The ecology in the hematopoietic stem cell niche determines the clinical outcome in chronic myeloid leukemia

Adam L. MacLean[1], Sarah Filippi[1], and Michael P. H. Stumpf[2]

Theoretical Systems Biology, Division of Molecular Biosciences, Imperial College London, London SW7 2AZ, United Kingdom

Edited* by Robert M. May, University of Oxford, Oxford, United Kingdom, and approved January 10, 2014 (received for review September 10, 2013)

Chronic myeloid leukemia (CML) is a blood disease that disrupts normal function of the hematopoietic system. Despite the great progress made in terms of molecular therapies for CML, there remain large gaps in our understanding. By comparing mathematical models that describe CML progression and etiology we sought to identify those models that provide the best description of disease dynamics and their underlying mechanisms. Data for two clinical outcomes—disease remission or relapse—are considered, and we investigate these using Bayesian inference techniques throughout. We find that it is not possible to choose between the models based on fits to the data alone; however, by studying model predictions we can discard models that fail to take niche effects into account. More detailed analysis of the remaining models reveals mechanistic differences: for one model, leukemia stem cell dynamics determine the disease outcome; and for the other model disease progression is determined at the stage of progenitor cells, in particular by differences in progenitor death rates. This analysis also reveals distinct transient dynamics that will be experimentally accessible, but are currently at the limits of what is possible to measure. To resolve these differences we need to be able to probe the hematopoietic stem cell niche directly. Our analysis highlights the importance of further mapping of the bone marrow hematopoietic niche microenvironment as the "ecological" interactions between cells in this niche appear to be intricately linked to disease outcome.

niche dynamics | competition | cancer progression | model selection

Hematopoiesis is the intricate process that provides the vast diversity of cells found in the blood and immune system of vertebrate organisms. To generate great numbers of very different cells—ranging from erythrocytes to highly specialized lymphocytes—a delicate balance of activity of transcriptional programs and hematopoietic growth factors is required (1). Within this system, hematopoietic stem cells (HSCs) give rise to (lymphoid and myeloid) progenitors, which in turn through further differentiation steps yield the panoply of cells found in the blood (2). Given the complexity of the hematopoietic system it is remarkable how robustly this process works in general.

The cellular and molecular processes involved in healthy hematopoiesis must be disrupted in disease (3). There is however solid evidence that disease processes operate in parallel to the normal hematopoietic system. Thus, a branch of the hematopoietic system will propagate deficient cells and generate the ensuing pathology, but other cells of the system will continue to operate alongside (4–7). This is, of course, akin to problems encountered in ecology. Based on this analogy we here attempt to understand the interplay between the healthy and chronic myeloid leukemia (CML) systems. This is a particularly helpful perspective if interactions between healthy and diseased cell types can affect the disease progression, as is now widely believed.

Interactions will include shared reliance on nutrients, cofactors, and physiological conditions required to maintain stem cell properties. These factors can be subsumed under the umbrella of the stem cell niche, first introduced in ref. 8. In adult mammals the bone marrow is the site of hematopoiesis and contains the HSC niche. Although much has been learnt about the HSC niche in the last decade, which particular elements constitute the niche continues to be debated (9–14). The dynamics in the niche impact the abundance and maintenance of stem cells, so in the case of cancer arising in the niche, stem cell dynamics will certainly be affected. The ensuing interactions between healthy and cancer stem cells can be considered as competitive interactions, given that they must share the influence of niche constituents (15).

Here we investigate models of hematopoiesis that describe both healthy and leukemic-related hematopoietic systems derived from leukemia stem cells (LSCs) in light of data for the changing levels of leukemia in patients over time. These models differ quite profoundly in the manner in which they consider interactions (or competition) between different cell types.

The level of leukemia in patients is quantified by measuring the proportion of differentiated leukemic cells in blood. The patients are receiving tyrosine kinase inhibitor (TKI) therapy (16), and our data fall into two different categories: data for patients that go into remission, and data for patients exhibiting relapse. We consider these cases separately to identify possible functional differences that give rise to—or at least are correlated with—the different clinical outcomes. We use a Bayesian computation framework throughout to calibrate models against data and compare the performance of the different models.

We will show that detailed analysis of the models allows us to reject those models that fail to account for niche dynamics explicitly and from the outset. Furthermore we can show that the available data suggest strongly an important role of the HSC niche dynamics for CML disease progression and outcome. We find evidence for the niche effects extending beyond the HSC/LSC level and also affecting the dynamics of progenitor and quite possible other cell types in the hematopoietic hierarchy.

Models for CML Dynamics

Three models for CML progression are studied. Each describes the dynamics of healthy and leukemic stem cell lineages and accounts for the impact of the niche in different ways.

Significance

Three contrasting models of the ecological interactions in the hematopoietic stem cell niche explain clinical progression of chronic myeloid leukemia equally well, but do so in different ways. We identify key differences between models and find that we can conclusively rule out those that fail to take competition between healthy and leukemic lineages explicitly into account. Detailed analysis of population dynamics within the bone marrow niche allows us to ascribe mechanisms to distinct disease outcomes and suggests experiments to distinguish between these mechanisms.

Author contributions: A.L.M., S.F., and M.P.H.S. designed research; A.L.M. and S.F. performed research; A.L.M. and S.F. analyzed data; and A.L.M., S.F., and M.P.H.S. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]A.L.M. and S.F. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: m.stumpf@imperial.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317072111/-/DCSupplemental.

