EVALUATING HETEROGENEITY AND PHENOTYPIC TRANSITIONS IN SMALL CELL LUNG CANCER VIA MECHANISTIC MODELING

By

Samantha Petrillo Beik

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Cancer Biology

August 11, 2023

Nashville, Tennessee

Approved:

Christine Lovly, Chair, M.D., Ph.D.

Bill Holmes, Ph.D.

Tony Capra, Ph.D.

Vito Quaranta, M.D.

Carlos F. Lopez, Ph.D.

ACKNOWLEDGMENTS

This work would not have been possible without many, many supporters throughout my time in graduate school. I thank my co-mentors, Dr. Carlos F. Lopez and Dr. Vito Quaranta. I am very grateful for Dr. Lopez's guidance during my PhD as well as his willingness to entertain my ideas for my research direction, along with providing a supportive lab environment along the way. Dr. Quaranta has provided excellent feedback and career advice throughout my PhD, bringing a complementary perspective from Dr. Lopez and supporting the experimental questions I addressed in this dissertation. I could not have achieved any measure of success in the Lopez and Quaranta labs without the support of Dr. Leonard Harris. I thank Dr. Harris for teaching me the theory behind multiple types of computational modeling and answering all of my follow-up questions, even if they required a significant time investment from him. Dr. Harris is extremely generous with his time and knowledge and I am very grateful for that.

The support of the Lopez lab members was crucial for time in graduate school. I thank Dr. Martina Prugger, for many theoretical discussions and statistical explanations including homework help. I would not have had nearly as much fun in graduate school without the hiking trips we went on as a group and the swing dancing events we went to! I thank Dr. Michael Irvin for our own discussions about statistical theory, with Dr. Irvin answering very specific questions and helping improve my work, and I thank him as well for excellent drink and recipe ideas and for providing me and the Lopez lab with oyster ice cream. I thank Dr. Michael Kochen for helping me use HypBuilder for my project, and for his coding explanations including implementation of Boolean modeling. I thank Dr. James Pino and Dr. Oscar Ortega for helping me learn theory, implementation, and troubleshooting of parameter estimation. My understanding and improvement in this area was a process and they were very patient. I thank Dr. Geena Ildefonso for her advice and productive discussions about code improvement. I thank Melaine Sebastian, who was a bright light in the lab and who we all miss. I thank Bryan Glazer and Perry Wasdin for many interesting discussions about the interface of computation/math and biology and the details of representing biology in code.

The Quaranta lab members provided experimental support and welcomed me into their lab more fully in my final year of the PhD. I would especially like to thank Dr. Sarah Groves, Dr. Clayton Wandishin, and Dr. Jing Hao. Dr. Groves was generous with her time helping me with both small cell lung cancer cell line culture details, tips on lab equipment, and computational advice with single-cell RNA sequencing data. Dr. Wandishin taught me all about molecular cloning and always answered my questions, no matter how basic, about bringing the experimental ideas in my dissertation as far as possible. Dr. Hao answered all of my questions, catching me up on how the Quaranta lab runs, where reagents are kept, and the nitty gritty details of bench work and determining the right supplies.

While mentioned already for individual contributions to my time in graduate school, I would like to extend another thank you to the Movie Night/Day group, Dr. Harris, Dr. Prugger, Dr. Kochen, and Dr. Ortega. Movie Day kept me socializing during COVID and now keeps the far-flung alumni of the labs in contact monthly, and introduces me to movies/media/culture I would not have found or considered on my own.

I am grateful to the members of my thesis committee, Dr. Christine Lovly, Dr. Bill Holmes, Dr. Tony Capra, and Dr. Pierre Massion. I thank them for the in-depth discussions we were able to have, which shaped my work and helped guide my dissertation. I thank them as well for being available outside of committee meetings for more specific and detailed discussions, and for career advice, especially from Dr. Lovly. I thank as well the two Directors of Graduate Study for the Cancer Biology Graduate Program during my tenure in graduate school, first Dr. Jin Chen and then Dr. Rachelle Johnson. They were extremely helpful in navigating the graduate program from the perspective of an MSTP student and for ensuring everything administrative could come together for the transition from graduate school to medical school. I thank Dr. Julien Sage for collaboration on my population dynamics modeling publication, providing insightful feedback from an experimental point of view. I thank Dr. Jonathan Lehman for his mentorship, both career-wise and that related to the timecourse mouse xenograft experiments and CyTOF data contained in my dissertation.

I thank the Vanderbilt Medical Scientist Training Program. I thank Dr. Chris Williams for his support, and Dr. Lourdes Estrada for providing advice throughout all levels of my training in the MSTP. I thank Bryn Sierra for providing answers to many last-minute questions over the years. I gratefully acknowledge the funding sources that have supported my education and my research, particularly F30CA247078, T32GM007347 (Vanderbilt MSTP), and T32LM012412 (Vanderbilt Training Program in Big Biomedical Data Science). I thank my MSTP entering cohort (class of 2023/2024), and I am extremely grateful for the friendship of Dr. Elizabeth Moore, Dr. Ayesha Muhammad, Dr. Danny Sack, and Dr. Thao Le, sustaining me both nutritionally and in commiseration over the years through ups and downs of all our PhD projects.

I am very lucky to have the support of family and friends throughout this lengthy process. I would like to thank Chelsea Bathurst and Quintin Marcus for their friendship despite the distance between us. I thank my extended family, Mimi and Joe Canio, Flossie Ierardi and Elinor Dalessandro, and my parents and siblings, Patti Ierardi, Dave Beik, Allie Beik, Tim Davis, and Peter Beik. I would like to thank my cats Westley and Buttercup, but they cannot read and so this would be wasted on them. Finally, I would like to thank my partner, Mike Tackeff, for always encouraging me to pursue my dreams and supporting me along the way.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

1.1 Theories of cancer persistence

Cancer is a major public health burden: it is the second leading cause of death in the United States, with 1,958,310 new cases and 609,820 cancer deaths projected for 2023 [2]. Nearly 50 years ago, Peter Nowell characterized cancer as an evolutionary process, undergoing Darwinian selection where a mutation in one cell leads to a fitter tumor subclone—better able to withstand nutrient-poor conditions and immune pressures—that expands to populate a large portion, if not majority, of the tumor population [3, 4]. Treatment based on mutations, successful in some situations, was thought to fail when tumors included multiple subclones and a fitter subclone, unaffected by drug, continued to grow during and after treatment. In the ensuing years, two views were built upon: the view of phenotypic rather than genetic heterogeneity - a fitter phenotype if not genotype that could continue to grow despite treatment; and the cancer stem cell (CSC) theory - where a multipotent cancer cell could repopulate an entire tumor if not eradicated [5, 6]. Phenotypic heterogeneity and the CSC view were considered analogous to cell behavior in organism development and upkeep, where less-differentiated cells populate a tissue with more-differentiated and tissue-specific cells, and tissues are made up of multiple phenotypes.

Lung cancer is the leading cause of cancer death among men and women in the United States, with small cell lung cancer (SCLC) the most aggressive subtype. SCLC on its own is the 8th leading cause of cancer death, with a projected 35,751 cases and 19,060 deaths in the United States in 2023 [7, 2]. Nearly all SCLC tumors have somatic mutations inactivating TP53 and RB1, [8] yet the knowledge of this relatively homogeneous mutational landscape did not lead to improved treatment options. SCLC has been determined to be phenotypically heterogeneous across tumors, with varying phenotypic subtypes that may exist singly or as multiple phenotypes within the tumor [9]. This heterogeneity, and phenotypic transitions and trophic support between subtypes, [10, 11] is thought to underlie typical SCLC behavior. Initial therapy overall response rates are high at 80%, yet, the vast majority of SCLC cases recur and lead to death within 2 years of diagnosis [12].

1.2 Mechanistic modeling applied to cancer

Mechanistic modeling has improved our understanding of cancer dynamics. In 2011, Gupta and colleagues [13] found that breast cancer cells they had characterized as non-stem were able to grow to recapitulate *in vitro* equilibrium populations that included CSCs - indicating de-differentiation activity. Their model predicted un-

expected cell growth dynamics during the growth process from the sorted, single-phenotype population to the equilibrium population, which they validated experimentally [13]. These results indicated that the hierarchical view of the CSC and more-differentiated tumor cell phenotypes that can arise from it was not accurate: any cell in the tumor population could undergo a phenotypic transition not only into the less-differentiated CSC, but into any of the phenotypes present in the population at equilibrium. Others continued to model heterogeneous populations and their return to equilibrium after perturbation, [14, 15, 16] or adaptation to a new equilibrium given a prolonged insult [17, 18, 19]. Paudel and colleagues [20] investigated, quantified, and modeled equilibrium populations and growth-dynamics phenotypes, evaluating transitions between these phenotypes in the presence and absence of drug. They found that the treated population re-equilibrates to a new phenotypic composition, and suggest the development of concurrent or sequential treatment with drug combinations to move the population from one equilibrium to another, and then take advantage of sensitivities in the new equilibrium ("targeted landscaping") [20].

Considering mathematical modeling's contributions to understanding tumor cell population heterogeneity, its potential for non-hierarchical differentiation via phenotypic transitions, and the tendency to return to treated or untreated population equilibrium, I aimed to apply these principles to SCLC. Evaluating the phenotypic landscape of SCLC and what kind of phenotypic transitions or cell-cell interactions support the tumor's phenotypic equilibrium during development, or after a seemingly successful treatment, could shed new light on this disease.

1.3 Mechanistic model generation and evaluation

A mathematical model should be interconnected with experimental results, whether experiments are designed specifically with the mathematical model in mind, or if results are drawn from the literature or related experiments (**Figure 1.1A**). After building the mathematical model (see below), the investigator will implement it computationally and run model simulations [21]. Next, the investigator will examine simulation results, (experiment *in silico*) and based on these results, will potentially revisit the assumptions used to encode the model or even collect more data. The investigator will then make predictions based on the *in silico* data, to be validated (or refuted) experimentally. Once such experiments are performed, experimental behavior can be compared to model behavior. Whether or not the model behavior was validated or refuted, the experimental data improves the model: if predictions were not correct, the model can be changed to more accurately reflect what occurs experimentally, and if they were correct, more features can be added to the model or a tighter concordance between data and simulations can be encoded.

One common way to build a mathematical model is to review literature and data related to the system of interest, whether published data or one's own generated data, and use this review to elucidate model variables

(for example, cells of a particular phenotype) and the ways in which they act and interact within the system (for example, cell division or trophic interactions) [21]. However, mathematical models by design require the specification of precise relationships within a biological system [22] and unless the data to be used for modeling was generated specifically with the mathematical model in mind, historical data is unlikely to provide the precision required. This can be as simple as literature not addressing whether or not two cell phenotypes in a population interact, and the modeler must decide whether the lack of evidence means a correct model will not include such an interaction. In a more complicated example, if one finding from an experimental model for a disease includes evidence for a model variable or action/interaction, but another finding from a different experimental model of the same disease did not find evidence for that variable or action, the modeler must determine whether to include or exclude the model variable or action.

Despite these potential pitfalls, the precision of relationships between biological variables in a model and their actions [22] enables a modeler to hypothesize, and subsequently test, the existence of these relationships through modeling. Model variables and their actions or interactions in a model can be considered hypotheses: if it is not clear whether a biological process occurs in the system of interest, including the mathematical representation of the process in the model represents the hypothesis that it does occur in that system. Comparing models with and without this process represented, and how well each model recapitulates experimental data, is a test of the hypothesis that such a process is present in the system from which the data was derived.

1.4 Model selection and model averaging

Model selection in biology typically involves building a set of plausible models and evaluating each candidate relative to a given experimental dataset. The goal is most often to identify the "best" model of the set - that with the highest result of a scoring function comparing it to the dataset of interest. This particular approach has a significant limitation in that there may not be one top model, and instead a (potentially large) group of models with similarly high scores. In this case, the investigator may compare the features common to the high-scoring group of models. This challenge has been addressed in several ways in biology. Investigators have noted the proportion of top-scoring models containing each possible model interaction [23]. Parameter importance analysis, where candidate model scores are transformed into probabilities and summed according to the presence of model variables or actions/interactions, is frequently used in conjunction with Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC), information theoretic model selection methods [24, 25]. However, in recent years, parameter importance analysis has been noted to have several flaws [1, 26].

