possible clonal hierarchies based on their compatibility with the data (i.e., the smaller the error when fitting the dataset to a given hierarchy, the better the rank of the respective hierarchy). For all datasets considered in this study the real hierarchy is among the upper 3–15% of this ranking. Taking into account that in case of n clones n^{n-1} possible hierarchies exist our algorithm allows to rule out a significant number of them. Our algorithm can also be applied to scenarios in which the disease is derived from multiple founding clones. However, due to its combinatorial nature the algorithm can only be applied to relatively small clone numbers.

Our computational approach can be used to study how sensitive the reconstructed hierarchies are to perturbations of the input data. By adding random errors to the input data obtained from an experiment and by repeating the reconstruction with the perturbed input data it turns out that some datasets are robust with respect to the perturbations. This means that the obtained optimal clonal hierarchies do not change if the input data is perturbed. For other datasets perturbations of the input data leads to a change of the reconstructed hierarchies, indicating that the reconstruction may be affected by measurement errors. The robustness of a given dataset can be checked using our proposed framework. It is straightforward to take into account that the measured frequencies may have different confidence intervals. In principle our approach can also be applied to clustered single nucleotide variants (SNVs). Since the number of detected SNVs is usually high, the variants are grouped into clusters according to their allele frequencies. Each cluster comprises all SNVs with a similar allele frequency. The cluster center is defined as the average allele frequency of all SNVs that belong to the respective cluster. In this setting our algorithm can be applied using cluster centers as input data.

Mechanistic mathematical models allow to extract relevant information from clonal hierarchies, such as estimates of proliferation rates and self-renewal of the different clones (Whichard et al., 2010; Stiehl et al., 2016). Correlating these estimates with detected mutations and clinical observations may provide new insights into AML pathophysiology (Stiehl et al., 2014, 2016). The proposed framework is a first attempt to quantify the ambiguity emerging during reconstruction of clonal hierarchies from bulk sequencing data. It allows to identify when such reconstructions are reliable and can be used as input data for mechanistic models. This knowledge helps to make available routine clinical data to studies that require clonally resolved input (Leung et al., 2017).

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Ediriwickrema et al. (2020).

AUTHOR CONTRIBUTIONS

TS and AM-C designed the research, discussed the results, and wrote the manuscript. TS implemented the algorithm and run simulations. Both authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by research funding from the German Research Foundation DFG (SFB 873; subproject B08).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys.2021.596194/full#supplementary-material

REFERENCES

Abdel-Wahab, O., Manshouri, T., Patel, J., Harris, K., Yao, J., Hedvat, C., et al. (2010). Genetic analysis of transforming events that convert chronic myeloproliferative neoplasms to leukemias. *Cancer Res.* 70, 447–452. doi: 10.1158/0008-5472.can-09-3783

Anderson, K., Lutz, C., van Delft, F. W., Bateman, C. M., Guo, Y., Colman, S. M., et al. (2011). Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature.* 469, 356–361. doi: 10.1038/nature09650

Andersen, M., Dahl, J., and Vandenberghe, L. (2018). *CVXOPT: A Python Package for Convex Optimization, Version 1.2.0*. Available online at: http://cvxopt.org/

Attolini, C. S., Cheng, Y. K., Beroukhim, R., Getz, G., Abdel-Wahab, O., Levine, R. L., et al. (2010). A mathematical framework to determine the temporal sequence of somatic genetic events in cancer. *Proc. Natl. Acad. Sci. U S A.* 107, 17604–17609. doi: 10.1073/pnas.1009117107

Bachas, C., Schuurhuis, G. J., Assaraf, Y. G., Kwidama, Z. J., Kelder, A., Wouters, F., et al. (2012). The role of minor subpopulations within the leukemic blast compartment of AML patients at initial diagnosis in the development of relapse. *Leukemia* 26, 1313–1320. doi: 10.1038/leu.2011.383

Bacher, U., Haferlach, C., Kern, W., Haferlach, T., and Schnittger, S. (2008). Prognostic relevance of FLT3-TKD mutations in AML: the combination matters–an analysis of 3082 patients. *Blood* 111, 2527–2537. doi: 10.1182/blood-2007-05-091215

Banck, J. C., and Görlich, D. (2019). In-silico comparison of two induction regimens (7 + 3 vs 7 + 3 plus additional bone marrow evaluation) in acute myeloid leukemia treatment. *BMC Syst. Biol.* 13:18. doi: 10.1186/s12918-019-0684-0

Brierley, C. K., and Mead, A. J. (2020). Single-cell sequencing in hematology. *Curr. Opin. Oncol.* 32, 139–145. doi: 10.1097/cco.0000000000000613