Model M1—Niche Competition. The first model describes the dynamics of five species of interacting cell populations (17): HSCs (x_0), healthy progenitor (x_1) and differentiated cells (x_2), LSCs (y_0), and differentiated leukemia cells (y_1)(Fig. 1*A*). The dynamics are given by:

$$\begin{aligned}\dot{x}_0 &= a_x x_0(k-z) - b_x x_0 & \dot{y}_0 &= a_y y_0(k-z) - b_y y_0\\ \dot{x}_1 &= b_x x_0 + c_x x_1(k-z) - d_x x_1 & \dot{y}_1 &= b_y y_0 - e_y y_1,\\ \dot{x}_2 &= d_x x_1 - e_x x_2\end{aligned}$$

where the parameters $(a_x, b_x, c_x, d_x, e_x, a_y, b_y, e_y)$ that characterize the phenotypes of the cells are described in Table 1. Here, k represents the so-called carrying capacity, i.e., the total number of cells that can exist within the bone marrow; in this analysis we use $k=1$ so population sizes are given as concentrations. Here z represents the total size of the niche, i.e., $z = x_0 + x_1 + x_2 + y_0 + y_1$, thus occupation of the niche is not limited to the stem cell species, but extends to progenitor and differentiated cells.

Model M2—Niche Independence. This model was developed to describe the dynamics of HSCs and LSCs and their temporal evolution under TKI therapy and contains eight species: HSCs with three progenitor populations, and LSCs and their corresponding three leukemic progenitor populations (18). The model is depicted in Fig. 1*A* and the equations that specify it are:

$$\begin{aligned}\dot{x}_0 &= (\lambda(x_0) - \delta_0)x_0 \equiv 0 & \dot{y}_0 &= (\rho_y(1-u) - \delta_0)y_0\\ \dot{x}_1 &= \alpha_x x_0 - \delta_1 x_1 & \dot{y}_1 &= \alpha_y y_0 - \delta_1 y_1\\ \dot{x}_2 &= \beta_x x_1 - \delta_2 x_2 & \dot{y}_2 &= \beta_y y_1 - \delta_2 y_2\\ \dot{x}_3 &= \gamma_x x_2 - \delta_3 x_3 & \dot{y}_3 &= \gamma_y y_2 - \delta_3 y_3,\end{aligned}$$

where x_0 are HSCs, y_0 are LSCs, and the x_i, y_i for $i \in (1,2,3)$ are progeny of the two stem cell populations, respectively. Also, $\lambda(x_0)$ describes a homeostatic function defined such that $\dot{x}_0 = 0$; the HSC population in this case is constant over time. Here, u is the number of resistance mutations, which is assumed to be 0 (a version of the model that did include resistance mutations was briefly considered, but mechanistically it offers no more than the form presented here, because there are no additional interactions between species). Note that the δ_i, which are referred to as death rates in the succeeding analysis, subsume the processes of differentiation and migration out of the niche; these are shared between healthy and leukemic lineages. The parameters that are shared by model M3 are described in Table 1.

Model M3—Partial Niche Dependence. Model M2 does not allow for competition between HSCs and LSCs, and in ref. 19 a model is introduced that describes the dynamics of the same eight species as in M2, but where the growth of stem cells is affected by the population size of HSCs and LSCs. We have rescaled model M3 compared with ref. 19 to allow for comparison of the three models considered here:

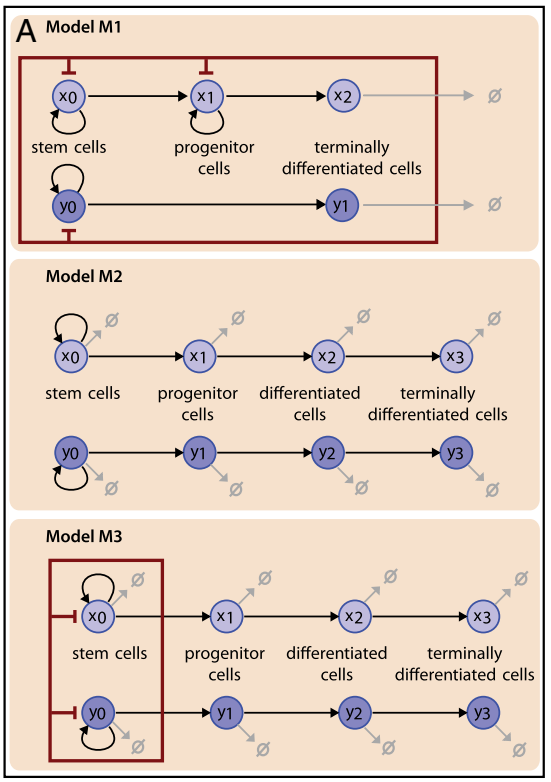
$$\begin{aligned}\dot{x}_0 &= (\Phi_x - 1)\delta_0 x_0 & \dot{y}_0 &= (\Phi_y - 1)\delta_0 y_0\\ \dot{x}_1 &= \alpha_x x_0 - \delta_1 x_1 & \dot{y}_1 &= \alpha_y y_0 - \delta_1 y_1\\ \dot{x}_2 &= \beta_x x_1 - \delta_2 x_2 & \dot{y}_2 &= \beta_y y_1 - \delta_2 y_2\\ \dot{x}_3 &= \gamma_x x_2 - \delta_3 x_3 & \dot{y}_3 &= \gamma_y y_2 - \delta_3 y_3\end{aligned}$$

and

$$\Phi_{x/y} = \frac{(\rho_{x/y} + 1)\kappa}{\kappa + \rho_{x/y}(x_0 + y_0)},$$

where the species $x_i, y_i, i \in (0,1,2,3)$ are the same as for M2. The parameters that specify this model are described in Table 1. We have also considered reduced forms of model M3—containing two or three species per lineage instead of four—to investigate the properties of the model's intermediate species. We perform a similar analysis on these reduced models as the analysis of M3 (*Supporting Information*).