Bayesian model averaging describes model uncertainty, enabling calculation of Bayesian posterior distributions for model parameters as well as models themselves [27]. Investigating model probability in the

context of Bayesian inference, and by extension probability of model variables and actions, requires that the model selection method used return a likelihood for use in Bayes' Theorem. As above, most model selection investigations in biology use information theoretic approaches like AIC, however, the nested sampling method developed by Skilling [28] and employed in astronomy and astrophysics, provides the Bayesian evidence, or marginal likelihood, per model. The implementation of nested sampling via Multinest [29, 30, 31] enables a Bayesian approach to model selection problems in biology, and in the case of this dissertation, SCLC.

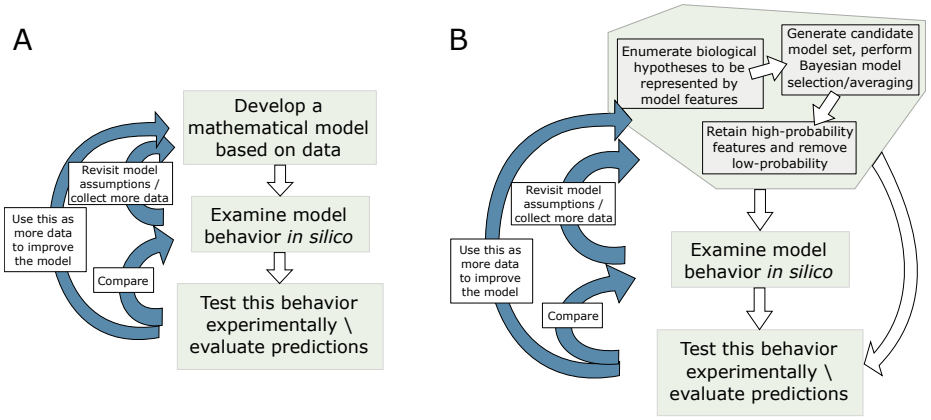


Figure 1.1: Model building and comparison to data

A To develop and use a model, the model should be built based on gathered data, simulated, then its results evaluated experimentally. The completion of each step may require revisiting the previous step. Whatever the experimental outcome, the data it provides can be used to improve the model.

B The development of a model can be built upon using Bayesian multimodel inference, engaging with the biological hypotheses to be represented in the model and building a candidate set with combinations of features representing these hypotheses. The probabilistic outcome of this process can be used to determine what model features to move forward with in the process or to test predictions of which model features or actions are present.

1.5 Using Bayesian inference and computational tools to evaluate biological systems, including small cell lung cancer

Given that features of models can stand as biological hypotheses to be tested computationally, here I move beyond the process of model selection and model averaging to consider the biologically processes represented in a set of candidate models. In developing this multimodel inference approach, I not only investigate Bayesian inference as it relates to model selection, (**Chapter 5**) but make predictions about processes in SCLC (**Chapter 6**). SCLC has multiple experimental model systems, and even the canonically recognized phenotypic subtypes that can be singly or multiply present in human tumors were determined using differ-

ent model systems per subtype (Rudin 2019). This is often the case in the research of a particular tumor, and the questions investigated about SCLC in this dissertation – differentiation hierarchy, phenotypic plasticity, trophic or other cell-cell interactions – are applicable to many cancer types. This multimodel inference methodology can be used to build models using experimental data not specifically designed for those models, or using data from multiple experimental systems; this process can be incorporated into the first step of the systems biology model-experiment cycle (**Figure 1.1B**).

In this dissertation, I aim to address both mechanistic modeling, and tumor features that must be contended with before optimal treatments can be designed. The optimal data for evaluating the aforementioned tumor features are measurements of the same tumor cells over time. This dissertation includes assessment of the same SCLC patient-derived xenograft tumors sampled over time with or without human-relevant treatment, (**Chapter 7**) from which I can also learn about phenotypic changes and tumor evolution over time. Without visualizing a phenotypic change, tumor behavior can be suggested rather than conclusively determined, and I put forward a means for visualizing predictions from my multimodel inference approach (**Appendix A**) to draw conclusions. With these various efforts, I enable a move toward better understanding features of SCLC and potentially cancer in general.

CHAPTER 2

Phenotypic plasticity and population equilibrium in tumors

2.1 The epigenetic landscape and non-genetic tumor evolution

Tumor development and response to therapy – often an incomplete response, or a response followed by recurrence – are still incompletely understood. Cancer has long been known as a dynamic process, where a tumor evolves over time, responding to and evading treatment. While this evolution encompasses the accumulation of genetic mutations, non-genetic evolution within tumors has been recognized as an important driver of the dynamics enabling tumors to withstand treatment [5]. Phenotypic plasticity, the ability of cells to transition between phenotypes without undergoing genetic changes, is an important part of this dynamic process. Plasticity allows tumor cells to maintain the phenotypic equilibrium within a tumor during growth, and maintain or develop new equilibria in response to treatment. Understanding the dynamics of these processes and the molecular mechanisms involved will be crucial for developing new treatment strategies to prevent cancer resistance and recurrence.

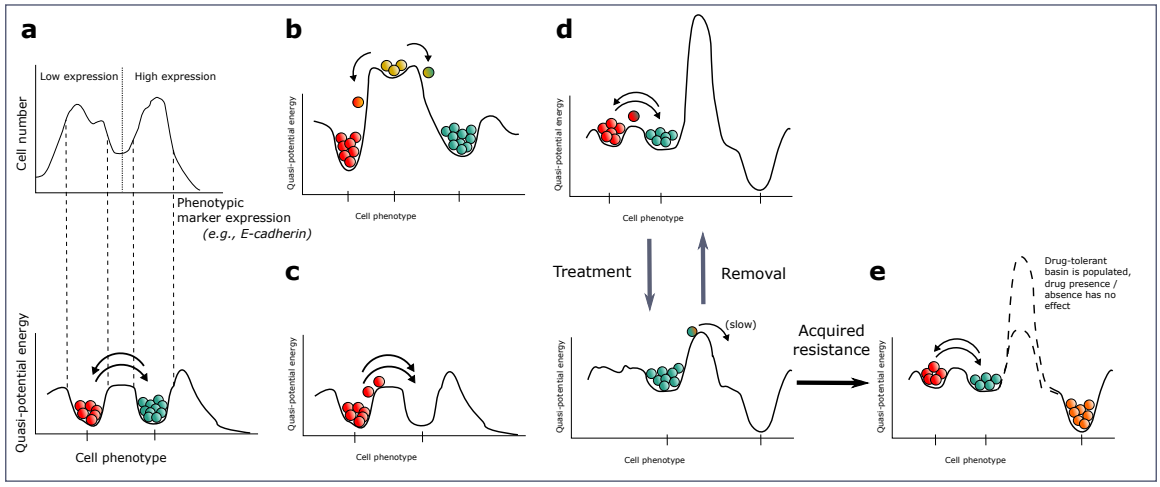


Figure 2.1: Phenotypes as basins in a quasi-potential energy landscape

A Gene expression levels (above) determine cell phenotype, which can be visualized as a basin in a quasi-potential-energy landscape (below)

B Cancer stem cell hypothesis represented as a landscape: shallow basin with few cells populates remaining basins; cells in these basins cannot transition back to the stem cell basin.

C Depleting cells of one phenotype results in re-equilibration of cells in the population

D Treatment changes the landscape, and removal after a short time changes it back.

E Landscape change under continuous treatment lasts long enough for rare cells to populate a previously inaccessible basin. This results in ”acquired resistance”, where cells populate a new, deep well and are able to remain there regardless of treatment presence or absence.

Adapted from Fig. 1 in [32].

Equilibrium within tumors can be conceptualized as an "epigenetic landscape", initially described by Waddington [33]. An epigenetic landscape represents distinct stable states – phenotypes – by "basins" of attraction, wells in the quasi-potential energy landscape where cells reside (**Figure 2.1A**). Within this landscape, transitions may occur between basins across "barriers": this represents cellular phenotype transitions [32]. The basins and barriers comprising the landscape are fundamentally based on molecular networks that regulate gene expression, generating transcriptional profiles representing each phenotype in a cell population [34]. Gene expression "noise" enables cells to stochastically transition between basins, thereby undergoing phenotypic transitions [35]. Thus, for a given cell population, "equilibrium" is the result of complex interactions between these gene regulatory networks, intracellular and extracellular signals, and gene expression noise. At equilibrium, these interactions maintain the phenotypic proportions within the cell population. If perturbations deplete a certain phenotype, cells will transition across basins to re-equilibrate within the landscape. Drug treatment alters molecular interactions and thus modifies the landscape, causing re-equilibration to the altered topography.

2.2 Cancer stem cells

Within a developing organism, embryonic stem cells have identical DNA, but are responsible for generating all the cells of different tissues, such that the adult organism contains cells of many different phenotypes. In adults, rare adult stem cells reside in tissue compartments and participate in tissue upkeep. Stem cells are the apex of a hierarchical lineage, where they generate increasingly differentiated cells. Stem cells maintain homeostasis, the phenotypic population equilibrium of tissues [36].

The cancer stem cell (CSC) hypothesis posits that tumors are arranged in a hierarchy similar to normal tissues, with CSCs at the apex, generating differentiated tumor cells. Several studies showed that even when tumor cells arise from the same genomic clone, only rare tumor cells – the CSCs – have the capacity to regenerate a tumor in an animal model [37]. By extension, CSCs are responsible for maintaining the tumor – including its phenotypic population equilibrium [37]. Taking this view, the epigenetic landscape of CSCs could resemble **Figure 2.1B**. CSCs occupy a shallow basin with high quasi-potential energy, accommodating only rare cells. This enables proliferating cells to transition to other basins and generate all phenotypes within the population. It is not possible for cells in the other basins to overcome the barrier height and transition into CSCs.

Several different signaling pathways have been characterized as responsible for the CSC phenotype, including Wnt, Notch, and JAK/STAT, implicating these networks in maintaining the CSC landscape [38]. For a more complete description of CSCs and current research, the reader is directed to [39].

More recently, it has been shown that phenotypic plasticity as it relates to tumor cell lineage is not limited

to CSCs, but that differentiated tumor cells are able to "de-differentiate". The fact that a transition toward the CSC phenotype exists indicates that the transition to the CSC basin is possible, and that the landscape may not be as specialized as **Figure 2.1B**. Studies investigating the transition toward CSCs will be addressed here.

2.3 Phenotypic plasticity and equilibrium during tumor growth

The cancer stem cell hypothesis postulates that the CSC initiates a tumor, generating different phenotypes, maintaining equilibrium within the tumor landscape. Yet in melanoma, researchers found that cells of different phenotypes – slowly-proliferating tumorigenic ("invasive"), or, highly proliferative – can each singly initiate tumors that, at steady state, include both phenotypes [40, 41]. This occurs through modulation of the Notch pathway [41]. Notch maintains stemness, indicating that one of these subpopulations is stem-like [42]. Another melanoma study found that sorted, non-stem cells are as likely to initiate tumors *in vivo* as sorted CSCs [43]. That each subpopulation can generate both phenotypes indicates that CSCs are not the only tumor-initiating cells (as the CSC hypothesis states). In the epigenetic landscape, sorting populations such that only one basin is filled nevertheless results in re-equilibration (**Figure 2.1C**), filling the remaining basin(s) such that the tumor maintains both phenotypes.

The transition to a CSC phenotype has since been studied in other tumors. In transformed human mammary epithelial cells, sorted differentiated cells generate populations of both differentiated and stem-like cells *in vitro* and *in vivo* [44]. Colorectal cancer cells transition from differentiated cells into CSCs, and sorted populations of CSCs or non-stem cells always achieve the same proportions after several weeks [45]. Colorectal cancer cell transition into CSCs is mediated by Wnt and NFKB signaling [46, 47]. In glioblastoma (GBM), cells without neurosphere-forming ability (non-CSCs) can gain this ability, transitioning into a population that includes CSCs [48]. SOX2 drives this ability [47]. In contrast to the former studies [46, 47], SOX2's silencing of the Wnt pathway enables the transition to CSCs in GBM [49].

The dynamics of the transition to CSCs from non-stem cells include an "overshoot." In a population of sorted non-stem cells, a stem cell proportion emerges and increases rapidly, then decreases to steady-state [16]. This indicates a feedback mechanism maintaining stem and non-stem population equilibrium, likely the result of extracellular signaling between these cells [50]. In melanoma, CSC return to steady-state is mediated by microRNAs targeting the Wnt signaling pathway [16].

