Intratumor heterogeneity alters most effective drugs in designed combinations

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1073/pnas.1323934111">http://dx.doi.org/10.1073/pnas.1323934111</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>National Academy of Sciences (U.S.)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Sat Dec 22 12:06:51 EST 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/93760">http://hdl.handle.net/1721.1/93760</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Intratumor heterogeneity alters most effective drugs in designed combinations

Boyang Zhao\textsuperscript{ab}, Michael T. Hemann\textsuperscript{b,c}, and Douglas A. Lauffenburger\textsuperscript{b,c,d,1}

\textsuperscript{a}Computational and Systems Biology Program, \textsuperscript{b}The David H. Koch Institute for Integrative Cancer Research, and Departments of \textsuperscript{c}Biology and \textsuperscript{d}Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Gordon B. Mills, The University of Texas, M. D. Anderson Cancer Center, Houston, TX, and accepted by the Editorial Board June 13, 2014 (received for review December 26, 2013)

The substantial spatial and temporal heterogeneity observed in patient tumors poses considerable challenges for the design of effective drug combinations with predictable outcomes. Currently, the implications of tissue heterogeneity and sampling bias during diagnosis are unclear for selection and subsequent performance of potential combination therapies. Here, we apply a multiobjective computational optimization approach integrated with empirical information on efficacy and toxicity for individual drugs with respect to a spectrum of genetic perturbations, enabling derivation of optimal drug combinations for heterogeneous tumors comprising distributions of subpopulations possessing these perturbations. Analysis across probabilistic samplings from the spectrum of various possible distributions reveals that the most beneficial (considering both efficacy and toxicity) set of drugs changes as the complexity of genetic heterogeneity increases. Importantly, a significant likelihood arises that a drug selected as the most beneficial single agent with respect to the predominant subpopulation in fact does not reside within the most broadly useful drug combinations for heterogeneous tumors. The underlying explanation appears to be that heterogeneity essentially homogenizes the benefit of drug combinations, reducing the special advantage of a particular drug on a specific subpopulation. Thus, this study underscores the importance of considering heterogeneity in choosing drug combinations and offers a principled approach toward designing the most likely beneficial set, even if the subpopulation distribution is not precisely known.

Significance

Tumors within each cancer patient have been found to be extensively heterogeneous both spatially across distinct regions and temporally in response to treatment. This poses challenges for prognostic/diagnostic biomarker identification and rational design of optimal drug combinations to minimize recurrence. Here we present a computational approach incorporating drug efficacy and drug side effects to derive effective drug combinations and study how tumor heterogeneity affects drug selection. We find that considering subpopulations beyond just the predominant subpopulation in a heterogeneous tumor may result in nonintuitive drug combinations. Additional analyses reveal general properties of effective drugs. This study highlights the importance of optimizing drug combinations in the context of intratumor heterogeneity and offers a principled approach toward their rational design.

Author contributions: B.Z., M.T.H., and D.A.L. designed research; B.Z. performed research; B.Z. and D.A.L. analyzed data; and B.Z., M.T.H., and D.A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. G.B.M. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

1To whom correspondence should be addressed. Email: lauffen@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323934111/-/DCSupplemental.
Our premise is that a profound knowledge of the subpopulations within a given tumor, as well as their dynamics and interactions, allows for the use of a linear multi-objective optimization approach to design optimal drug combinations. Each drug has the potential to minimize toxicity using a multiobjective optimization approach, scalarized using an augmented weighted Tchebycheff method, and solved iteratively using linear programming (SI Appendix). Additional drug or tumor properties (derived from prior knowledge or machine learning) may also be incorporated into this framework. Solving a multiobjective optimization problem, concomitantly incorporating efficacy and toxicity, results in a discrete set of Pareto optimal solutions, which represent an integer number of drugs included in the combination. The Pareto frontier is a surface for which any increase in one objective results in a decrease in the other objective.