In summary, these three models can be split according to whether they assume complete independence between lineages (M2) or assume that the evolution of one lineage has an impact on the evolution of the other (M1 and M3). In model M1, all species (z) have an influence on the dynamics of healthy and leukemic stem and progenitor cells, thus all species are effectors



C

	Effectors	Niche reliant	Unconstrained by the niche
M1	all cell types	stem and progenitor cells	terminally differentiated cells
M2			all cell types
M3	stem cells	stem cells	progenitor and terminally differentiated cells

Fig. 1. Mechanistic models for hematopoiesis in CML and experimental measurements of leukemia progression in patients. (*A*) Graphical description of models M1, M2, and M3. The hematopoietic niche is represented by a red rectangle: every species inside the rectangle influences the niche, whereas only species marked by red inhibitory arrows are affected by it. Cell death is illustrated by light gray arrows. (*B*) The change in the BCR-ABL level over time (in months) is represented by gray lines for each of the 69 patients under CML treatment. The average change over all patients is illustrated by a dashed green line; the change in the BCR-ABL level for a patient who exhibits disease relapse is illustrated by a dashed red line. Our model comparison analysis is based on these two trajectories. (*C*) Description of how niche competition is treated by each model. Each species can impact the niche (effector), be affected by the niche (niche reliant), and/or be fully unconstrained by the niche.

Table 1. Model parameters

Model	Symbol	Definition
M1	$a_{x/y}$*	Self-renewal rate of stem cells
	$b_{x/y}$	Production rate of progenitor cells
	c_x	Self-renewal rate of healthy progenitor cells
	d_x	Production rate of healthy differentiated cells
	e_x	Death rate of healthy differentiated cells
	e_y	Death rate of leukemic differentiated cells
M3	κ	Carrying capacity of the niche
	$\alpha_{x/y}$*	Production rate of progenitor cells
	$\beta_{x/y}$*	Production rate of differentiated cells
	$\gamma_{x/y}$*	Production rate of fully differentiated cells
	$\rho_{x/y}$*	Growth rate of HSCs/LSCs (rescaled)
	δ_0	Death rate of stem cells†
	δ_1	Death rate of progenitor cells†
	δ_2	Death rate of differentiated cells†
	δ_3	Death rate of fully differentiated cells†

*x denotes healthy lineage; y denotes leukemic lineage.
†Death rates for model M3 are shared between healthy and leukemic lineages.

of the niche, and stem and progenitor cells are affected by the niche (Fig. 1*C*). In model M3, the niche is only composed of stem cells; explicit interactions occur only between these two species. All others (progenitor and differentiated and terminally differentiated cells) are not directly constrained by the niche although their evolution does have niche dependence through the stem cell species that produced them.

In addition to these three models, we have studied one further established model that describes interactions between HSCs and LSCs, including competition, presented in ref. 20. Here similar cellular processes are modeled but using an agent-based modeling approach. The large computational cost of simulating this model, however, precludes the Bayesian analysis attempted here.

Results and Discussion

We study how well each model can explain the CML outcomes of disease remission or disease relapse (Fig. 1*B*). Behavior is classified as remission when the level of CML in the blood is decreasing over time. This is the case for the majority of patient trajectories, therefore we used the mean over 69 patients (averaged at each time point where there were at least 20 measurements taken) to represent remission behavior. Behavior is classified as relapse if an initial decrease in CML is followed by a pronounced and continuing increase in CML levels. Due to variability among patients in the time at which relapse occurs, each relapse patient should be considered independently. We base the subsequent analysis on a single relapse patient trajectory. We performed similar analyses on other relapse patients and found qualitatively similar behavior. We compare how each model explains the data, and from this we are able to draw conclusions about mechanisms of the dynamics of the disease.

Models Cannot be Distinguished Based on Fit to the Data Alone. We use Bayesian parameter inference and model selection algorithms to determine which of the three models best captures the two observed clinical progressions, remission and relapse, as measured by the levels of the BCR-ABL tyrosine kinase—the product of the BCR-ABL gene fusion that is a hallmark of CML—in CML patients being treated with TKIs.