Busse, J. E., Gwiazda, P., and Marciniak-Czochra, A. (2016). Mass concentration in a nonlocal model of clonal selection. *J. Math. Biol.* 73, 1001–1033. doi: 10.1007/s00285-016-0979-3

Cancer Genome Atlas Research Network. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* 368, 2059–2074. doi: 10.1056/nejmoa1301689

Caravagna, G., Sanguinetti, G., Graham, T. A., and Sottoriva, A. (2020). The MOBSTER R package for tumour subclonal deconvolution from bulk DNA whole-genome sequencing data. *BMC Bioinform.* 21:531. doi: 10.1186/s12859-020-03863-1

Cocciardi, S., Dolnik, A., Kapp-Schwoerer, S., Rücker, F. G., Lux, S., Blätte, T. J., et al. (2019). Clonal evolution patterns in acute myeloid leukemia with NPM1 mutation. *Nat. Commun.* 10:2031.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., et al. (2009). Mutation in TET2 in myeloid cancers. *N. Engl. J. Med.* 360, 2289–2301.

Diestel, R. (2017). *Graph Theory*. Cham: Springer.

Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., et al. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481, 506–510.

Dinh, K., Corey, S. J., and Kimmel, M. (2019). Predicting minimal residual disease in acute myeloid leukemia through stochastic modeling of clonality. *Blood* 134:1448. doi: 10.1182/blood-2019-127457

Dinh, K. N., Corey, S. J., and Kimmel, M. (2020). Application of the moran model in estimating selection coefficient of mutated CSF3R clones in the evolution of severe congenital neutropenia to myeloid neoplasia. *Front. Physiol.* 11:806. doi: 10.3389/fphys.2020.00806

Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Büchner, T., et al. (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood* 129, 424–447. doi: 10.1182/blood-2016-08-733196

Ediriwickrema, A., Aleshin, A., Reiter, J. G., Corces, M. R., Köhnke, T., Stafford, M., et al. (2020). Single-cell mutational profiling enhances the clinical evaluation of AML MRD. *Blood Adv.* 4, 943–952. doi: 10.1182/bloodadvances.2019001181

Gigli, A. (1996). Efficient bootstrap methods: A review. *J. Ital. Statist. Soc.* 1, 99–127. doi: 10.1007/bf02589584

Greif, P. A., Hartmann, L., Vosberg, S., Stief, S. M., Mattes, R., Hellmann, I., et al. (2018). Evolution of cytogenetically normal acute myeloid leukemia during therapy and relapse: an exome sequencing study of 50 patients. *Clin. Cancer Res.* 24, 1716–1726. doi: 10.1158/1078-0432.ccr-17-2344

Grimwade, D., Walker, H., Oliver, F., Wheatley, K., Harrison, C., Harrison, G., et al. (1998). The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The medical research council adult and children's leukaemia working parties. *Blood* 92, 2322–2333. doi: 10.1182/blood.v92.7.2322

Herudkova, Z., Culen, M., Folta, A., Jeziskova, I., Cerna, J., Loja, T., et al. (2020). Clonal hierarchy of main molecular lesions in acute myeloid leukaemia. *Br. J. Haematol.* 190, 562–572. doi: 10.1111/bjh.16341

Jonas, B. A. (2017). From MDS/AML to iPSC and back again. *Sci. Transl. Med.* 9:eaam9861. doi: 10.1126/scitranslmed.aam9861

Kuipers, J., Jahn, K., and Beerenwinkel, N. (2017). Advances in understanding tumour evolution through single-cell sequencing. *Biochim. Biophys. Acta Rev. Cancer* 1867, 127–138. doi: 10.1016/j.bbcan.2017.02.001

Leung, M. L., Davis, A., Gao, R., Casasent, A., Wang, Y., Sei, E., et al. (2017). Single-cell DNA sequencing reveals a late-dissemination model in metastatic colorectal cancer. *Genome Res.* 27, 1287–1299. doi: 10.1101/gr.209973.116

Lorenzi, T., Marciniak-Czochra, A., and Stiehl, T. (2019). A structured population model of clonal selection in acute leukemias with multiple maturation stages. *J. Math Biol.* 79, 1587–1621. doi: 10.1007/s00285-019-01404-w

Lutz, C., Hoang, V. T., Buss, E., and Ho, A. D. (2013). Identifying leukemia stem cells–is it feasible and does it matter? *Cancer Lett.* 338, 10–14. doi: 10.1016/j.canlet.2012.07.014

Nobile, M. S., Vlachou, T., Spolaor, S., Bossi, D., Cazzaniga, P., Lanfrancone, L., et al. (2019). Modeling cell proliferation in human acute myeloid leukemia xenografts. *Bioinformatics* 35, 3378–3386. doi: 10.1093/bioinformatics/btz063

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., et al. (2016). Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* 374, 2209–2221.