subpopulation distributions for a given tumor, as would be typical in clinical situations. Among the insights gained, a key principle is that significant differences may exist between the combinations of subpopulations, projected to be most beneficial for the predominant single subpopulation and those more beneficial when multiple subpopulations are taken into account. This suggests that gaining at least some degree of information concerning the heterogeneity of a primary tumor upon diagnosis, even if not quantitative or complete, may lead to an indicated drug regimen different from what would be viewed as best if only the predominant genetic signature was obtained. Moreover, consistent with our recent experimental findings (26), our simulations demonstrate that drug effects are homogenized by tumor heterogeneity over essentially a full range of conceivable subpopulation distributions. Consequently, the effectiveness of drugs most likely to be found beneficial in combination for heterogeneous tumors may be characterized by average efficacy and robustness across the subpopulations rather than by extreme efficacy in a particular subpopulation. Taken together, our results offer conceptual principles for designing drug combinations in the context of intratumor heterogeneity.

Results

Conceptual and Computational Approach. Our premise is that a typical tumor comprises multiple diverse subclones along with a dominant subpopulation, and that dynamic changes in the distribution of these populations may be influenced by drug treatments. Currently, the prevalent clinical practice for diagnosing tumor type and drug regimen is based on histological and/or biomarker identification and generally is biased toward the predominant subpopulation as a result of its greater representation in the population (2, 27). However, minor subpopulations are important in determining how a tumor responds to treatments and may be responsible for resistance and relapse (28). An immediate critical question is, how should consideration of heterogeneity, instead of simply the predominant subpopulation, affect drug combination design (Fig. 1A)?

Our approach integrates a theoretical framework with empirical experimental information. The theoretical framework is a multiobjective linear optimization algorithm that simultaneously incorporates efficacy (the desired tumor cell killing) and toxicity (adverse side effects in patients) for drug combination effects on many heterogeneous tumor compositions (see SI Appendix for methodological details). The calculations involved in the optimization algorithm incorporate the effects of each drug on each tumor subpopulation, sampling from among the conceivable compositions of subpopulations and combinations of drugs (Fig. 1B). These compositions and drugs come from previous experimental information (26, 29). The tumor subpopulations are taken from a battery of shRNA knockdowns of genes in the DNA damage repair and apoptosis pathways in the murine Eμ-myc; p19Arf−/− lymphoma cell line [derived from a well-established preclinical Eμ-myc mouse model of human Burkitt lymphoma (30, 31)], and the drugs are taken from a series of commonly used chemotherapeutics and targeted therapeutics (Fig. S1). A key finding in the previous work is that the overall efficacy of multiple drugs in combinations typically is approximated by linear combinations of the individual drug efficacies when they are applied together at overall LD₉₀ concentration. Such additivity allows for the use of a linear objective function for efficacy with linear constraints in our optimization algorithm.

Of course, the design of drug combinations involves constraints and conflicting objectives, resulting in multiple tradeoffs to consider simultaneously. For instance, the tradeoff between efficacy and toxicity prohibits maximizing the number of drugs without exceeding the tolerable level of toxicity. Previous analyses of optimal control theory-based design of chemotherapy regimens have long used toxicity (e.g., in terms of maximal/ cumulative drug concentrations) as constraints or secondary objectives in their mathematical formulations (14). Here, our framework instead posits multiobjective optimization to maximize overall efficacy and minimize overall toxicity concomitantly (Fig. 1C and SI Appendix).

As one of the primary considerations in clinical drug combination regimens is nonoverlapping toxicity (again, adverse side effects), we formulated our drug toxicity model as a linear model, with the goal of minimizing overall toxicity with constraints on a maximal allowable toxicity for each side effect (e.g., myelosuppression, gastrointestinal effects). This formulation effectively captures nonlinear behaviors as a result of nonoverlapping drug combinations. Subsequent analyses first are based on a symmetric toxicity profile—that is, each drug has the same toxicity unit of 1. Later, this is relaxed to include the actual asymmetric toxicity profile of drugs in our efficacy dataset.