We first use likelihood-free sequential Monte Carlo (SMC) Bayesian inference (21) to perform parameter inference independently for the three models. As shown in Fig. 2*A*, the simulated behaviors obtained for each of the three models and two outcomes are very similar to the observed behaviors. The acceptance threshold between simulated and observed data was chosen so that the final Euclidian distance between the two was lower than 4.0 in the case of remission and 15.0 for relapse. From the

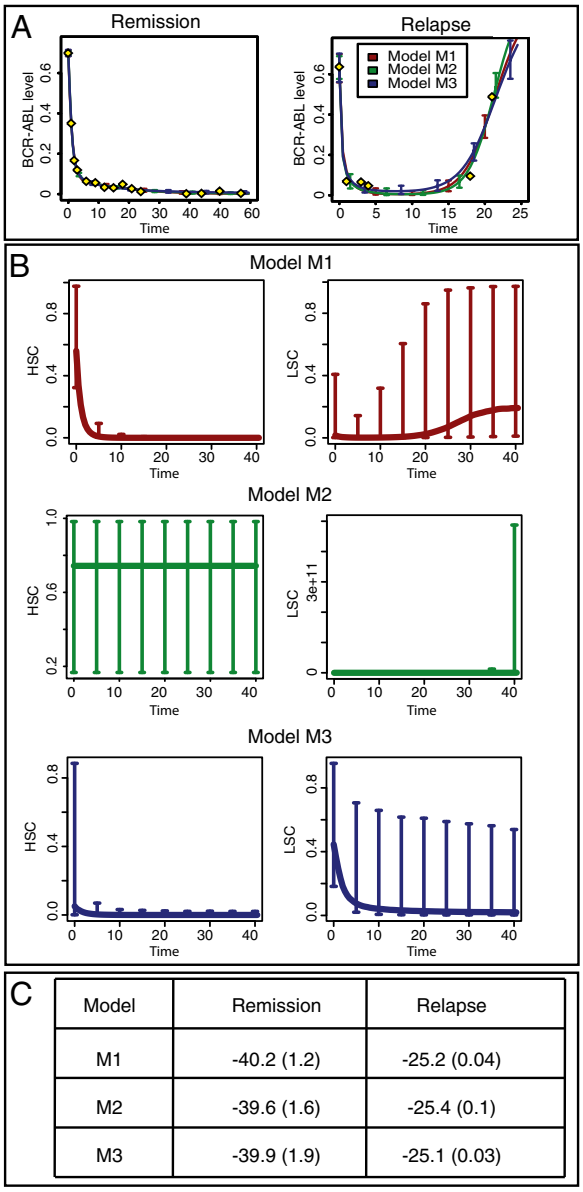


Fig. 2. Model comparison based on the fit to the data and the importance of niche competition. (*A*) Fits to the data (yellow diamonds) for the three models and the two outcomes. All three models are able to explain the observed disease progression for both outcomes. (*B*) Predicted trajectories for stem cell species for each model. The lines represent the median of the evolution and the error bars designate the 5% and 95% percentiles for a set of 1,000 parameters sampled from the posterior distribution. For model M2, LSCs exhibit unbounded exponential growth for all sampled parameter values. (*C*) Models are ranked by the probability of each to explain the observed datasets. For each model and each outcome, the logarithm of the model probability is computed using a SMC sampler; the average over 10 independent runs is displayed and the number in brackets is the SD.

similarity of the fits, we conclude that all three models are able to explain these outcomes, over some range of parameters.

To further compare the models, we use a likelihood-based Bayesian SMC sampler (22) assuming additive Gaussian measurement noise and again find no evidence in favor of any one of the three models. The model probability (log likelihood of the model given the data) for each case is shown in Fig. 2*C*, and there is no statistically significant difference between them (23). It should be noted that the data measure the ratio of leukemic and healthy terminally differentiated cells, and given information about the unobserved (stem and progenitor) species, the model probabilities

could change dramatically. As the available data alone does not allow us to discriminate between the three models considered here, we need to consider other properties of these models coupled with biological reasoning to elucidate the mechanisms underlying CML etiology and progression.

Model Checking—The Importance of Competition. The posterior distributions obtained from Bayesian parameter inference allow us to predict cell-type behavior: trajectories of individual species are simulated using parameter values sampled from the posterior. In Fig. 2*B*, the trajectories for HSCs and LSCs are plotted for each model in the case of relapse. Whereas for models M1 and M3 both stem cell species reach finite and biological plausible steady states, in the case of model M2 we find that LSCs and their progeny species grow exponentially. Thus, this model cannot represent the underlying biological mechanisms and cannot describe the dynamics of the hematopoietic system. Because our goal is to explain mechanisms of leukemia growth and decline, we cannot hope to gain much insight out of a model with such drastically implausible growth kinetics. In light of the behavior found we can reject model M2 and focus our further attention on the models M1 and M3: a model with two independent stem cell lineages fails to describe the dynamics of CML. This rejection suggests that interactions between the two lineages are important. The models considered here (M1 and M3) describe these interactions as competition between species within a niche; other models may explain the same behavior in different ways.

Model Comparison—M1 and M3 Offer Contrasting Mechanisms for CML Progression. To compare the two mechanistic models, we exploit the posterior distributions obtained from the parameter inference. It is difficult to directly identify differences between models from these distributions because none of the model parameters are tightly constrained by the data, and most of the parameters are highly correlated (see *Supporting Information*); although this may be seen as failure of the inference, it is in fact, a reflection of the data (their nature, quality, and quantity) and is itself informative. More crucially, perhaps, it highlights the importance of assessing the uncertainty of estimates instead of merely obtaining "optimal estimates" of parameter values. In the Bayesian framework we can still derive useful insights by making the posterior distribution the center of the analysis and using it to predict the behavior of the different cell types over time. The corresponding predicted trajectories for each outcome are shown in Fig. 3.