In some tumor systems, cells generate populations of supportive cells through transdifferentiation, phenotypic transitions across cellular lineages. Glioblastoma stem cells (GSCs) have been shown to transdifferentiate into endothelial cells to enable tumor neovascularization *in vitro* [51] and *in vivo* [52]. The transition to endothelial progenitors is mediated by Notch [51]. GSC-derived angiogenesis is essential for tumor survival, indicating the necessity for transdifferentiation in this tumor [52]. In small-cell lung cancer (SCLC),

a neuroendocrine, tumor-propagating cell subpopulation generates all other SCLC subpopulations in mice [53]. Neuroendocrine SCLC cells can transition into non-neuroendocrine cells via Notch signaling, whereupon non-neuroendocrine cells provide trophic support [10]. Though these transitions occur between cells of different lineages, they represent transitions between basins in a landscape representing the tumor and supportive cell population.

2.4 Response to treatment: drug tolerance

Application of drug changes the epigenetic landscape, altering barriers along with basins. Tumor cells are depleted by treatment, but the remaining population will re-equilibrate within the new landscape. In several tumors, treatment response – re-equilibration – is a transition toward the CSC phenotype. Leukemia cells develop a drug-tolerant, stem-like phenotype after radiation [54]. Pisco et al. [15] showed that after vincristine treatment, cultured leukemia cells shift toward a stem-like population where drug is rapidly effluxed. This transition is dependent on Wnt signaling [15]. Sorted hepatocellular carcinoma non-stem cells generate spheroids upon radiation treatment, indicative of a transition to CSCs [55]. This is mediated by SOX2 and OCT4 [55]. Androgen deprivation induces a CSC phenotype in prostate adenocarcinoma cells, through activation of STAT3 [56]. In GBM, treatment with PDGFR inhibitors induces a small proportion of GBM cells to transition to a reversible, slow-cycling, drug persister state [57]. These cells have increased expression of stemness genes and are dependent on Notch signaling [57]. These treatments shift the landscape and create a new equilibrium where a stem-like state is prevalent, perhaps by decreasing the height of the barriers and deepening the CSC basin (**Figure 2.1D**).

The ability to tolerate drug application is not always the result of transitions toward a stem-like state, but instead transition toward other drug-tolerant states. In contrast to [57], treatment of a GBM stem-cell-enriched population with an alkylating agent shifts the tumor proportion from a mostly stem-like phenotype to a differentiated phenotype that can survive in the presence of this treatment; treatment removal causes the population to shift back to its former composition [58]. Comparing [57] and [58], it is likely that the different treatment modes result in different effects on GBM cells. In pancreatic ductal adenocarcinoma, tumor cells tolerate KRAS-inhibition via enhanced focal adhesion kinase signaling, which returns to normal levels when KRAS-inhibition is ceased [59]. Risom et al. [60] showed that cell-state transitions in breast cancer between different phenotypes – luminal, basal, mesenchymal – occur upon treatment, and when drug is removed the tumor reverts to its drug-naive composition. The authors confirm that these transitions are the major driver of drug tolerance [60]. These observations cannot be ignored, given efforts to supersede tolerance by developing CSC-targeting therapies. Such therapies would not affect a population whose continued survival does not depend on transitions toward CSC phenotypes.

Up to this point, observed treatment tolerance is reversible. It is likely that treatment depletes cells in particular basins while altering the landscape, and the remaining cells populate basins representing drug-tolerant phenotype(s). When treatment is removed, the landscape reverts to its status pre-drug, and cells re-equilibrate back to their original phenotypic proportions within the population (**Figure 2.1D**). Importantly, phenotypic proportions will be the same as pre-treatment, but absolute cell number will be lower. This is the basis for the "drug holiday," where treatment is stopped for a period of time, the tumor is "re-sensitized" in its absence, and further treatment has an observable effect [61].

However, there is a time-dependence to treatment reversibility: with continuous treatment, irreversible tolerance occurs. A subpopulation in non-small cell lung cancer (NSCLC) achieves a negligible growth state for several weeks under treatment, eventually developing stable tolerance [62]. These "drug-tolerant persisters" go on to develop irreversible tolerance via several mechanisms, including genetic changes [63, 64] and non-genetic changes [62]. In melanoma, "sporadic [high] expression" of resistance genes enables rare cells to survive a single treatment application [65]. Without further treatment, these cells repopulate the tumor with a similar proportion of tolerant and non-tolerant cells as the treatment-naive tumor. However, with continuous treatment, the tolerant cells develop a stably changed transcriptome and repopulate the tumor [65]. In the epigenetic landscape, this may indicate that treatment lowers a formerly impossibly high barrier to a height that cells may now overcome. The rate of transitions across this barrier into a new basin will be low; thus, over short treatment time, a transition will not occur. Stopping treatment will again raise the barrier (**Figure 2.1D**). With continuous treatment, transition(s) will occur, and cells populate the new, drug-tolerant basin. Treatment removal at this time will raise the barrier, and cells will remain in the new basin (**Figure 2.1E**).

2.5 Response to treatment: transdifferentiation

In contrast to the former studies, where cells achieved stemness properties or changed behaviors due to treatment, some tumor types have been shown to undergo phenotypic transitions across lineages, or transdifferentiation. As in [51, 52, 53, 10], these transitions between lineages are yet more cases of transitions between basins in an epigenetic landscape. While small cell lung cancer (SCLC) cells undergo phenotypic transitions at baseline, treatment increases transitions from neuroendocrine to non-neuroendocrine phenotypes [53, 10]. Chemotherapy in mice leads to an increased proportion of non-neuroendocrine cells, suggesting that the SCLC phenotypic landscape shifts toward these cells under stress [10]. In oral squamous cell carcinoma, cisplatin causes epithelial cells to transition into mesenchymal cells, inducing cisplatin resistance [66]. Epithelial cells express SOX2 before their transition, but in the mesenchymal state they lose SOX2 and gain SOX9 expression [66]. Advanced prostate adenocarcinoma cells with TP53 and Rb1 loss have been shown to

transdifferentiate from a luminal cell phenotype into a neuroendocrine phenotype upon anti-androgen treatment [67]. Loss of these proteins activates SOX2 and thereby increases the lineage plasticity of tumor cells with this alteration [68]. This plasticity enables cells to transdifferentiate into a neuroendocrine phenotype when challenged with antiandrogen treatment [68, 69].

Studies are still revealing the processes driving these transitions. Prostate adenocarcinoma cells transition through a stem-like state between the luminal phenotype and the neuroendocrine phenotype, with increased expression of SOX9. These neural stem-like cells are enriched at nadir in prostate cancer xenografts treated with androgen deprivation, indicating that cells such as these eventually repopulate the tumor [70]. In melanoma, long-term BRAF inhibition causes cells to transition from melanocytic differentiation, through a slow-cycling state with neural crest-like gene expression, to a more proliferative phenotype with mesenchymal gene expression [19]. The transitions between lineages in these examples have the common feature of transitioning to a stem-like state, and then to a new, drug-tolerant phenotype. Thus, the epigenetic landscape may be arranged in such a way that the stem-like basin is between the basins of different lineages.

2.6 Modeling phenotypic plasticity

The complex behaviors described in this review are best understood via mathematical modeling. Modeling enables researchers to synthesize observed results and behaviors, and probe the system to confirm knowledge or make additional predictions. Several types of models have been used to represent cellular dynamics in cancer, including differential equations [71], Boolean modeling [72], agent-based modeling [73], stochastic modeling [71], and combinations of these as multi-scale models [74].

Mathematical models improve characterization of the tumor system and confirm understanding of the processes that occur. Several studies compared mathematical models of tumor cell populations with and without transitions from non-stem cells to CSCs. Only models with this transition toward CSCs match observed tumor population dynamics [75, 14]. Zhou et al. [76] further characterized the transition toward CSCs: their model exhibits a negative correlation between self-renewal of CSCs and dedifferentiation of non-stem cells, suggesting a "balancing mechanism" in maintaining equilibrium [76]. This is in agreement with the hypothesis that feedback mechanisms play a role in CSC subpopulations returning to steady state [16, 50].

Modeling enables calculation of relative features of the epigenetic landscape itself. Paudel et al. [20] characterize "basins" in the melanoma epigenetic landscape and the likelihood of transitions between states. Such a system could model treatments that change the landscape favorably, for example those that cause transitions into a single, treatment-sensitive state [20].

Models are especially useful in predicting cell behavior, either to tailor experimental design or lead to additional investigations. Gupta et al. [13] used mathematical modeling and experiments to characterize

phenotypic transitions in breast cancer cells. Their model predicted that differentiated cell types, luminal and basal, were able to transition to stem cell types, which they confirmed *in vivo* [13]. This prediction and subsequent validation altered the fundamental understanding of cancer stem cells. Goldman et al. [18] developed a model representing their observation that in breast cancer, chemotherapy induces a phenotypic transition to a therapy-tolerant state that is not stem-like. To confirm model predictions, depleting CSCs with a small molecule inhibitor had no effect on the development of drug-tolerant cells when treatment was applied [18]. Both of these studies have important implications for clinical practice; CSC-targeted agents are currently in trials, [39] yet these models and validating experiments indicate that such treatment would be ineffective.

Mathematical models can also be used to predict treatment response. Chapman et al. [17] developed a computational model of transitions between triple-negative breast cancer phenotypes based on time-course measurements with and without treatment. The model predicted which phenotypic transitions occur given particular treatments, and which states will be present in the tumor after drug. The knowledge gained from this model could be used to design treatment strategies: the first treatment can induce cells to transition into a particular phenotype, then a second can be applied to effectively manage tumor growth [17]. This will be extremely important as treatments are developed to prevent tumor recurrence: these treatments must either prevent phenotypic plasticity and prevent transitions, or make use of transitions as a strategy.

2.7 The epigenetic landscape could be used to evaluate treatment options

The epigenetic landscape enables researchers to envision heterogeneity and drug tolerance outside of the Darwinian selection paradigm, where pre-existing genetic mutations enable subpopulations to survive and then expand. Nongenetic means enable small tumor subpopulations to tolerate drug, and thereby survive for an extended time; eventually, enough time and divisions pass for these cells to undergo a genetic mutation towards drug tolerance. Treating a tumor based on its epigenetic landscape can target tumor subpopulations before they develop genetic resistance. However, the epigenetic landscape remains an underused conceptual paradigm. Transitions between basins are based on quasi-potential energy, which is calculated on a two-dimensional landscape (as in Fig 1). In terms of gene expression, the state space where cells exist is 20,000-dimensions (one axis per gene), and information is lost as these dimensions are reduced into two. Due to lost information, quasi-potential energy cannot be measured directly and as such a landscape from one study cannot be compared to another. It is currently unknown how all molecular networks in the cell make different phenotypes available, and thus the gene expression underlying each possible basin is unknown. This is a feature of the current state of research into cellular signaling pathways and molecular interactions: investigations in this area are active and ongoing. Further development in both areas can bring the epigenetic

landscape into use with regard to treatment decisions. More optimal means of dimensionality reduction could enable generation of more robust landscapes with potentially universal quasi-potential energy measurements that can be compared across studies. Additional characterization of the molecular networks available to cells as they develop drug tolerance will give information about transcriptional activity underlying each basin. Together, these will enable us to target the landscape features directly in order to manage tumors.

CHAPTER 3

Modeling heterogeneous tumor growth dynamics and cell–cell interactions at single-cell and cell-population resolution [1]

3.1 Introduction

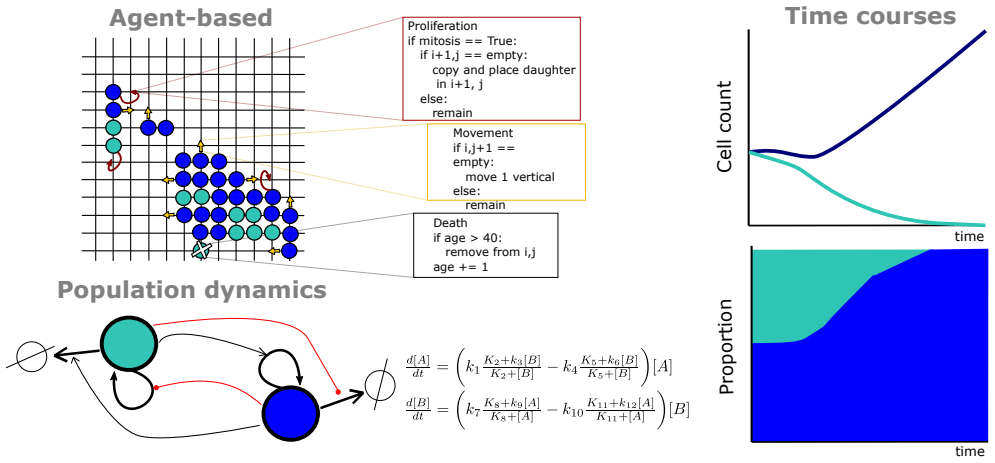


Figure 3.1: Mathematical models of tumor growth dynamics

Tumors are comprised of a heterogeneous mix of cell subpopulations with common traits and heritable phenotypes [77, 5] (e.g. morphological, proliferative, tumorigenic, transcriptional) that compete for resources and interact with each other and with noncancer cells in complex ways [3, 4]. Tumor heterogeneity has been implicated as a source of therapeutic resistance and treatment failure and is seen both across tumors (among patients and between primary and metastatic sites within patients) and within individual tumors [77, 5]. Factors contributing to intertumoral heterogeneity include genomic and epigenomic alterations, intrinsic differences between individuals, cell type of origin, and microenvironmental differences across anatomical (primary and secondary) sites. Intratumoral heterogeneity (**Figure 3.2**) is driven by a complex mix of genomic instability, cellular plasticity, regional differences in the microenvironment (e.g. oxygen availability), cell-cell interactions, and intrinsic fluctuations in cell fate and gene expression.