Solving a multiobjective optimization problem, concomitantly incorporating efficacy and toxicity, results in a discrete set (known as a “Pareto optimal set”) of feasible solutions (i.e., optimal drug combinations) rather than a single “absolute best” solution (Fig. S2). The set is discrete, because each solution represents an integer number of drugs included in the combination, but it resides upon a continuous curve termed the “Pareto frontier.” Each solution in the Pareto optimal set has the property that an increase in one objective (say, efficacy) results in a decrease in at least one or more objectives (then, toxicity). We enumerate the Pareto optimal solutions via linear programming calculations (see SI Appendix for technical details).

Effects of Tumor Heterogeneity on Drug Combination Efficacy and Toxicity Tradeoffs. As the first manifestation of our analysis, we used our empirical efficacy dataset (26, 29) in concert with a symmetric toxicity profile. In this circumstance, we expect the algorithm to produce as optimal solutions drug combinations with the most efficacious drug plus additional drugs incorporated
Intratumor heterogeneity linearizes the Pareto frontier and homogenizes drug combination efficacy. (A) Representative objective space showing the Pareto frontier and the tradeoff between toxicity and efficacy for a homogeneous population of murine Eμ-MYC p19ARF−/− lymphoma cells infected with shp53 hairpin. The plot was generated based on the optimization described in Fig. 1C using efficacy data composed of single-drug efficacy for individual subpopulations (Fig. S1) and a symmetric toxicity profile for each drug. (B) Representative objective space showing the effects of a heterogeneous population containing a predominant shp53 subpopulation and 13 minor subpopulations on the Pareto frontier. (C and D) Representative objective and solution space (shown for the same population as in B), with a maximum toxicity constraint of 6, which is set based on the number of drugs present in commonly used combination regimens. In the context of the mathematical model, this refers to the value for the parameter $C_{\text{max}}$ (SI Appendix). Compromise solution (colored green) refers to the point closest to the utopia based on an $L_\infty$ norm distance metric.

...into the combination in order of decreasing efficacy. Partnering a symmetric toxicity profile with decreasing efficacy would suggest a deviation from linearity from the Pareto frontier curve as the number of drugs in the combination is increased. Fig. 24 shows a representative example of this anticipated tradeoff between toxicity and efficacy on the Pareto frontier for the homogeneous tumor population (100% shp53). Interestingly, however, as heterogeneity is introduced to the population (55% shp53 plus 13 minor subpopulations), the shape of the Pareto frontier becomes surprisingly linear (Fig. 2B and Table S2). This behavior, the shift from nonlinear Pareto frontier curve to linear as a tumor becomes more heterogeneous, was also observed with other subpopulation distributions (Fig. S4 and Table S3). The meaning of this in tumor treatment terms is that heterogeneity homogenizes the benefit of drug combinations, because each additional drug has differential effects on each subpopulation.

In fact, most clinically used chemotherapeutic regimens consist of multiple drugs, such as ABVD (doxorubicin, vinblastine, bleomycin, and dacarbazine), Stanford V (vinblastine, doxorubicin, vincristine, bleomycin, mechlorethamine, etoposide, and prednisone), and BEACOPP (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) for Hodgkin lymphoma. Accordingly, we imposed an upper limit of six drugs in our subsequent analyses. The consequence of this limit for the same heterogeneous tumor in Fig. 2B is shown in Fig. 2C (the corresponding Pareto frontier) for direct comparison. Fig. 2D enumerates the drugs elicited by the algorithm for each of the regimens, from one drug to six drugs.