We first focus on the mechanisms leading to remission under both models. For model M1, the level of leukemic species declines over time, with differentiated leukemia cell levels being tightly constrained (Fig. 3*A*). In the healthy lineage, the HSC population initially decreases, leading to an initial peak in the level of healthy progenitor and (with greater variance and delay) differentiated cells. After this peak the behavior for all healthy species is unconstrained. Therefore, model M1 explains disease remission by suppression of leukemic species and expansion of the healthy differentiated cell populations through transient progenitor growth. This prediction would be directly testable by studying healthy blood progenitor abundances over the course of CML treatment to determine whether a growth peak occurs after initial administration of TKIs. In contrast to model M1, model M3 permits a large variety of behaviors and high species' concentrations ($\gg 1$), shown in Fig. 3*B*; this makes its behavior harder to understand and interpret. Similarly to model M1, the level of healthy cells is higher than their leukemic counterparts, which leads to the outcome of remission. It is interesting to note that the level of LSCs permitted at steady state can constitute a substantial proportion of the niche. This is surprising to see under remission (in contrast with M1) and could be tested by directly measuring species' abundances in the niche.

For CML relapse, the species' behaviors are more constrained than for remission—even though the posterior distributions are still broad. We observe similarities between the two models, in

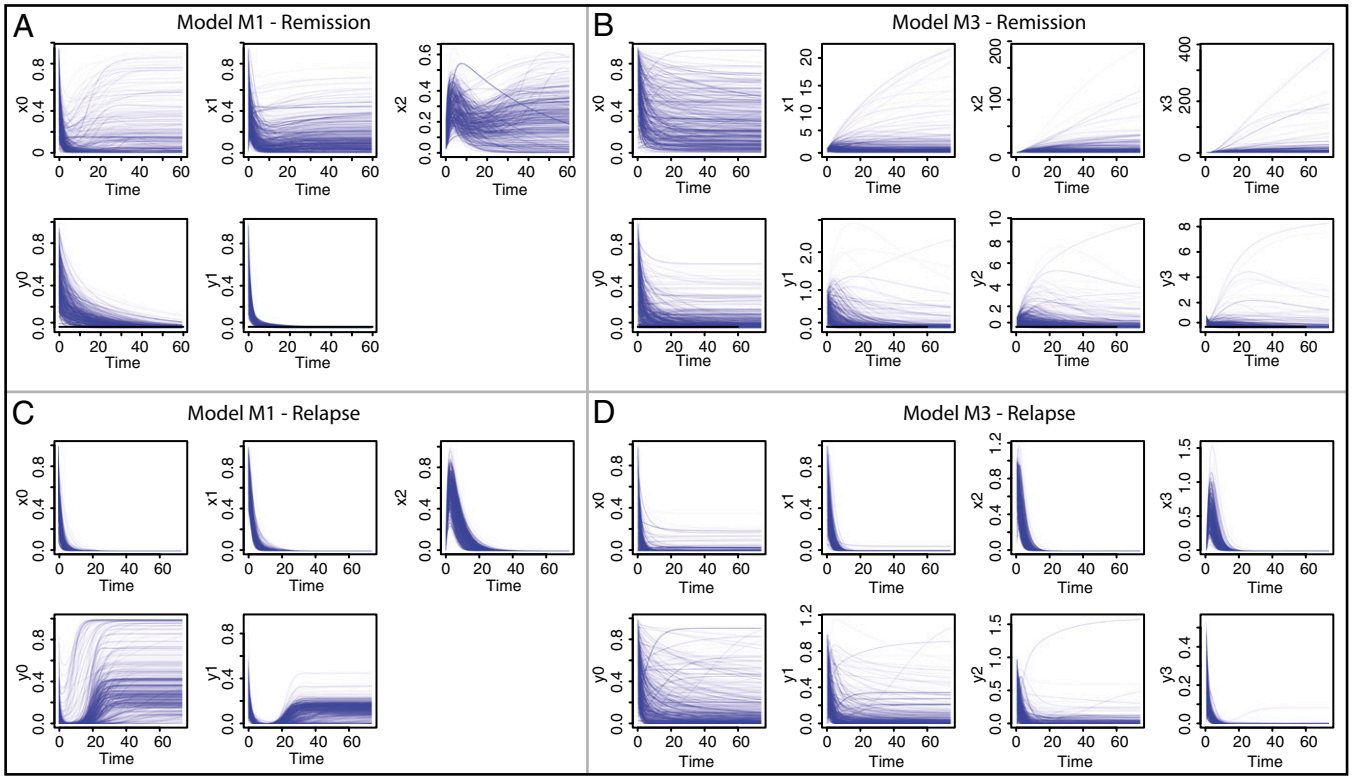


Fig. 3. The predicted change over time of each species for a timespan of 6 y (the total period of the trial), for (*A* and *C*) model M1 and (*B* and *D*) model M3 and (*A* and *B*) remission and (*C* and *D*) relapse. For each model and each outcome, 1,000 parameter sets were sampled from the posterior distribution and each line corresponds to the simulated trajectory for one of these parameter sets.

particular for the behavior of the healthy species (Fig. 3 *C* and *D*): both models require low levels of healthy differentiated cells at steady state as well as an initial peak in the level of healthy progenitor and differentiated cells, explaining the temporary suppression of leukemia.