The complex nature of tumor heterogeneity can be understood within the framework of 'epigenetic landscapes,' first proposed by Waddington [33] as a conceptual tool for understanding cellular differentiation during development. More recently, it has been applied to tumors, where 'stem-like' cells have the capacity to differentiate into multiple phenotypes and seed tumor growth. Viewing tumor heterogeneity through the

[1] This chapter is adapted from 'Modeling heterogeneous tumor growth dynamics and cell–cell interactions at single-cell and cell-population resolution' published in *Current Opinion in Systems Biology* and has been reproduced with the permission of the publisher and my co-authors, Leonard A. Harris, Patricia M. M. Ozawa, Lizandra Jimenez, and Alissa M. Weaver.

lens of epigenetic landscapes may serve as a means for better understanding tumor dynamics and therapy response [5]. Borrowing ideas from physical chemistry, an epigenetic landscape is a quasi-potential energy surface where local minima, or 'basins of attraction,' correspond to cellular phenotypes [78, 79]. The genetic state of a cell sets the topography of the landscape, and genetic mutations can modify it, for example, by changing the depths of the basins (**Figure 3.2**). Cells can transition among these phenotypes at rates dependent on the heights of the barriers separating basins (**Figure 3.2**). Fluctuations in intrinsic (e.g. gene expression [80]) and/ or extrinsic (e.g. secreted factors [81], intracellular protein concentrations [82], oxygen production [83]) processes may serve as sources of noise that drive these transitions. This combination of epigenetic landscape heterogeneity and intrinsic/extrinsic stochasticity, together with genomic diversity (either acquired during normal tumor progression or over the course of therapy), results in a highly heterogeneous tumor (**Figure 3.2**). This view of tumor heterogeneity is consistent with both known genetic clonality of tumors [84] and a growing body of literature on nongenetic (epigenetic) inheritance in cancer cells. The latter includes the cancer stem cell (CSC) hypothesis, which posits that phenotypically distinct tumorigenic subpopulations sit atop a hierarchy of nontumorigenic progeny [85], as well as reports of stem-like cancer cells that are not organized in a classical hierarchy but are metastable and can reversibly shift phenotypes [13, 53]. Notably, stem cells are known to function within niches, and interactions between stem-like and niche-like cells may govern tumor dynamics in an analogous manner to their tissue of origin. These cell-cell interactions may also influence switching between epigenetic basins in response to tumor therapy.

In light of this confluence of genetic, epigenetic, and stochastic factors that underlie tumor heterogeneity, it is no surprise that durable anticancer therapies remain elusive. To tackle this complexity, a systems biology approach is required and has been steadily growing in popularity [86]. By combining high-throughput experimentation, 'Big Data' analysis, and mathematical modeling with *in vitro* and *in vivo* validation experiments, cancer systems biologists aim to disentangle the complex web of interactions, feedbacks, and dysregulated control mechanisms that has stymied clinical progress [77, 5]. Mathematical models are essential to this endeavor because they can predict behaviors outside the range of the experimental conditions on which they are based and can explore the effects of drug dosing, scheduling, and changes in the microenvironment on tumor response [87]. Moreover, mathematical models can explicitly account for both intertumoral and intratumoral heterogeneity by varying model structure and parameter values. Intertumoral heterogeneity, for example, can be captured by creating variants of a model with different cellular and/or molecular species or by varying the values of rate parameters to account for genetic or microenvironmental differences across individuals or disease sites. Intratumoral heterogeneity can be accounted for by including multiple cell subtypes, interactions through physical contact or secreted factors, and by including spatial inhomogeneity of biological factors, such as oxygen, over the computational domain. In the following sections, we provide a brief review of

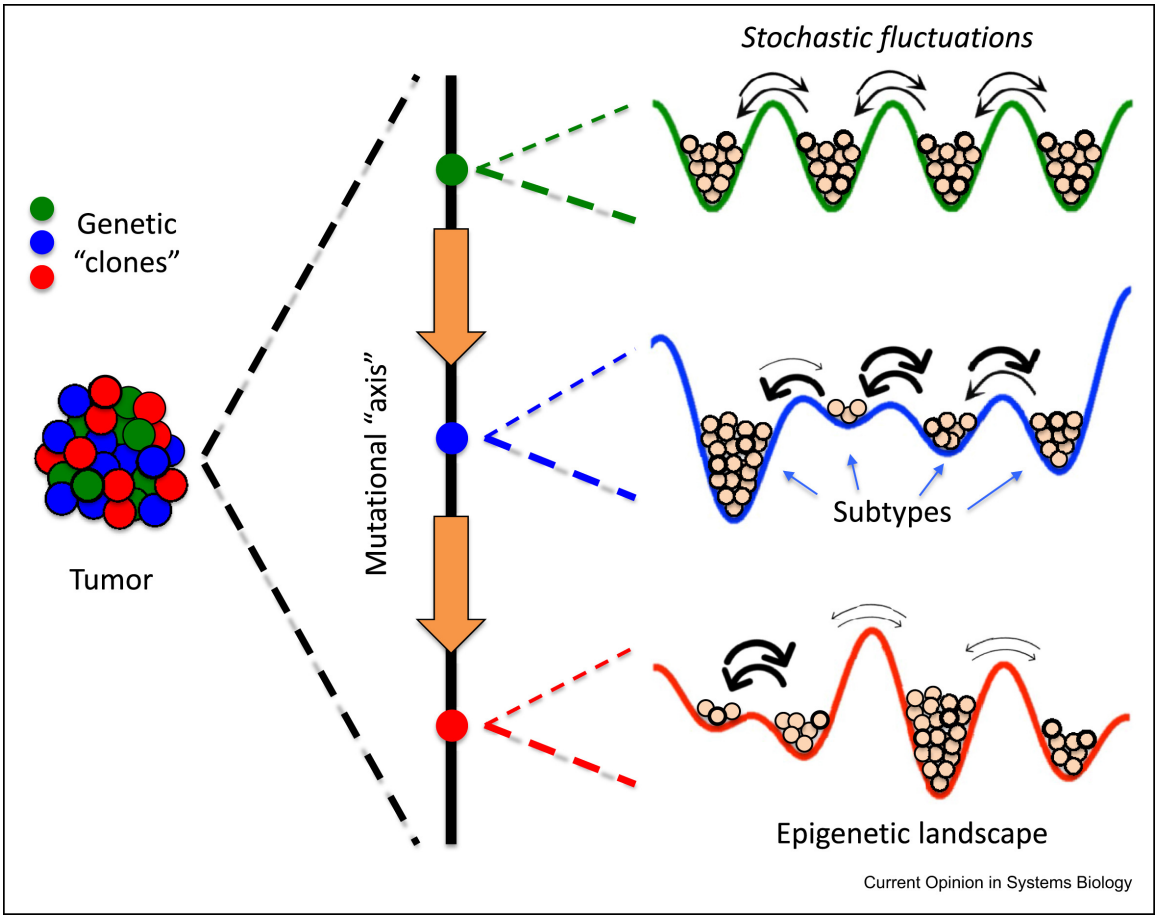


Figure 3.2: Three sources of intratumoral heterogeneity

Genetic clones are cells with common origin but genetically diversified by mutations that either pre-exist or are acquired during the course of therapy [77, 5]. Each genetic clone can be envisioned as existing along a 'mutational axis' and having an associated 'epigenetic landscape,' i.e. a quasi-potential energy surface where local minima, or 'basins of attraction,' correspond to cellular subtypes [33]. From a molecular perspective, the epigenetic landscape is the consequence of the complex biochemical interaction networks that underlie cell fate decisions [78, 79]. Thus, gene expression noise [80] and other sources of intrinsic (e.g. fluctuations in the production and contents of secreted factors) and extrinsic stochasticity can drive transitions between subtypes (thick arrows represent fast transitions, and vice versa). Altogether, at any point in time, the subtype composition of a tumor will depend on the genetic clones present within the tumor, the topographies (depths of basins and heights of barriers) of the associated epigenetic landscapes, and the magnitudes of intracellular fluctuations within individual cells.

recent approaches for modeling tumor heterogeneity, followed by a short example from our own work.

3.2 Mathematical formalisms for modeling tumor growth

3.2.1 Agent-based models

At the resolution of single cells, agent-based models (ABMs) aim to capture complex and emergent behaviors of cell populations through simple and intuitive rules [88]. ABMs are particularly useful when detailed mechanisms of cell behavior are not known but there is a qualitative understanding of the conditions under which certain cellular behaviors are observed. The simplest type of ABM is the cellular automaton [73], in

which cells are organized in a regular grid and can take on one of a finite number of discrete states. Cell states change based on user-defined rules that account for the current state value of a cell and those of its neighbors. These updates can be performed synchronously or asynchronously (i.e. stochastically). Modern ABMs extend the basic cellular automaton by permitting cells to have multiple properties (e.g. age, phenotypic state, mutation), with discrete or continuous state values, and allowing rules to be either deterministic (occurring if a certain condition is met) or probabilistic (occurring with some probability if the condition is met) [89]. An example of a rule governing cell death based on the age of a cell and its phenotypic state is as follows (ccl.northwestern.edu/netlogo/models/Tumor).

```
if (not stem?) and (not metastatic?) and (age > 20)
    [ die ]
if (not stem?) and metastatic? and (age > 4)
    [ die ]
```

ABMs can also be on-lattice (i.e., grid-based) or off-lattice. On-lattice models are often used to model cell movement within a tumor microenvironment [90, 75, 91, 92, 93, 94]. Off-lattice models can be beneficial for modeling cells as clusters or functional units, such as colon crypts [73, 95, 96, 97]. ABMs can also be nonspatial, where the cell state is informed by its environment and the states of other cells but not by location or distance from other cells [98, 99, 100]. ABMs can be implemented in custom computer code or using one of numerous publicly available software packages [101, 102, 103, 104].

Examples of ABMs applied in cancer include that of Enderling and Hahnfeldt [90], who constructed an on-lattice ABM of CSCs and non-CSCs, including crowding effects, and predicted that low proliferative capacity and a high rate of spontaneous cell death – which might be expected to keep tumor growth and metastasis in check – actually leads to a less densely seeded tumor where CSCs have the physical space to divide, produce progeny, and eventually escape and seed metastasis. Stichel et al. [95] developed an off-lattice ABM that includes displacement of cells in response to external forces from neighboring cells (e.g. repulsive forces at close cell-cell distances) and a vector-valued velocity term based on these forces. The model was able to reproduce diverse *in vitro* cell speed dynamics and migratory behaviors, such as straight cell fronts and the formation of cellular bridges.

3.2.2 Population dynamics models

In contrast to ABMs, population dynamics models do not track individual cells over time but rather the size of the cell population [71]. The classic model of population growth is as follows:

$$Cell \xrightarrow{k} 2Cell \tag{3.1}$$

which, if one assumes continuous and deterministic dynamics, gives the exponential growth equation:

$$N(t) = N(0)e^{kt} \tag{3.2}$$

where $N(t)$ is the cell count at time t and the growth rate constant k can be positive or negative. Since its inception, numerous extensions to the exponential growth model have been proposed, including adding an upper limit on the population size ('carrying capacity') to account for nutrient consumption (e.g. logistic and Monod kinetics) [105] and reducing growth rates at small population sizes to account for the positive effects of cell-cell interactions on growth (e.g. the Allee effect) [106]. Models explicitly accounting for progression through the cell cycle, the apoptotic program, and entry into quiescence/senescence have also been developed [107, 108]. Spatial heterogeneity can be accounted for by including cell diffusion and migration over either a continuous or discrete domain [109], and the inherent randomness of cell fate decisions can be modeled using stochastic methods, such as the chemical Master Equation or the Gillespie algorithm [110]. A variety of tools are available for constructing and simulating population dynamics models, including environments such as Matlab and Python and numerous general-purpose software packages [111, 112, 113, 114, 115, 116, 117].