Prüfer, H. (1918). Neuer Beweis eines Satzes über Permutationen. *Arch. Math. Phys.* 27, 742–744.

Rahman, M. S., Nicholson, A. E., and Haffari, G. (2018). HetFHMM: A novel approach to infer tumor heterogeneity using factorial hidden markov models. *J. Comput. Biol.* 25, 182–193. doi: 10.1089/cmb.2017.0101

Ran, D., Schubert, M., Taubert, I., Eckstein, V., Bellos, F., Jauch, A., et al. (2012). Heterogeneity of leukemia stem cell candidates at diagnosis of acute myeloid leukemia and their clinical significance. *Exp. Hematol.* 40, 155–165. doi: 10.1016/j.exphem.2011.10.005

Röllig, C., Bornhäuser, M., Thiede, C., Taube, F., Kramer, M., Mohr, B., et al. (2011). Long-term prognosis of acute myeloid leukemia according to the new genetic risk classification of the European LeukemiaNet recommendations: evaluation of the proposed reporting system. *J. Clin. Oncol.* 29, 2758–2765. doi: 10.1200/jco.2010.32.8500

Roloff, G. W., and Griffiths, E. A. (2018). When to obtain genomic data in Acute Myeloid Leukemia (AML) and which mutations matter. *Blood Adv.* 2, 3070–3080. doi: 10.1182/bloodadvances.2018020206

Roth, A., Khattra, J., Yap, D., Wan, A., Laks, E., Biele, J., et al. (2014). PyClone: statistical inference of clonal population structure in cancer. *Nat. Methods* 11, 396–398. doi: 10.1038/nmeth.2883

Salichos, L., Meyerson, W., Warrell, J., and Gerstein, M. (2020). Estimating growth patterns and driver effects in tumor evolution from individual samples. *Nat. Commun.* 11:732.

Sandén, C., Lilljebjörn, H., Orsmark Pietras, C., Henningsson, R., Saba, K. H., Landberg, N., et al. (2020). Clonal competition within complex evolutionary hierarchies shapes AML over time. *Nat. Commun.* 11:579.

Stiehl, T., Baran, N., Ho, A. D., and Marciniak-Czochra, A. (2014). Clonal selection and therapy resistance in acute leukaemias: mathematical modelling explains different proliferation patterns at diagnosis and relapse. *J. R. Soc. Interface* 11:20140079. doi: 10.1098/rsif.2014.0079

Stiehl, T., Baran, N., Ho, A. D., and Marciniak-Czochra, A. (2015). Cell division patterns in acute myeloid leukemia stem-like cells determine clinical course: a model to predict patient survival. *Cancer Res.* 75, 940–949. doi: 10.1158/0008-5472.can-14-2508

Stiehl, T., Ho, A. D., and Marciniak-Czochra, A. (2018). Mathematical modeling of the impact of cytokine response of acute myeloid leukemia cells on patient prognosis. *Sci. Rep.* 8:2809.

Stiehl, T., Lutz, C., and Marciniak-Czochra, A. (2016). Emergence of heterogeneity in acute leukemias. *Biol. Direct.* 11:51.

Stiehl, T., and Marciniak-Czochra, A. (2017). Stem cell self-renewal in regeneration and cancer: Insights from mathematical modeling. *Curr. Opin. Syst. Biol.* 5, 112–120. doi: 10.1016/j.coisb.2017.09.006

Stiehl, T., Wang, W., Lutz, C., and Marciniak-Czochra, A. (2020). Mathematical modeling provides evidence for niche competition in human AML and serves as a tool to improve risk stratification. *Cancer Res.* 80, 3983–3992. doi: 10.1158/0008-5472.can-20-0283

Wang, W., Stiehl, T., Raffel, S., Hoang, V. T., Hoffmann, I., Poisa-Beiro, L., et al. (2017). Reduced hematopoietic stem cell frequency predicts outcome in acute myeloid leukemia. *Haematologica* 102, 1567–1577. doi: 10.3324/haematol.2016.163584

Whichard, Z. L., Sarkar, C. A., Kimmel, M., and Corey, S. J. (2010). Hematopoiesis and its disorders: a systems biology approach. *Blood* 115, 2339–2347. doi: 10.1182/blood-2009-08-215798

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.