We also find significant differences between the two models under relapse: for model M3, leukemic stem and progenitor cells can exhibit wide-ranging behaviors, whereas, for M1, leukemic stem and differentiated cells always initially decrease before increasing to fill (at least a proportion of) the niche. This transient behavior of leukemia suppression before a return of the disease is in agreement with the literature on recurrence of CML (24). The variety of species' behaviors we see for HSCs (which can occupy up to $\sim 40\%$ of niche space) and leukemic stem and progenitor cells under model M3 could be due to the number of layers in the hierarchy of M3 and the fact that we only measure fully differentiated cells. As a result, effects originating at the stem cell level are dampened in successive progenitor populations. For model M3, the differentiated leukemic cells decrease to a low steady state, similarly to the differentiated healthy cells, and the inferred rates of death/differentiation (coupled between healthy and leukemic lineages) are all high. Thus, we suggest that an explanation for the CML relapse could be that healthy and leukemia cells are pushed toward differentiation, with the leukemia lineage slightly ahead. This high rate of differentiation (or progression through the hematopoietic hierarchy) should be measurable in principle if it does indeed exist. It is intriguing that both models predict that relapse is linked to very low levels (also in the periphery) of lymphocytes, as measured in ref. 25; this should also be measurable.

Model Comparison—Disease Outcome Is Controlled by Either LSC Growth Rate (for M1) or Progenitor Death Rate (for M3). We next determine which combinations of parameters best explain the differences between remission and relapse. For each model we compare the posterior parameter distributions associated to each outcome to determine the parameter combination that best separates the two clinical outcomes. Linear discriminant analysis provides a convenient way to do this; the results of this analysis are summarized in Fig. 4. For both models M1 and M3 a clear separation was found despite the fact that the individual posterior distributions were broad.

For M1 we find that parameters a_y, b_y, and d_x explain most of the differences seen between clinical outcomes. The pair of parameters contributing most to the linear discriminant, (a_y, b_y), describe growth and differentiation of LSCs. In the case of remission both parameters have to be low for leukemic cell production to be kept in check, whereas for relapse $a_y > b_y$ is required, which maintains growth of the LSC pool. Parameter d_x, although not constrained for remission, should be relatively high for relapse; this is required to produce the sharp peak in healthy differentiated cells. Overall, for M1, controlling leukemia species is more important than controlling healthy species in determining clinical outcome.

This leads us to predict that disease outcome could be controlled by variation only in parameters related to LSC function. To test this we simulate model M1 under relapse conditions until the onset of disease relapse, and study the effect of varying a_y. We see that by decreasing a_y while keeping all other parameters constant, we can stop relapse from occurring and the patient goes back into remission (see *Supporting Information*). This intervention—selectively suppressing the growth of LSCs—thus seems promising, but is probably not yet possible to achieve clinically.

For M3 we find that a single parameter, δ_2, explains most of the variation between outcomes (Fig. 4). Relapse requires all of the differentiation/death rates to be high to produce the desired peak in healthy differentiated cells from peaks in the progenitor cell species; for remission, δ_2, should be moderate or low. We observe that finite populations of partially differentiated cells (x_2, y_2) must be maintained (explaining the low value of δ_2) so that fully differentiated cells (x_3, y_3) can be produced. Given the relative importance of δ_2, we perform a similar in silico treatment analysis as that above and test how M3 responds to a change in δ_2 at the point of relapse. We observe that although the rate of relapse is slowed, disease eventually returns and the patient does not recover (see