In cancer, heterogeneous tumor growth and drug-response dynamics can be modeled by including multiple cell subpopulations that compete and interact [118]: Clonal competition can be modeled by varying growth rates across subpopulations [119, 120, 20]; genetic mutations can be included by allowing for spontaneous creation of new subpopulations with prespecified growth rates [121]; epigenetic states can be accounted for by allowing cells to reversibly transition between subpopulations [13, 75, 99, 119, 20, 122, 123]; and cell-cell communication can be incorporated by including density-dependent growth rates [124, 125, 126, 127] or by modeling (explicitly or implicitly) the secretion of factors that modify the growth dynamics of other cells [128, 129, 130, 131, 132]. As an example, Gupta et al. [13] constructed a population dynamics model describing phenotypic transitions among luminal, basal, and stem-like subpopulations in two breast cancer cell lines and recapitulated the experimental observation that isolated subpopulations recover the cell-state proportions of the parental cell line over time. Based on this, they predicted that after cessation of CSC-targeted therapies, surviving non-CSCs would transition back into CSCs and restart tumor growth, thereby reversing any therapeutic gains. In a subsequent study, Zhou et al. [119] extended this model to include subpopulation-specific growth rates and derived analytical expressions for the temporal dynamics of cell-state proportions as well as the conditions necessary for distinct subpopulations to coexist. In the following section (**Section 3.3**), we present a similar model of subpopulation dynamics in tumors with the added feature of subtype interactions

mediated by secreted factors.

3.2.3 Multiscale models

Biological processes span multiple scales, from genes to proteins to cells to tissues to organisms [74]. However, most mathematical models of biological processes operate at only a single scale and either ignore processes at other scales or try to incorporate their effects in some approximate way. Multiscale models attempt to address this shortcoming by explicitly modeling processes at different spatial and/or temporal scales and connecting these models through some type of information-exchange interface [74, 133, 134, 135, 136]. In multiscale tumor models, the 'cell-level' model can either be an ABM or a population dynamics model, which can be coupled to, for example, a reaction-diffusion model governing background concentrations of bioactive species (e.g. oxygen, nutrients, extracellular matrix) [137, 138, 139, 140, 141, 142, 143] or a kinetic or Boolean model describing the intracellular biochemical networks driving cell fate (e.g. progression through the cell cycle, apoptosis) [73, 144]. Typically, simulations are performed by running a short simulation (possibly a single step) of a model at one scale, passing information generated from that simulation to another model at a different scale, running a short simulation of that model with updated initial and/or boundary conditions, and repeating. While in the majority of cases multiscale models are constructed by composing single-scale models in well-established modeling tools and linking them together with specialty purpose code [145], there is a rapidly growing list of general-purpose software packages for multiscale modeling [146, 147, 148, 149, 150, 151]. These represent a community-wide effort to develop easy-to-use platforms that facilitate calibration, validation, reproducibility, and exchange of multiscale models using methods firmly rooted in a mathematical formalism [145]. As an example, Yan et al. [152] used a multiscale hybrid discrete-continuum model to investigate vascular tumor signaling in glioblastoma (GBM). The continuum (spatial population dynamics) model includes various GBM cells (stem, progenitor, terminally differentiated, dead), host tissue, and human primary endothelial cells, as well as nutrients and signaling molecules. Angiogenesis is modeled via a discrete, off-lattice ABM: Blood vessels sprout, grow, and branch at rates dependent on tumor density and background concentrations of vascular endothelial growth factor (VEGF) and other signaling molecules. The vasculature also produces a factor that maintains the pool of GBM stem cells by increasing their proliferation rate and self-renewal capability. Central to the model is a positive feedback loop whereby nutrient-poor tumor cells produce VEGF, inducing vasculature growth, which increases GBM stem cell growth and leads to even more VEGF production. Using the model, the authors confirmed the experimental observation that antiangiogenesis treatments have a negative effect of inducing the growth of invasive 'fingers' that eventually develop into multifocal tumors. Blocking vasculature-secreted factors, however, disrupts growth and reduces tumor size without increasing invasiveness, as also seen experimentally

[153, 154].

3.3 Example: modeling subtype interactions in small-cell lung cancer

Small-cell lung cancer (SCLC) is an aggressive and highly metastatic neuroendocrine (NE) carcinoma [155, 9] for which the standard of care (etoposide + cisplatinum and radiation), until very recently [156], had not changed in over 30 years. SCLC is characterized by loss-of-function mutations in TP53 and RB1 and is often described as a homogeneous-looking cancer, with 'small blue round cells' and minimal presence of nontumor cells. Indeed, genomic analyses indicate that, on average, $\sim 84\%$ of cells in primary SCLC tumors are cancer cells [8]. However, over time, it has become clear that SCLC tumors are actually highly heterogeneous, with various cell subtypes coexisting, some of which may provide the trophic support that noncancer cells often provide in other cancer types [10, 157].

Owing to the challenges of acquiring human SCLC samples, heterogeneity in SCLC has been primarily studied in animal models [158]. A study using a genetically engineered mouse model of SCLC, with induced loss of p53, Rb, and p130 tumor suppressor genes in lung epithelial cells, demonstrated that the most abundant tumor cells in that model are NE self-renewing 'tumor-propagating cells' (TPCs), characterized by $CD24^{high}$, $CD44^{low}$, and $EpCAM^{high}$ surface expression [53]. A non-NE population with high expression of Hes1 and lacking NE markers but active in the Notch signaling pathway was also identified [10]. This subtype was shown to have increased resistance to chemotherapy in comparison with NE cells (TPCs), suggesting a possible role in chemoresistance. A third population, also non-NE, with high levels of CD44 and characterized by expression of mesenchymal markers, such as vimentin, has also been reported [159]. Furthermore, *in vitro* coculture experiments of NE and $Hes1^+$ non-NE cells show that NE cell growth is enhanced in the presence of non-NE cells [10, 159]. *In vivo* experiments have been reported showing that mice injected with an admixture of NE and $CD44^+$ non-NE cells develop metastases but mice injected with only a single subpopulation do not [159]. Taken together, these data suggest that TPCs, $Hes1^+$, and $CD44^+$ cell subtypes form an SCLC tumor ecosystem that supports growth, metastasis, and treatment evasion.

Improving our understanding of the dynamics of SCLC tumors and the interactions between cell subtypes through mathematical modeling could lead to improved treatment options in the future [160]. Therefore, we present here a population dynamics model of the SCLC tumor ecosystem as an illustrative example of the model building and analysis process. We begin by enumerating a list of biological rules (**Table 3.1**) that encompass, to the best of our ability, all that is presently known about SCLC tumor heterogeneity and subtype interactions. Notably, these rules, which form the basis of our model, are derived from the experimental results described previously [53, 10, 159] and are primarily qualitative in nature. This is an important point as most biological data, especially in the early stages of a research project, are not quantitative [161]. As we

	Description and reference(s)
1	Tumor-propagating cells (TPCs) divide approximately once per day in culture [53].
2	TPCs have a 20–30% apoptosis rate in culture and a 0–5% apoptosis rate *in vivo* [53].
3	TPCs can differentiate into other cell types, e.g. $Hes1^+$/Notch-active cells and $CD44^+$ cells (and possibly other subtypes as well) [10].
4	$Hes1^+$ cells have a lower division rate than TPCs [10].
5	$Hes1^+$ cells support the growth of the TPC population by both enhancing cell division and reducing apoptosis [10].
6	There is evidence that TPCs inhibit proliferation of $Hes1^+$ populations *in vitro* (30–50% reduction as assessed by BrdU incorporation) [10].
7	Role of $CD44^+$ cells is unclear, but they may be prometastatic and promote invasion by secretion of fibroblast growth factor, which activates ERK1/2 signaling in TPCs [8, 159].
8	*In vivo* tumors are comprised of approximately 50% TPCs, 25% $Hes1^+$ cells, 5% $CD44^+$ cells, and 20% unknown cell types [53, 10].
9	Interactions among cell types may involve secretion of soluble factors [10, 159].

Table 3.1: Biological rules used to construct the small-cell lung cancer population dynamics model.

shall show, it is nevertheless possible to construct a predictive mathematical model based on these data that can provide valuable insight and guide future, more quantitative experiments. These follow-up experiments can, in turn, provide the necessary data to perform more advanced modeling tasks, such as parameter estimation and identifiability analysis [162], which can significantly improve the scope and quality of the model predictions. This iterative process of model construction, experimentation, and refinement lies at the heart of the systems biology approach [86, 87].

With these biological rules in hand (**Table 3.1**), we next synthesize this knowledge into a simple diagram (**Figure 3.3A**) that illustrates the basic biological processes we aim to model, namely, division, death, differentiation, and inhibition/enhancement of division and death via cell-cell interactions. We then translate this diagram into mathematical form by writing down eight kinetic equations representing division and death of each subtype and differentiation of TPCs into $Hes1^+$ and $CD44^+$ subtypes (**Table 3.2**). For five of these processes ($Hes1^+$ death, $CD44^+$ division/death, TPC differentiation), we currently have no information regarding any potential mediating factors (**Table 3.1**). Therefore, we make the simplest possible assumption that the rates of these processes are directly proportional to the population sizes of the subtypes (known as the 'law of mass action'). For the other three processes ($Hes1^+$ division, TPC division/death), experimental data suggest that the rates are dependent on cell-cell interactions (**Table 3.1**). We assume these interactions to be mediated by secreted factors according to the following simple reaction motif,

$$
\begin{array}{lll}
A \xrightarrow{k_x} A + x & B + x \xrightarrow{k_f} B^* + x & B \xrightarrow{k_{fate}} 2B \text{ or } \varnothing \\
x \xrightarrow{k_{-x}} \varnothing & B^* \xrightarrow{k_r} B & B^* \xrightarrow{k_{fate*}} 2B^* \text{ or } \varnothing
\end{array}
\tag{3.3}
$$

where A and B are cell types, A secretes a factor x that interacts with B and converts it into an 'active' form B^* that has an increased or decreased rate of division or death, and $\varnothing$ represents cell death. Assuming that the concentration of x is stable, that is, its rates of production and degradation are equal (known as the 'quasi-steady-state assumption') and that the rates of interconversion between B and B^* are equal (known as the 'partial equilibrium assumption'), we can derive an analytical expression for the rate of the cell fate (division or death) of cell type B as a function of the population size of cell type A,

$$
\nu_{fate} = \left(\frac{k_{fate} K_D K_x^{eq} + k_{fate}^* [A]}{K_D K_x^{eq} + [A]} \right) [B]_T \tag{3.4}
$$

where $K_D \equiv \frac{k_r}{k_f}$, $K_x^{eq} \equiv \frac{k_{-x}}{k_x}$, and $[B]_T \equiv [B] + [B^*]$ is the total amount of B. Note that Eq. (3.4) models an increase in the rate of division or death by the diffusible factor x if $k_{fate}^* > k_{fate}$ and reduction in the rate if $k_{fate}^* < k_{fate}$. If $k_{fate}^* = k_{fate}$, then $\nu_{fate} > k_{fate}[B]_T$ (mass-action kinetics). Reactions 1-3 in **Table 3.2** have rate expressions of this form, with the parameters k_{fate}, k_{fate}^*, and $K_D K_x^{eq}$ written as numbered parameters k_i or K_i, and $[B]_T = [TPC]$ *or* $[Hes1^+]$.