Monte Carlo Analyses of Drug Combinations in Diverse Heterogeneous Tumor Populations. The results above were obtained for single heterogeneous tumors possessing a specified subpopulation distribution. In practice, from a clinical tumor biopsy, one might gain data on the genetic variants present but not with explicit quantitative distribution proportions. To deal with this challenge, we performed Monte Carlo sampling of heterogeneous tumors from among the possible quantitative distributions for a given set of genetic variants and analyzed the resulting frequency distribution of component drugs in six-drug combinations. Overall, 10,000 heterogeneous populations were generated with each population sampled, based first on the number of subpopulations, followed by the subpopulation genetic variants, and subsequently the subpopulation proportions. Upon drug combination optimization, we obtain predictions of the six-drug combinations most likely to be beneficial for a heterogeneous tumor with the specified genetic variants, averaged over potential unknown quantitative proportions. These genetic variants are characterized by a set of up to 30 shRNA knockdowns, targeting genes in the DNA damage repair and apoptosis pathways (Fig. S1). The component drugs most frequently included in the optimal drug combination for this class of heterogeneous tumor were rapamycin, 5-fluorouracil (5-FU), vincristine, and sunitinib (Fig. 34). We term these the “dominant” drugs. Although these drugs do not covary based on clustering analysis over all populations (Fig. S5), when we analyzed the frequency of drug inclusion as a function of tumor complexity (i.e., the number of subpopulations in the heterogeneous tumor), we observed a strong dependence for a select number of drugs: some positively and some negatively correlating with tumor complexity (Fig. 3 B and C). For instance, with increasing tumor complexity, we observed an increase in the frequencies of inclusion of the four drugs noted above but a decrease in the frequencies of some other drugs (such as NSC3852, temozolomide, roscovitine, cisplatin, mitomycin C, and camptothecin). In Fig. 3B, these relationships are illustrated with positive or negative trends, respectively, and Fig. 3C quantifies them in terms of Spearman correlation coefficient values. Furthermore, tumor complexity is a crucial characteristic, because this strong correlation is abrogated when we control for tumor complexity (by examining the subsequent partial correlation) (Fig. S6).

To expand the generalizability of our insights further, we also analyzed combinations of numbers of drugs other than six; i.e., representing other regimens along a Pareto frontier. The
frequency distributions of the component drugs were found to be similar among the various $N$-drug combinations (annotated as "Regimen N"), as calculated in terms of the correlation between the drug frequency distribution in Regimen N vs. that in Regimen M (Fig. 3D and Fig. S8). In other words, the drugs more and less dominant in a regimen combining $N$ drugs are similarly more and less dominant in a regimen combining $M$ drugs. Thus, the behavior of drug dominance among a constellation of potential drugs as a function of heterogeneous tumor complexity is consistent across a broad range of multidrug regimens.

We next asked how the drug frequency distributions might change if we determine the optimal combinations based solely on the predominant tumor subpopulation instead of considering all the subpopulations present in each heterogeneous tumor. Using the same set of 10,000 theoretical tumors, we found only small changes in the overall distribution of drug frequencies within optimal combinations under this approach (Fig. 4A, in comparison with Fig. 3A). However, quite surprisingly, the correlation between drug dominance and tumor complexity is abrogated dramatically (Fig. 4B in comparison with Fig. 3B, and Fig. 4C in comparison with Fig. 3C). For example, the choice of whether to include drugs such as rapamycin, 5-FU, vincristine, and sunitinib no longer strongly depends on tumor complexity, but instead varies sensitively with the specific individual predominant subpopulation. As a consequence, the drug distributions found in optimal combinations across $N$-drug regimens may be very different from those for $M$-drug regimens (Fig. 4D and Fig. S9, in striking contrast to Fig. 3D and Fig. S8). In addition, changes in correlations are minimal when compared before and after controlling for tumor complexity, further suggesting that tumor complexity no longer is an important factor (Fig. S7). Hence, the optimal design of combination drugs based on only a predominant subpopulation—i.e., a tumor assumed to be homogeneous—potentially may exclude drugs that work effectively for highly heterogeneous tumors.