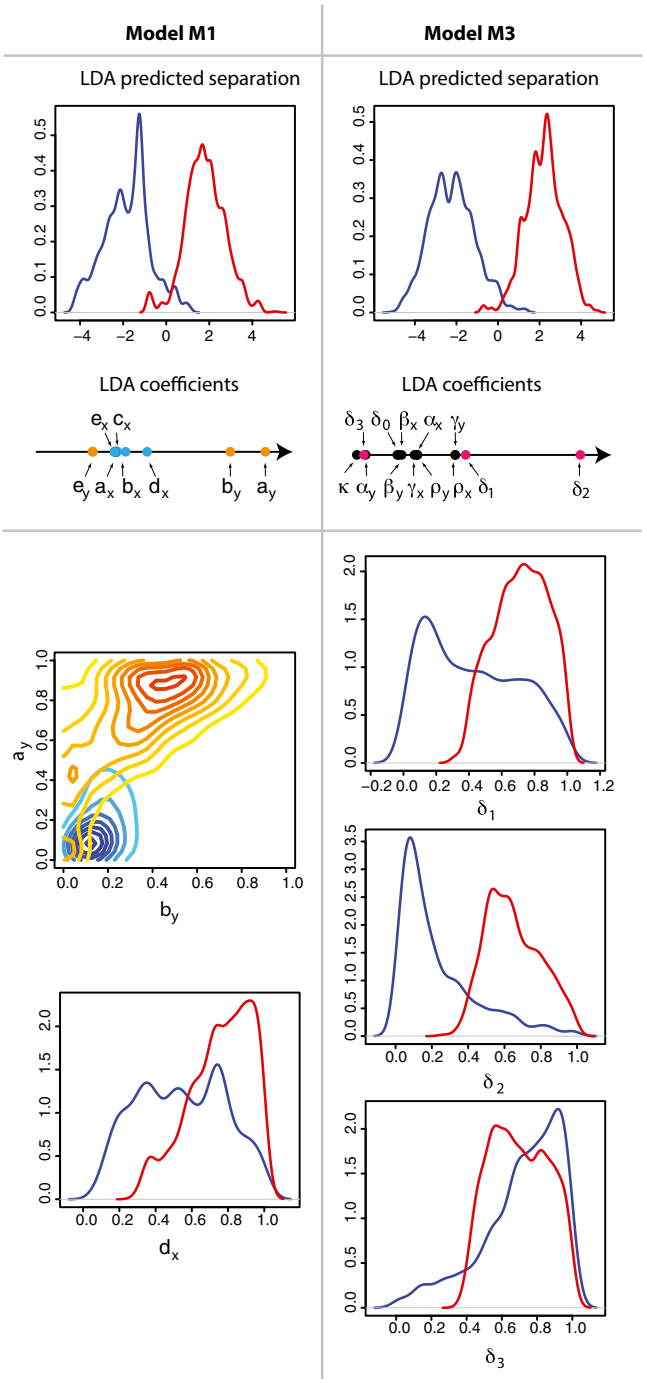


Fig. 4. Linear discriminant analysis was used to predict the key differences between remission and relapse. For (*Left*) model M1 and (*Right*) model M3, clear separation between disease outcomes can be made (*Upper*). The parameter contributions (coefficients) to the linear discriminant are shown on a horizontal axis. For model M1, the healthy parameters are represented by blue dots and the leukemic parameters by orange dots; for model M3, the pink dots designate the death rates. (*Lower*) Shown for model M1 is a density plot of a_y against b_y with the location of remission outcomes marked in blue and the relapse outcomes in red: using only these two parameters a good separation is already achieved. We also show the posterior density distribution of parameter d_x alone in case of remission (blue) and relapse (red)—moderate separability is observed. For M3 we show the posterior density distribution for each of the death rates and we see that for δ_2, remission (blue) and relapse (red) can be well distinguished.

Supporting Information). As a possible explanation for this, we note that for M3 at the point of relapse, the levels of all cell species (healthy and leukemic) are very low; at this point a return to healthy conditions might not be possible. This also suggests that there are intricate dependencies between the parameters of M3, requiring more subtle interventions to reverse clinical outcome.

Conclusion

CML has been studied in great detail and a number of models describing different aspects of disease progression have been developed in the past decade (17–20, 25–28). However, there remains uncertainty about the cellular processes that control disease etiology and outcome. Here we have approached this question by comparing three models of CML dynamics and their ability to explain clinical data on CML progression.

With our Bayesian framework we found that all of the models can be fitted to data for both remission and relapse and receive the same level of statistical support. Because of the limitations of the data we thus have to adopt more detailed model analysis and model-checking tools to discriminate between these different mechanistic hypotheses.

Model M2, for example, can be ruled out as it requires sustained exponential growth of LSCs to explain the data. The key difference between this and the other two models is that M2 does not allow for interactions in the stem cell niche; instead all cell types evolve independently. This suggests that cellular competition is an important and perhaps crucial aspect of a model seeking to describe the ecology of hematopoietic and leukemic cells inside CML patients.

Despite the limited nature of the data, our Bayesian framework allowed us to analyze the remaining models in more detail, and we used posterior simulations to explore the behavior of the different models in patients exhibiting relapse and remission. Posterior distributions, even though frequently broad, can nevertheless give rise to constrained behavior of the cell types modeled here. Crucially, our analysis has allowed us to show the effects that the niche can exert on disease progression: as more cell types—HSCs and LSCs reside necessarily inside the niche microenvironment—are exposed or contribute to the cell population dynamics inside the niche, model M3's behavior increasingly resembles that of M1. If the niche does indeed extend its influence also over progenitor cells, then our analysis suggests that clinical outcome is primarily determined by how effectively TKI therapy targets LSCs. In M3, by contrast, we found that the differences between patients showing remission or relapse are less pronounced, with the death rate of progenitor cells being the parameter exerting most influence over a patient's prognosis.

Despite the large body of knowledge gathered, and despite the successful treatments for CML, there remain large gaps in our understanding of the cellular processes that underpin this disease. Here we have identified mechanistic differences at the cellular level between different mechanistic hypotheses of CML progression that we hope will become experimentally testable as methods for probing the HSC niche advance. The ability to map the hematopoietic system and its dynamics in vivo will almost certainly also benefit our understanding of other hematological disorders.

Materials and Methods

Experimental Data. In the dataset used for analysis, CML levels are assayed through measurements of transcript levels of the BCR-ABL protein in the blood (Fig. 1), which is a proven marker for CML. Each of the 69 patients in the study is being treated with tyrosine kinase inhibitors that selectively bind to the BCR-ABL protein present on CML cells (16). Blood samples are taken every few months although not always at the same time points across all patients.

Bayesian Parameter Inference and Model Comparison. We use a sequential approximate Bayesian computation (ABC) algorithm (21) to infer the model parameters and the initial conditions for each species. ABC approaches use systematic comparison between simulated trajectories and the observed data to identify parameters associated with simulated trajectories that are closest to the observed data. Here, the time series for BCR-ABL (described above) is compared (using Euclidian distance) to the ratio $y/(y + 2x)$, where y is the number of leukemic differentiated cells and x is the number of healthy differentiated cells (20). We use the *ABC – SysBio* package, which provides a python implementation of this algorithm with graphical processing unit (GPU) support (29).