Index	Reaction	Rate expression	Description
1	$TPC \rightarrow 2TPC$	$\left(\frac{k_1 K_2 + k_3 [Hes1^+]}{K_2 + [Hes1^+]}\right)[TPC]$	TPC cell division ($k_3 > k_1$) *($Hes1^+$ promotes growth)*
2	$TPC \rightarrow \varnothing$	$\left(\frac{k_4 K_5 + k_6 [Hes1^+]}{K_5 + [Hes1^+]}\right)[TPC]$	TPC cell death ($k_6 < k_4$) *($Hes1^+$ inhibits apoptosis)*
3	$Hes1^+ \rightarrow 2Hes1^+$	$\left(\frac{k_7 K_8 + k_9 [TPC]}{K_8 + [TPC]}\right)[Hes1^+]$	$Hes1^+$ cell division ($k_9 < k_7$) *(TPC inhibits growth)*
4	$Hes1^+ \rightarrow \varnothing$	$k_{10}[Hes1^+]$	$Hes1^+$ cell death
5	$CD44^+ \rightarrow 2CD44^+$	$k_{11}[CD44^+]$	$CD44^+$ cell division
6	$CD44^+ \rightarrow \varnothing$	$k_{12}[CD44^+]$	$CD44^+$ cell death
7	$TPC \rightarrow Hes1^+$	$k_{13}[TPC]$	TPC differentiation into $Hes1^+$
8	$TPC \rightarrow CD44^+$	$k_{14}[TPC]$	TPC differentiation into $CD44^+$

Table 3.2: Kinetic formulation of the small-cell lung cancer population dynamics model.
TPC, tumor-propagating cell.
$[X]$ is the population of cell type X; $\varnothing$ represents cell death.

It is important to emphasize that the model in **Table 3.2** is certainly not the only possible mathematical interpretation of the diagram in **Figure 3.3A**. It is simply our first attempt at codifying our knowledge of SCLC subtype interactions and dynamics in mathematical form. Over time, we expect to include

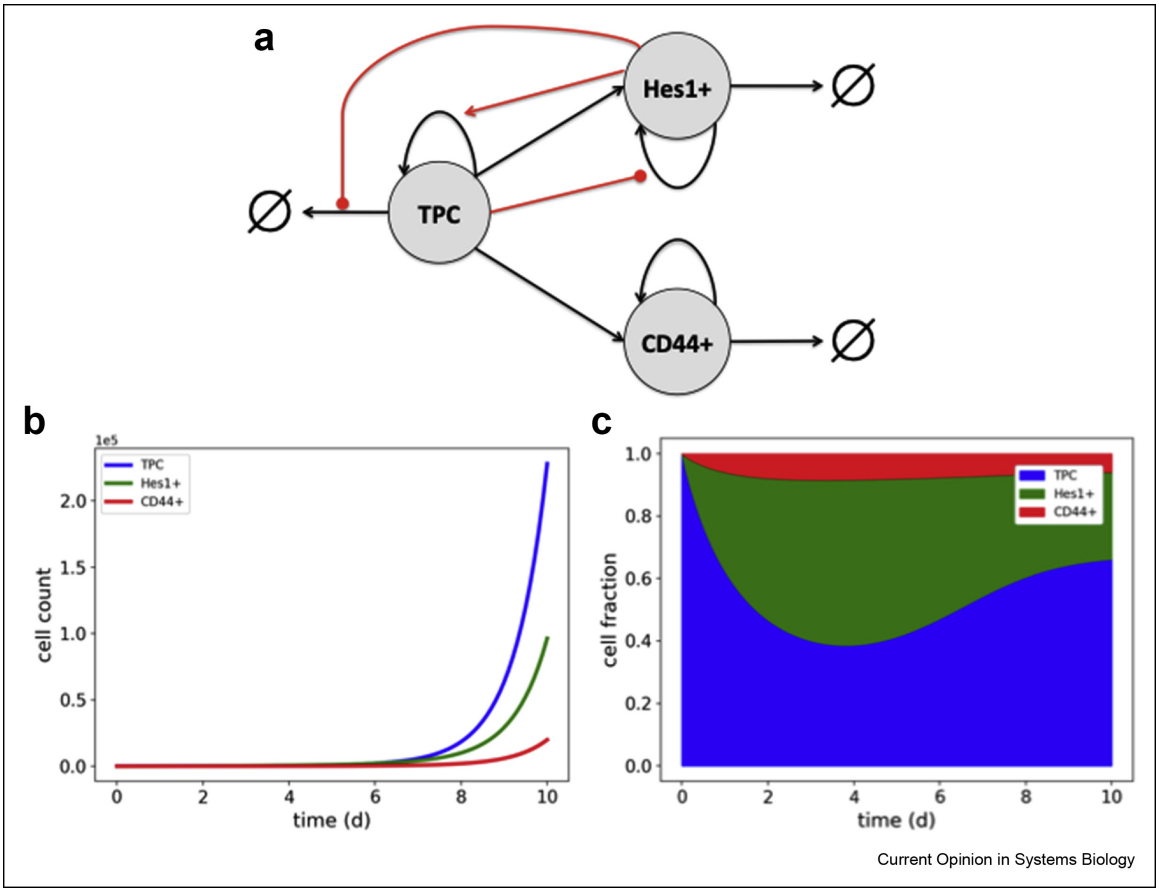


Figure 3.3: Population dynamics model of subtype interactions in small-cell lung cancer.

A Schematic representation. Black arrows represent cell fates (division, death, differentiation); red arrows/circles represent enhancement/inhibition of cell fate by secreted factors; ∅ represents cell death.

B Simulated time courses of cell counts,

C and fractions.

Ordinary differential equation simulations were performed in PySB [116] with initial conditions $[TPC]_0$ = 100, $[Hes1^+]_0 = 0$, $[CD44^+]_0 = 0$ and the following parameter values (**Table 3.2**): $k_1 = 1$, $K_2 = 1000$, $k_3 = 2$, $k_4 = 0.2$, $K_5 = 1000$, $k_6 = 0.1$, $k_7 = 0.5$, $K_8 = 1000$, $k_9 = 0.25$, $k_10 = 0.2$, $k_11 = 0.3$, $k_12 = 0.2$, $k_13 = 0.5$, $k_14 = 0.1$. These values were chosen to produce subtype proportions similar to those seen *in vivo* (**Table 3.2**, rule 8; excluding the 20% unknown cell types).

additional interactions, cell types, and processes as experimental observations that cannot be captured by this preliminary model are made. A common approach when the details of a biological process are not well known is to construct numerous 'candidate' models and then systematically eliminate unlikely candidates using various 'model selection' metrics, such as Akaike or Bayesian information criteria [163]. Thus, rather than a single model, we expect to construct a compendium of SCLC tumor models that will grow and change over time. Numerous software tools and environments exist to facilitate such efforts [116, 164, 165, 166, 167]. One such tool is PySB [116], a 'rule-based' modeling platform [168, 169] writ-

ten in Python that includes an easy-to-use modeling language and a variety of deterministic and stochastic simulators [113, 117, 170, 171] and analysis tools [172]. We constructed the model in **Table 3.2** in PySB (available at github.com/lh64/CurrOpinSysBiol_2019) and used it to automatically generate the following set of coupled ordinary differential equations,

$$\frac{d[TPC]}{dt} = \left(\frac{k_1 K_2 + k_3 [Hes1^+]}{K_2 + [Hes1^+]} - \frac{k_4 K_5 + k_6 [Hes1^+]}{K_5 + [Hes1^+]} - k_{13} - k_{14} \right) [TPC] \tag{3.5}$$

$$\frac{d[Hes1^+]}{dt} = \left(\frac{k_7 K_8 + k_9 [Hes1^+]}{K_8 + [Hes1^+]} - k_{10} \right) [Hes1^+] + k_{13} [TPC] \tag{3.6}$$

$$\frac{d[CD44^+]}{dt} = (k_{11} - k_{12}) [CD44^+] + k_{14} [TPC] \tag{3.7}$$

These equations were then numerically integrated using algorithms available in the Python package `SciPy` [170].

As our experimental data are qualitative (**Table 3.1**), for a preliminary simulation analysis (**Figures 3.3B** and **3.3C**), we decided to *manually* choose parameter values using the experimental observations as a guide (see caption to Figure 2 for values). For example, the division rate for TPCs unaffected by secreted factors (k_1 in **Table 3.2**) was set to 1/day, consistent with experimental observations (**Table 3.1**, rule 1). The division rate for unaffected Hes1$^+$ cells (k_7 in **Table 3.2**) was then set to 0.5/ day based on the observation that they divide at a lower rate than TPCs (**Table 3.1**, rule 4). Initial values for all other rate parameters were chosen based on similar reasoning. All parameter values were then manually varied until values were found that produced steady-state subtype proportions similar to those seen in *in vivo* mouse tumors (**Table 3.1**, rule 8; **Figure 3.3C**). Of note is that the rate of differentiation from TPCs to Hes1$^+$ cells (k_{13} in **Table 3.2**) had to be set to a value five times larger than that for TPCs to CD44$^+$ cells (k_{14} in **Table 3.2**) because modulating the division and death rates of the subtypes alone could not achieve the desired subtype proportions.

This result is a simple illustration of the nontrivial relationships that often exist between parameter values and model outputs in systems biology models [173]. An obvious next step in the model analysis process would be to quantify these relationships through a 'sensitivity analysis' [174] that would identify parameters (or sets of parameters) that act as 'control points' for modulating model behavior. These could then be tested experimentally to either validate or refute the current version of the model and suggest possible refinements. As mentioned previously, these validation experiments may be quantitative. In such cases, one could then move beyond sensitivity analysis and perform a formal 'parameter estimation' [175], which would allow for quantitative model predictions that could be further tested experimentally. Modern param-

eter estimation algorithms utilize Monte Carlo (or similar) techniques to generate 'ensembles' of parameter sets that provide similar fits to experimental data [172, 173, 176]. Model predictions are then generated by sampling over the ensemble, producing a predictive 'envelope' [20] that hopefully is not excessively broad (if it is, further experiments could be designed to reduce the breadth of the envelope [177]). We do not go into this level of sophistication here as our goal is simply to demonstrate how a preliminary model can be built and analyzed, but we will use these techniques in the future as the model grows in complexity and additional experimental data are collected.

Finally, even though our model is based on qualitative observations and static measurements of the compositions of *in vivo* mouse tumors, the population dynamics formalism that we use allows us to explore and speculate on the temporal dynamics of tumor growth. In **Figures 3.3B** and **3.3C**, we show time courses for a 10-day *in silico* experiment starting with a pure population of 100 TPCs. Interestingly, we see that while the total cell count grows continuously throughout the simulation (because the current model lacks a carrying capacity; **Figure 3.3B**), the cell-state *proportions* stabilize around day 10 (**Figure 3.3C**). This is consistent with prior theoretical work [13, 119], which was experimentally validated [13], showing that phenotypic state transitions can lead to stable steady-state proportions in heterogeneous cell populations. However, our simulation also shows pronounced nonlinear dynamics before day 10, with the $Hes1^+$ proportion peaking around day 4 (**Figure 3.3C**). This is a direct result of modeling cell-cell interactions (which were not considered in the studies by Gupta et al. [13] and Zhou et al [119]) between the TPC and $Hes1^+$ subtypes (**Figure 3.3A** and **Table 3.2**). Thus, these nonlinear dynamics represent a prediction of the model that could be tested experimentally: if observed, it would provide support for the underlying assumptions of our model and verify the importance of cell-cell interactions in SCLC tumor dynamics.

3.4 Conclusion

Significant progress has been made in recent years in cancer therapy, including the development of new chemotherapies, targeted agents, and immunotherapies. However, cancer morbidity remains stubbornly high, in large part due to intertumoral and intratumoral heterogeneities. To overcome this challenge, an integrative systems approach involving experimentation, bioinformatics, and mathematical modeling is beginning to take root and produce results [86]. ABMs, population dynamics, and multiscale models are central to this effort as they provide a window onto the underlying mechanisms of therapy resistance [87]. Going forward, a major goal in the field will be to develop patient-specific tumor models, parameterized by clinical data and spanning genetic, cellular, and tissue scales, which can act as *in silico* platforms for testing and designing personalized anticancer therapies. This approach holds the promise of significantly improving clinical outcomes and quality of life for millions worldwide suffering from this deadly disease.

CHAPTER 4

Methods and frameworks developed for the present work

For this body of work, I developed several workflows using existing methods in novel ways or applied to novel situations. I present these methods as an individual chapter so that future readers may refer to this section for reference. The methods used are presented in chapter order.

4.1 Methods: Didactic example to contrast AIC vs posterior probability calculated by Bayes-MMI for model selection and multi-model inference

4.1.1 Estimating regression coefficients and evidence calculation by nested sampling using data from (Galipaud et al., 2014)

Galipaud and colleagues provide R code in the Supporting Information of their article for their simulations, and we used their `data.simulation` method to generate a dataset with `sample.size=100`, and using the same tapering predictor variable effects, $r_{y,x_1} = 0.70$, $r_{y,x_2} = 0.50$, $r_{y,x_3} = 0.10$, $r_{y,x_4} = 0.00$ [1]. This generates the dataset used for our analyses (see **Appendix B.1**, **Table B.1**).