We therefore proceeded to analyze in detail the effects of these two alternative optimization perspectives—consideration of the entire heterogeneous tumor or just the predominant subpopulation—on the drug combinations predicted to be optimal for tumors whose complexity resides within the sampled 10,000 populations. We observed that there was a large proportion of tumors for which the optimal drug combinations were disparate between the two perspectives (Fig. 5A), with the fraction of optimal drug combinations for a given tumor found to be different when taking heterogeneity into account vs. not taking it into account, increasing from about 40% for one-drug regimens to almost 70% for six-drug regimens. Quantitatively, the difference in efficacy between the drug combinations optimized based on the two perspectives followed a roughly exponential profile (Fig. 5B). This means that cases for which the efficacy difference falls toward the tail of the profile dramatic disparities in therapeutic outcome when only the predominant subpopulation is considered in determining the best drug.

Now, it is conceivable that a drug most frequently deemed worthy of inclusion within a regimen optimized for a predominant subpopulation might generally also be found frequently within a regimen determined from considering the entire tumor complexity. Therefore, we probed this notion by using our 10,000 heterogeneous tumor sample, for each case asking whether the drug combination optimized based on the entire tumor complexity also contained the single-best drug for the predominant subpopulation. Notably, there were many regimens for which the drug combination optimized based on the entire heterogeneity did not contain the single-best drug for the predominant subpopulation (Fig. 5C). This proportion expectedly decreased with an increasing number of drugs in the combinations; for instance, for two- and three-drug regimens, 39% and 23% of the cases, respectively, did not include the drug that would have been best for the predominant subpopulation, and we obtained similar conclusions when we performed the same analyses on drug combinations optimized based also on the second largest subpopulation (Fig. S10). To confirm the robustness of these conclusions, we performed the same analyses but decomposed the results with respect to the predominant and second largest subpopulation proportions (Fig. S11). We found that the proportion of regimens with different drug combinations, as well as those containing single-best drugs for predominant (and second largest) subpopulations, is now near zero in many cases across the predominant subpopulation proportions—and begins to decrease expectedly only when the predominant subpopulation becomes very large compared with the remaining subpopulations (Fig. S12). We also performed Monte Carlo sampling with more than 13,000 heterogeneous populations near the boundaries, with similar results (Fig. S13). This firmly demonstrates that selecting drug combinations for a heterogeneous tumor often may not follow from simple intuition—which seemingly would involve including the best drug for the predominant subpopulation.

Taken together, these analyses indicate that intratumor heterogeneity may influence drug combinations dramatically such that the chosen regimen may have strongly disparate outcomes based on whether we examine the entire heterogeneity or just the predominant subpopulation. This insight argues for the necessity of considering at least even qualitative features of heterogeneity in selecting combination regimens.

### Sensitivity Analysis of Monte Carlo Sampling

Next we sought to examine in greater detail the specific dependencies of a particular drug choice on the tumor subpopulations present. To accomplish this, we performed a global sensitivity analysis by calculating the correlation between a given tumor subpopulation (i.e., shRNA knockdown) and the categorical outcome of whether a particular drug was chosen by the optimization algorithm to be included in the combination regimen for the
that is, are asymmetric for each drug across systemic tissue types for mouse (5) and human (6). Together, these observations suggest that drug combinations optimized based on consideration of the entire heterogeneity still contains the single best drug for the predominant subpopulation.

### Discussion

The substantial spatial and temporal intratumor heterogeneity that inevitably exists in cancer patients presents a fundamental challenge for the rational design of effective drug combinations. Here, we applied a multiobjective optimization approach, grounded in empirical experimental data. These data comprised quantitative effects of commonly used chemotherapeutics and targeted therapeutics on subpopulations of the murine Eμ-myc; p19ARF−/− lymphoma cell line generated via shRNA knockdowns of genes in the DNA damage repair and apoptosis pathways. The cell line was derived from the well-established Eμ-myc mouse model of human Burkitt lymphoma (30), and the existing p19ARF−/− potentially captures the early events in tumor evolution (31). Our computational analysis on these data generated predictions of optimal drug combinations for tumors in which the genetic heterogeneity is qualitatively characterized (in terms of the shRNA knockdowns) but quantitatively uncertain with respect to the various relative proportions in the tumor. An initial key result is that tumor heterogeneity can effectively homogenize drug efficacy, so the most effective drug combinations are those that best kill the broadest range of subpopulations. Notably, we found that optimizing drug combinations based on consideration of the entire tumor heterogeneity instead of just the predominant subpopulation may result in nonintuitive optimal drug combinations. As such, knowledge of the single “best agent” for each subpopulation does not promise that it will be an optimal choice for inclusion within the overall drug combination when the entire heterogeneous tumor is considered. We recently demonstrated successful validation of our optimization approach for selected tumor subpopulation distributions in both in vitro cell culture and in vivo mouse contexts (28).