Use of a likelihood-based SMC sampler (22) enables us to compute the model evidence, which is the probability to observe a dataset under a given model. We used our own implementation of the SMC sampler algorithm in python as well as an interface to simulate the models in a computationally efficient manner using a GPU accelerated ordinary differential equation solver (30).

ACKNOWLEDGMENTS. The authors would like to thank Ingo Roeder who kindly supplied the dataset used for analysis. A.L.M. acknowledges financial support from a Biotechnological and Biological Research Council Systems Biology PhD studentship. S.F. is funded through a Medical Research Council Computational Biology Research Fellowship. M.P.H.S. is a Royal Society Wolfson Merit Award Holder.

1. Morrison SJ, Uchida N, Weissman IL (1995) The biology of hematopoietic stem cells. *Annu Rev Cell Dev Biol* 11:35–71.
2. Wang LD, Wagers AJ (2011) Dynamic niches in the origination and differentiation of haematopoietic stem cells. *Nat Rev Mol Cell Biol* 12(10):643–655.
3. Hu X, et al. (2009) Kinetics of normal hematopoietic stem and progenitor cells in a Notch1-induced leukemia model. *Blood* 114(18):3783–3792.
4. Lapidot T, et al. (1994) A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature* 367(6464):645–648.
5. Hope KJ, Jin L, Dick JE (2004) Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. *Nat Immunol* 5(7):738–743.
6. Jamieson CHM, et al. (2004) Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. *N Engl J Med* 351(7):657–667.
7. Wang JCY, Dick JE (2005) Cancer stem cells: Lessons from leukemia. *Trends Cell Biol* 15(9):494–501.
8. Schofield R (1978) The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells* 4(1-2):7–25.
9. Zhang J, et al. (2003) Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* 425(6960):836–841.
10. Calvi LM, et al. (2003) Osteoblastic cells regulate the haematopoietic stem cell niche. *Nature* 425(6960):841–846.
11. Kiel MJ, et al. (2005) SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* 121(7):1109–1121.
12. Adams GB, et al. (2006) Stem cell engraftment at the endosteal niche is specified by the calcium-sensing receptor. *Nature* 439(7076):599–603.
13. Wilson A, Trumpp A (2006) Bone-marrow haematopoietic-stem-cell niches. *Nat Rev Immunol* 6(2):93–106.
14. Fleming HE, et al. (2008) Wnt signaling in the niche enforces hematopoietic stem cell quiescence and is necessary to preserve self-renewal in vivo. *Cell Stem Cell* 2(3):274–283.
15. Mangel M, Bonsall MB (2008) Phenotypic evolutionary models in stem cell biology: replacement, quiescence, and variability. *PLoS ONE* 3(2):e1591.
16. Müller MC, et al. (2003) Dynamics of BCR-ABL mRNA expression in first-line therapy of chronic myelogenous leukemia patients with imatinib or interferon α/ara-C. *Leukemia* 17(12):2392–2400.
17. MacLean AL, Lo Celso C, Stumpf MPH (2013) Population dynamics of normal and leukaemia stem cells in the haematopoietic stem cell niche show distinct regimes where leukaemia will be controlled. *J R Soc Interface* 10(81):20120968.
18. Michor F, et al. (2005) Dynamics of chronic myeloid leukaemia. *Nature* 435(7046):1267–1270.
19. Foo J, Drummond MW, Clarkson B, Holyoake T, Michor F (2009) Eradication of chronic myeloid leukemia stem cells: A novel mathematical model predicts no therapeutic benefit of adding G-CSF to imatinib. *PLOS Comput Biol* 5(9):e1000503.
20. Roeder I, et al. (2006) Dynamic modeling of imatinib-treated chronic myeloid leukemia: Functional insights and clinical implications. *Nat Med* 12(10):1181–1184.
21. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH (2009) Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J R Soc Interface* 6(31):187–202.
22. Del Moral P, Doucet A, Jasra A (2006) Sequential Monte Carlo samplers. *J R Stat Soc Series B Stat Methodol* 68:411–436.
23. Jeffreys H (1961) *Theory of Probability* (Oxford Univ Press, Oxford, UK).
24. Mahon FX (2012) Is going for cure in chronic myeloid leukemia possible and justifiable? *American Soc Hematol Educ Progr* 2012:122–128.
25. Moore H, Li NK (2004) A mathematical model for chronic myelogenous leukemia (CML) and T cell interaction. *J Theor Biol* 227(4):513–523.
26. Colijn C, Mackey MC (2005) A mathematical model of hematopoiesis—I. Periodic chronic myelogenous leukemia. *J Theor Biol* 237(2):117–132.
27. Kim PS, Lee PP, Levy D (2008) A PDE model for imatinib-treated chronic myelogenous leukemia. *Bull Math Biol* 70(7):1994–2016.
28. Werner B, Dingli D, Lenaerts T, Pacheco JM, Traulsen A (2011) Dynamics of mutant cells in hierarchical organized tissues. *PLOS Comput Biol* 7(12):e1002290.
29. Liepe J, et al. (2010) ABC-SysBio—Approximate Bayesian computation in Python with GPU support. *Bioinformatics* 26(14):1797–1799.
30. Zhou Y, Liepe J, Sheng X, Stumpf MPH, Barnes CP (2011) GPU accelerated biochemical network simulation. *Bioinformatics* 27(6):874–876.