For this example with simulated ground truth data, since we have generated the data using four predictor variables, we can assume that all 16 combinations of the 4 variables ($2^4 = 16$) will result in the set of 16 models representing the entire probability space that one of the models is correct (see **Chapter 5**, Eq. (5.5)).

For our regression coefficient estimation and model selection, we used the Multinest algorithm [29, 30, 31]. Multinest samples multi-dimensional parameter space, bounding its search by parameter values along each axis in each of the multiple dimensions based on prior expectation of parameters, $\Pr\left(\vec{\theta}\right)$, input by the user. It removes the lowest-probability parameter set and chooses a new one from within the bounded parameter space, subsequently re-drawing the search space with the bounds incorporating the new parameter set. This continues until all parameter sets representing the bounds of the remaining search space have approximately equal probability, and the algorithm estimates that the remaining probability of parameter sets within the bounds is less than a user-defined tolerance. Each parameter set is evaluated based on a user-defined likelihood function. Finally, the likelihood values that correspond to each sampled parameter set are arranged in the order they were replaced, and the integral over these is taken to approximate the integral over all possible models, that is, the marginal likelihood or Bayesian model evidence. We used the Multinest-returned "vanilla" nested sampling evidence value [31].

As a Bayesian method, Multinest utilizes a prior rate parameter distribution, $\Pr\left(\vec{\theta}\right)$, where $\vec{\theta}$ represents the set of n parameters $\theta_1, \theta_2, ..., \theta_n$, and a likelihood function to assess each potential parameter set

$$L(D \mid \vec{\theta}_{\mathbf{i}}) = \Pr\left(D \mid \vec{\theta}_{\mathbf{i}}\right) \tag{4.1}$$

where $\vec{\theta}_{\mathbf{i}}$ is the i^{th} parameter set and D represents the data being used for fitting. Parameter set $\vec{\theta}_{\mathbf{i}}$ is scored via the likelihood function $L(\vec{\theta}_{\mathbf{i}})$.

For the didactic example in **Chapter 5**, we use a least-squares likelihood function, evaluating the mathematical distance between the model simulation $\vec{\mathbf{y}}$ values $\vec{\mathbf{y}_{\mathbf{sim}}}$ and the simulated data $\vec{\mathbf{y}}$ values $\vec{\mathbf{y}_{\mathbf{data}}}$. Here,

$$\vec{\mathbf{y}_{\mathbf{sim}}} = (\vec{\mathbf{x}_{\mathbf{1}_{\mathbf{data}}}} * \beta_1 + \vec{\mathbf{x}_{\mathbf{2}_{\mathbf{data}}}} * \beta_2 + \vec{\mathbf{x}_{\mathbf{3}_{\mathbf{data}}}} * \beta_3 + \vec{\mathbf{x}_{\mathbf{4}_{\mathbf{data}}}} * \beta_4 + \beta_0) \tag{4.2}$$

where each β_i scalar in Eq. (4.2) is equivalent to a parameter θ_i (as in Eq. (4.1)) and represents the value in each i^{th} dimension of parameter space that Multinest is searching. For simulating y values $\vec{\mathbf{y}_{\mathbf{sim}}}$ during Multinest's parameter search, each predictor variable array $\vec{\mathbf{x}_{\mathbf{i}_{\mathbf{data}}}}$ and the response variable array $\vec{\mathbf{y}_{\mathbf{data}}}$ values in Eq. (4.2) are from the simulated data represented in part in **Table B.1**.

For the prior distribution of each parameter/regression coefficient β_i to provide to Multinest, we used uniform prior distributions from 0 to 10 for regression coefficients β_1 - β_4 and a uniform prior distribution from -10 to +10 for the intercept β_0. For the model selection problem, if a predictor variable x_i was not included in the model, its regression coefficient β_i was set to zero.

The mathematical distance calculated via least squares is

$$\frac{(\vec{\mathbf{y}_{\mathbf{data}}} - \vec{\mathbf{y}_{\mathbf{sim}}})^2}{\sigma^2_{\vec{\mathbf{y}_{\mathbf{data}}}}} \tag{4.3}$$

Multinest minimizes this distance, resulting in sets of parameters/regression coefficient values $\{\vec{\beta}_{\mathbf{1}}, \vec{\beta}_{\mathbf{2}}, \vec{\beta}_{\mathbf{3}}, \vec{\beta}_{\mathbf{4}}\}$ that result in the least distance between the model-simulated values $\vec{\mathbf{y}_{\mathbf{sim}}}$ and the data values $\vec{\mathbf{y}_{\mathbf{data}}}$. These represent the best-matching regression coefficients.

Using the Multinest algorithm results, the minimal distance, corresponding to maximum log-likelihood, can be determined out of all regression coefficient sets. For each model, Multinest also returns the Bayesian evidence (see **Chapter 5 Section 5.2** for more detail about Bayesian model evidence).

4.1.2 Calculating Akaike Information Criterion (AIC)

AIC for each of the 16 models in this example can be calculated using its maximum log-likelihood value $L(\vec{\beta_{\mathbf{jbest}}} \mid D)$ as

$$AIC_j = -2\ln\left(L(\vec{\beta_{\mathbf{jbest}}} \mid D)\right) + 2K_j \tag{4.4}$$

where $\vec{\beta}_{\mathbf{j_{best}}}$ is the best-fitting parameter set $\vec{\beta}_{\mathbf{best}}$ for model j, and K_j is the number of parameters in model j (see **Chapter 5 Section 5.3** for more detail about AIC).

As in [1], we calculated the "corrected AIC", or AICc, which is corrected for small sample sizes,

$$AICc_j = AIC_j + \frac{2K_j^2 + 2K_j}{n_{samples} - K_j - 1} \tag{4.5}$$

with $n_{samples} = 100$, since we used the same sample size as described in [1].

4.1.3 Calculating Bayesian posterior probability per candidate model

Each candidate model is considered equally likely prior to fitting by Multinest. That is, every candidate model has an equal prior probability of being the optimal model to represent the simulated data,

$$\Pr(M_j) = \frac{1}{|M|} \tag{4.6}$$

where M is the set of all candidate models. Model evidence, or marginal likelihood, $\Pr(D \mid M_k)$ is estimated by Multinest, [29, 30, 31] as discussed in **Section 4.1.1**. With a prior probability and marginal likelihood, the posterior probability per model can be calculated as

$$\Pr(M_k \mid D) = \frac{\Pr(D \mid M_k)\Pr(M_k)}{\sum_j \Pr(D \mid M_j)\Pr(M_j)} \tag{4.7}$$

4.1.4 Calculating sums of AIC weights (SW) per predictor variable

Using all candidate models, sums of AIC weights (SW) can be calculated. SW describes which predictor variables, with their fitted coefficients, best match the response variable. The first step toward calculating SW involves scaling models with respect to the minimum AIC value [24, 25]. This results in AIC differences,

$$\Delta_i = AIC_i - min(AIC) \tag{4.8}$$

where $min(AIC)$ is the AIC value of the lowest-scoring (best) model candidate. Using AIC differences, Akaike weights are calculated,

$$w_i = \frac{e^{-\frac{\Delta_i}{2}}}{\sum_{r=1}^{R} e^{-\frac{\Delta_i}{2}}} \tag{4.9}$$

where R is the number of models in the candidate set. Akaike weights represent relative likelihoods of the models given the data, and are interpreted as probabilities [24].

Sums of AIC weights (SW) can then be calculated using

$$SW(j) = \sum_{k=1}^{R} \{w_k \text{ if } j \in M_k\} \tag{4.10}$$

where $SW(j)$ is the sum of weights for variable j, and each model's weight w_k is summed only if variable j is present in that model.

4.1.5 Calculating Bayesian posterior probability per predictor variable

Each predictor variable that could be included in the model has an assigned prior probability based on our prior expectations. We take an approach where we consider each predictor variable as equally likely to be part of the "true" model as it is to not be part of the true model. This means a prior probability of 50% (50% likely the variable is part of a model that is the best representation of the system that generated the data).

Each candidate model can then be assigned a prior probability conditional on the hypothesis being considered, $\Pr(M_k \mid H_i)$, where M_k is the k^{th} candidate model and H_i represents the inclusion of one of the predictor variables. The calculation of $\Pr(M_k \mid H_i)$ is based on the number of candidate models that fall under the hypothesis being considered,

$$\Pr(M_k \mid H_i) = \frac{\Pr(H_i)}{|M_j \in M, H_i|} \tag{4.11}$$

where $M_j \in M, H_i$ is the set of all models assigned to H_i. For example, if H_i is the hypothesis that predictor variable x_1 belongs in the true model, then $M_j \in M, H_i$ is the set of all candidate models that include x_1.

Using this prior probability, the posterior probability for an individual model, conditional on the hypothesis being considered, can be calculated as

$$\Pr(M_k \mid D, H_i) = \frac{\Pr(D \mid M_k, H_i) \Pr(M_k \mid H_i)}{\sum_j \Pr(D \mid M_j, H_i) \Pr(M_k \mid H_i)} \tag{4.12}$$

where $\Pr(D \mid M_k, H_i)$ is the Bayesian model evidence (marginal likelihood) for $Model_k$.

The posterior probability for an individual model k under hypothesis H_i, $\Pr(M_k \mid D, H_i)$, is not directly used, as the posterior probability of H_i itself, $\Pr(H_i \mid D)$ is of principal interest. Under Bayes' Theorem,

$$\Pr(H_i \mid D) = \frac{\Pr(D \mid H_i) \Pr(H_i)}{\sum_j \Pr(D \mid H_j) \Pr(H_j)} \tag{4.13}$$

where $P(D \mid H_i)$ is the marginal likelihood of H_i over all models to which it applies, $M_j \in M, H_i$. According to [178], this can be calculated as

$$\Pr(D \mid H_i) = \sum_k \Pr(D \mid M_k, H_i) \Pr(M_k \mid H_i) \tag{4.14}$$

with a summation instead of an integral because each model has a discrete probability as calculated in Eq. (4.12). Using the results of Eq. (4.14) in Eq. (4.13), we then calculate the posterior probability for each predictor variable, pictured in **Figure 5.1E** and noted in **Table 5.2**.

4.2 Methods: Unified Tumor Growth Mechanisms from Multimodel Inference and Dataset Integration

4.2.1 CIBERSORT deconvolution of RNA sequencing data

Data from two GEMM models provide multiple replicates of tumors from two genetic backgrounds: one genetic background being $p53^{fl/fl};Rb^{fl/fl};p130^{fl/fl}$ (triple-knockout, or TKO) GEMM tumors [10], and another as $Rb1^{fl/fl};Trp53^{fl/fl};Lox-Stop-Lox[LSL]-Myc^{T58A}$ (RPM) GEMM tumors [179]. We also used publicly available SCLC cell line data. Having been originally derived from human tumors, each cell line has a different genetic background, and therefore we have only one (genetically identical) replicate per cell line sequencing event. To approximate genetic similarity between cell lines, and thus approximate multiple replicates, we expect that cell lines exhibiting similar steady state composition will be more genetically similar than those whose steady state compositions differ. Previously, in [180], we both clustered publicly available SCLC cell line data into clusters that align with the different SCLC subtypes and used CIBERSORT to deconvolute the proportions of cell line data and tumor samples into SCLC subtypes from their RNA sequencing signatures.

4.2.2 Population dynamics modeling in PySB

A population dynamics model represents the abundance of species over time, whether increase or decrease due to birth/growth or death. We use ordinary differential equation (ODE) models coded via PySB to generate population dynamics models [116]. PySB is a rule-based modeling language, where one will encode

$$A \xrightarrow{k} 2A,\ k = 0.469 \tag{4.15}$$

to indicate that A doubles at a rate of 0.469 doublings per day.

Inter-subtype effects are represented by the increase or decrease of the rate of affected reaction. For example, the above division rule has a baseline rate of 0.469 doublings per day, but in the presence of an effector subtype the division rule will have a rate of $0.469 * 1.05 = 0.493$ doublings per day. In this case the effector subtype has increased the division rate by 5%. Thus the rule-based representation is

$$A + Y \xrightarrow{k} 2A + Y,\ k = 0.493 \tag{4.16}$$

to indicate that A doubles at a rate of 0.493 doublings per day in the presence of Y.