The substantial tumor complexity that exists in patients and the detection limit of diagnostic tools undoubtedly cast uncertainty and incomplete information on the underlying tumor composition for each patient. However, our statistical sampling and sensitivity analyses offer guiding principles for the characteristics of effective drugs under tumor diversity. In particular, we discovered that a certain set of drugs dominated the solutions in 10,000 sampled tumors; this is termed a point-biserial correlation. In general, there were varying degrees of subpopulation dependency for the different drugs, negative as well as positive (Fig. 6A). In addition, the more frequently chosen drugs exhibit broader distributions of point-biserial correlation values than do the less frequently chosen drugs (Fig. 6B). Interestingly, the less frequently chosen drugs typically were characterized by a very tight range of dependencies, and were characterized most distinctly by outliers with respect to a particular tumor. In other words, the more frequently chosen drugs are relatively robust to the uncertainty concerning quantitative features of tumor heterogeneity, whereas the less frequently chosen drugs are problematically sensitive to precisely which subpopulations are present and in what quantitative proportions. This insight can be captured by the kurtosis of the distribution of the point-biserial correlation values, which was found to be strongly negatively correlated with the drug’s frequency (Fig. 6C). In addition to robustness, we also observed that, as expected, drug frequency was correlated with mean efficacy averaged over the entire set of 10,000 tumor samples. Taken together, all the results from the sensitivity analysis indicate that the most frequently chosen drugs for optimal drug combinations for heterogeneous tumors of uncertain subpopulation distribution are characterized by their mean and robustness in efficacy.

### Multiobjective Optimization Comprising Particular Drug Toxicity Along with Particular Drug Efficacy

Our multiobjective optimization model allows for the incorporation of multiple objectives in deriving optimal drug combinations, with whatever sophistication in objective might be desired. As such, we extended our analysis with greater nuance by recognizing that the toxicities differ—that is, are asymmetric—across systemic tissue types for each given drug and include this recognition in our objective functions. Accordingly, we incorporated actual dose-limiting toxicity information for each drug in the dataset (Fig. S15A). We observed that now a larger set of drug choices could be incorporated for each multidrug regimen along the Pareto frontier (Fig. S14) and that for myelosuppression and gastrointestinal effects, two of the most common side effects of the drugs in our dataset, the distribution in the number of overlaps between drugs in toxicity was effectively shifted to minimize such overlaps (Fig. S15 B and C and S16). Hence, a multiobjective optimization approach enables the incorporation of multiple drug characteristcs in examining tradeoffs in the rational design of drug combinations in the context of intratumor heterogeneity.

**Fig. 5.** Drug combinations optimized based on consideration of the entire heterogeneity may be nonintuitive. (A) Breakdown of the optimal drug combination regimens for 10,000 Monte Carlo-sampled heterogeneous tumor populations, showing proportions of solutions that are the same or different depending on whether the entire heterogeneity is considered vs. just the predominant subpopulation. (B) Distribution of the difference in efficacy between drug combinations optimized based on the entire heterogeneity vs. just the predominant subpopulation for the six-drug combination. (C) For each regimen in the Pareto optimal set, breakdown of solutions that are different based on the two optimization approaches, showing the proportions for which drug combinations optimized based on entire heterogeneity still contains the single best drug for the predominant subpopulation.

**Fig. 6.** Sensitivity analysis reveals optimal drugs are most robust and, on average, most efficacious. (A) Point-biserial correlation ($r_{pb}$) showing the dependence of component drug choice on subpopulation. The drugs are ordered according to their frequency in the optimal drug combination (Fig. S4A). (B) Distribution of point-biserial correlations for each component drug. (C) Correlation matrix of various drug characteristics. Component drug frequency is strongly positively correlated with mean efficacy and negatively correlated with kurtosis of $r_{pb}$, a metric used here for robustness.
for optimal drug combination at increasing tumor complexity, and such drugs were associated with greater average efficacy and robustness. This indicates that in the absence of complete knowledge of a tumor’s composition, we nonetheless can apply a de novo design of optimal drug combinations based on drugs’ average efficacy and robustness properties. Ultimately this requires experimental validation in a clinically relevant model. Although our dataset, acquired based on in vitro assays using clinically relevant genes and chemotherapies, is comprehensive with regard to knowledge of the efficacy of single agents in single subpopulations, such complete knowledge has yet to be realized in the clinical setting for experimental validation. Nevertheless, recent large-scale efforts, such as The Cancer Genome Atlas and the Cancer Cell Line Encyclopedia (32), are beginning to elucidate patient tumor composition and subpopulation responses. The accumulation of this type of information may provide opportunities to deconvolute and derive efficacy data for a single drug on individuals subpopulations, allowing this methodology to be used in the derivation of drug combinations sampled based on a prior distribution of known genetic variants of a cancer type. This would enable the experimental validation of many clinical samples, with responses tracked longitudinally following treatment.

Additional considerations of drug–drug interactions, including nonlinear drug–drug efficacy (e.g., drug synergy or antagonism) and toxicity interactions, and tumor dynamics undoubtedly add to the complexity of drug combination optimization. Specific genetic alterations also may affect subpopulation interactions and mutation rates, with consequences on the evolutionary trajectories of the tumor, in the absence of drug selection. These different rate parameters, generally used in stochastic and deterministic differential equation-based models of tumor evolutionary dynamics, would be particularly useful to consider in the design of sequential treatment strategies. Moreover, although our model optimized for overall efficacy by targeting all subpopulations, an optimal drug combination would require further understanding of how differential selective pressures imposed by drugs affect the potential for secondary resistance. This issue has been studied particularly in antibiotic resistance, in which strong selection by synergistic drug combinations actually may increase the risk of resistance (33, 34). Ultimately, just as multiobjective optimization approaches may be applied to derive small molecular structures with the desired polypharmacological profiles (35), a multiobjective optimization with considerations of these additional properties provides a potentially promising approach to optimizing drug combination in the context of tumor heterogeneity.

Methods

Full details of methods and mathematical models may be found in SI Appendix. Briefly, the efficacy data were derived previously by using an in vitro competition assay with murine Eµ-myc p19Arf−/− lymphoma cells partially infected with a specific shRNA and treated with a broad range of chemotherapies (26, 29). The toxicity dataset was a binary matrix of known dose-limiting toxicities for each drug. The optimization was formulated as a multiobjective optimization model with the objective of maximizing efficacy while minimizing toxicity. The optimization was solved iteratively with linear programming using MATLAB 2012b (MathWorks) and CPLEX 12.5. Monte Carlo sampling was performed with simulation of 10,000 heterogeneous tumors drawn from a specified distribution of subpopulations, followed by drug combination optimization on each tumor composition. Statistical and sensitivity analyses were performed using standard packages in MATLAB. All MATLAB codes are publicly available at http://cicbp.mit.edu/data-models.

Acknowledgments. We thank Dane Wittrup and Leona Samson for their insightful comments. Funding was provided by Integrative Cancer Biology Program Grant U54-CAI12967 (to M.T.H. and D.A.L.). B.Z. is supported by National Institutes of Health/National Institute of General Medical Sciences Interdepartmental Biotechnology Training Program 5T32GM008334.