To simulate the passage of time, the speed at which the division/death/transition reaction occurs – its rate, k, of cells per unit time – must be assigned as in the equations above. While a literature search reveals approximate rates of division and death among different SCLC subtypes (**Table C.1**), each of these are in a different context than the system we model here – for example, division rates for the A subtype are measured *in vitro* in the presence of only that one subtype, whereas our population dynamics model is meant to simulate this subtype in the presence of others as well as *in vivo* in a mouse tumor. Therefore, we use the rates in the literature as our prior expectations for division and death, that is, we use these values as approximate starting values for these parameters during the estimation process. Other rates, such as those indicating the speed of transitions between subtypes, or any rates including the effects of Non-NE subtypes, have not previously been noted in the literature and we used much wider ranges for each as our prior expectations. Rate prior expectations (**Figure C.2**) are then provided to the Multinest algorithm to perform nested sampling.

4.2.3 Multiple hypothesis generation via HypBuilder

Because we perform model selection, we use 5,891 ODE models coded via rule-based modeling in PySB. Each model is generated to include or exclude from 44 reaction rules. There are eight rules that represent division and death for each subtype, and with the potential for three different inter-subtype effects (including none) to have an impact on division or death, each division and death reaction has 3 options, leading to 24 potential rules relating to division/death in total. There are four rules that represent hierarchical phenotypic transitions, which likewise have three potential inter-subtype effects, for 12 rules in total representing hierarchical phenotypic transitions. There are eight rules related to non-hierarchical phenotypic transitions, for 20 total potential phenotypic transition rules out of the 44 rule options.

We use HypBuilder (https://github.com/LoLab-VU/HypBuilder) to generate the 5,891 PySB models that we would otherwise have to code by hand. HypBuilder is software for the automatic generation user-specified collections of mechanistic rule-based models in the PySB format. The input CSV file contains a global list of all possible model components, and reactions, as well as any instructions regarding model creation. The instructions dictate which subsets of model components and reactions will be combinatorially enumerated to create the collection of models. The reactions are parsed via HypBuilder's molecular interaction library, a library of defined reaction rule sets that is outfitted with common PySB interactions and is customizable to include more interactions should the user need them. Once parsed and enumerated each combination of

rules is exported as an executable model via PySB. The instructions for model construction used in this work direct HypBuilder to use a "list" method to enumerate all candidate models of interest using prior knowledge of likely combinations of model variables (see https://github.com/LoLab-MSM/Bayes-MMI for code used to enumerate candidate models and create the list for HypBuilder).

If the candidate model set contains every relevant biologically plausible possibility, we can consider the entire set of models as representative of 100% of the probability that one of the candidate models explains, or provides the mathematical basis underlying, the data. This is an assumption that cannot truly be met, and most model selection literature acknowledges that one cannot find the "true" model [24, 25]. However, prior knowledge enables us to determine that all 5,891 models represent all possibilities with regard to outstanding SCLC hypotheses to the best of our ability. We visualize the prior expectations for the 44 rate parameters as a probabilistic distribution per parameter (prior marginal distribution) (**Figure C.2**). Correspondingly, a probabilistic representation of best-fitting rates for each model is returned by the Multinest algorithm (posterior marginal distribution) (e.g., **Figures 6.5B** and **6.6E**; **Figures C.5B** and **C.4**).

4.2.4 Parameter estimation and evidence calculation by nested sampling

As noted in Eqs. (4.15) and (4.16), rate parameters must be set in order to run simulations of a mathematical model. Parameter estimation is the process of determining optimal rates that result in a model simulation recapitulating the data it is meant to represent. Multiple methods exist for parameter fitting or model optimization, [181, 182] with Bayesian methods utilizing a prior rate parameter distribution, $\Pr\left(\vec{\theta}\right)$, and a likelihood function to assess a parameter set $L(D \mid \vec{\theta}_{\mathbf{i}})$ as in Eq. (4.1).

For our likelihood function, we represent SCLC tumor steady-state proportion probabilistically, generating a Beta distribution (bounded by 0 and 1) to represent the means and standard deviations of sample replicate subtype proportions, accounting for noise in the proportional space. We test $n-1$ subtype proportions to ensure independence of each sampled subtype proportion, to result in the probability of $n-1$ independent events (tumor subtype proportions) occurring.

With a prior probability, $\Pr\left(\vec{\theta}\right)$, and a likelihood, (Eq. (4.1)) the posterior probability for each parameter set $\vec{\theta}_{\mathbf{i}}$ can be calculated via Bayes' Theorem,

$$\Pr\left(\vec{\theta}_{\mathbf{i}} \mid D\right) = \frac{\Pr\left(D \mid \vec{\theta}_{\mathbf{i}}\right)\Pr\left(\vec{\theta}_{\mathbf{i}}\right)}{\int \Pr\left(D \mid \vec{\theta}\right)\Pr\left(\vec{\theta}\right) d\vec{\theta}} \tag{4.17}$$

The denominator of Bayes' Theorem represents the likelihood integrated over all parameter sets, called the marginal likelihood or model evidence. Nested sampling computes this value [28].

To perform nested sampling, we utilize the Multinest algorithm [29, 30, 31], itself described in **Sec-**

tion 4.1.1.

For our likelihood function we represented simulation outcomes – proportions of subtypes at steady state – by a Beta distribution, calculating α and β using the mean μ and variance σ^2 of each dataset [183]:

$$\nu = \alpha + \beta = \frac{\mu(1-\mu)}{\sigma^2} - 1, \text{ where } \nu = (\alpha+\beta) > 0 \text{ (meaning } \sigma^2 < \mu(1-\mu)) \tag{4.18}$$

$$\alpha = \mu\nu = \mu\left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right), \text{ if } \sigma^2 < \mu(1-\mu) \tag{4.19}$$

$$\beta = (1-\mu)\nu = (1-\mu)\left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right), \text{ if } \sigma^2 < \mu(1-\mu) \tag{4.20}$$

We calculated the log-likelihood of each subtype mean from the dataset being fit against the simulated subtype's beta distribution,

$$\sum_{s}^{S-1} \begin{cases} \log\left(\frac{x_{s_{data}}^{\alpha_{s_{sim}}-1}\left(1-x_{s_{data}}\right)^{\beta_{s_{sim}}-1}}{B(\alpha,\beta)}\right), & \text{if } s \in \text{simulated model topology} \\ \log\left(\lambda e^{-\lambda x_{s_{data}}}\right), \text{ where } \lambda = \frac{1}{\sigma_{s_{data}}}, & \text{otherwise} \end{cases} \tag{4.21}$$

where S is the set of subtypes, $\alpha_{s_{sim}}$ and $\beta_{s_{sim}}$ indicate the calculation of α (Eq. (4.19)) and β (Eq. (4.20)) using the proportion of subtype s from the simulation and the variance of subtype s in the data; $x_{s_{data}}$ is the mean proportion of subtype s in the dataset, $B(\alpha,\beta)$ is $\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ and Γ is the Gamma function, and $\sigma_{s_{data}}$ is the standard deviation of the data. Using the exponential function ($\log\left(\lambda e^{-\lambda x_{s_{data}}}\right)$) as part of the log-likelihood (Eq. (4.21)) enabled us to calculate a likelihood value for subtypes not present in a model's topology, which should be a poor log-likelihood if the subtype has a high proportion in the data but was not included in the model topology, or a better log-likelihood if the subtype has a low proportion in the data but was not included in the model topology (and therefore potentially contributing to overfitting). The Python module `scipy.stats` was used to calculate the Beta log likelihood (Eq. (4.21), above) and the exponential log likelihood (Eq. (4.21), below). A simulation would not be scored (return `NaN` and thus be thrown out by the Multinest fitting algorithm) if the tumor subtype proportions did not reach steady state (calculated by whether a proportion timecourse had a slope of zero for the last 7.5% of the simulation).

Multinest is run per model per dataset, which equates to performing 5,891 mechanistic interpretations, 3 times each. CPU time for one model fitting was on average 19 hours ($\sim$ 0.80 days), with a range of 5 minutes to 28 days. If Multinest had not reached its stopping point by 28 days, we assumed that all regions of parameter space were similarly unlikely and that further running of the algorithm would only continue

to refine the search of the unlikely space; models with this difficulty are very likely to have low marginal likelihood due to the unlikeliness of the parameter space. We do not include these incompletely-searched models in our multimodel inference analyses (Figs 3-5) and we confirmed that all models that reached 28 days of CPU time without reaching the Multinest stopping point have an extremely low evidence value at the time they were terminated.

4.2.5 Candidate model prior and posterior probabilities and confidence interval calculation

Each candidate model has an equal prior probability of being the optimal model to represent the underlying SCLC tumor system, as shown in Eq. (4.6), where $|M|$ is the set of all candidate models. With the model evidence, or marginal likelihood, $\Pr(D \mid M_k)$ estimated by Multinest, [29, 30, 31] the posterior probability per model $\Pr(M_k \mid D)$ can be calculated as shown in Eq. (4.7).

With a posterior probability per model, (**Figure 6.4**) we calculate a 95% confidence interval. This is accomplished by summing decreasing model posterior probabilities until the sum is 0.95, then considering those models as our 95% CI [24] (**Figure 6.4**, orange). Using this confidence interval results in $\sim$ 1000 models per dataset, a considerable decrease from the initial 5,891. This is a more traditional approach to determining a confidence set of models.

We also took an approach discussed in [24]. In this approach, a CI is informed by use of the Bayes Factor between the highest-scoring model and consecutively decreasing scoring models, until the Bayes Factor is larger than a particular cutoff. The models in this CI would be those models i for which $\frac{\Pr(M_{highest})}{\Pr(M_i)} > cutoff$. Burnham and Anderson [24] denote such a method as a "relative likelihood confidence interval" and discuss its support by statistical theory, noting that it is uncommonly found in the model selection literature. We used a cutoff of $10^{\frac{1}{2}}$, the lowest Bayes Factor at which a difference may be determined [178]. Even with this permissive cutoff, the relative likelihood CI includes only tens of models, an even greater decrease from the initial number of candidates.

4.2.6 Prior and posterior probabilities per hypothesis being investigated

For all hypotheses, similar to **Section 4.1.5**, we considered each hypothesis as equally likely compared to competing hypotheses: for inclusion of most model terms, this was a prior probability of 50%, where it is 50% likely the model term is part of a model that is the best representation of the tumor system, and 50% likely that same term is not part of that model. For the inclusion of effects in the candidate models, the prior probability for a given effect is 33%, where it is equally likely that an effect is generated by Y, generated by A2 and Y, or that no effect is present. The comparison between effect types (including none) is included in **Tables C.2** to **C.4**, while the comparison of any effect at all vs. no effect (50% vs. 50%) is included in the

main text.

For the model topology analysis, we considered it equally likely that any model topology could best represent the tumor system that generated each dataset, and with 11 possible model topologies this resulted in a 9% prior probability per model topology (**Figure 6.5A**). For model initiating subtype hypotheses, (**Figure C.1G**) with 15 potential combinations of initiating subtypes, each initiating subtype combination has a 6.67% prior probability.

Each candidate model is then assigned a prior probability conditional on the hypothesis being considered, $\Pr(M_k \mid H_i)$, as in Eq. (4.11). The calculation of $\Pr(M_k \mid H_i)$ is based on the number of candidate models that fall under the hypothesis being considered. For example, if H_i is the hypothesis that the model term "A to Y transition" is part of the model that would best represent the SCLC tumor system, then $M_j \in M, H_i$ is the set of all candidate models that include the "A to Y transition" model term.

The posterior probability for an individual model, conditional on the hypothesis being considered, $\Pr(M_k \mid D, H_i)$, is then calculated as in Eq. (4.12). However, we are principally interested in the posterior probability of the hypothesis H_i being investigated, $\Pr(H_i \mid D)$, calculated using Bayes' Theorem as detailed in Eq. (4.13). $\Pr(D \mid H_i)$, the marginal likelihood of H_i over all models to which it applies, $M_j \in M, H_i$, can be calculated according to [178] as in Eq. (4.14).

Using the results of Eq. (4.14) in Eq. (4.13), we then calculate the posterior probability for each hypothesis, pictured in **Figures 6.6A** to **6.6D** and noted in **Table 6.2**. In this way, we can use Bayesian calculation rather than parameter importance analysis (sums of AIC weights) [1, 26] (see **Chapter 5** and methods in **Section 4.1**) to determine the posterior probability of each model term. This also enables us to avoid bias in considering models with and without certain model term, if an uneven number of candidate models contain a model term vs. do not contain the term [184].

4.2.7 Posterior odds per hypothesis being investigated

All model terms and variables begin with a prior probability of 0.5. With equal prior probabilities across all model term hypotheses, the posterior odds represented by $\frac{posterior\ probability}{(1-posterior\ probability)}$ is equivalent to the Bayes Factor (see **Section 5.6**). Therefore, calculation of the posterior odds and the Bayes Factors for each model term are equivalent.

A posterior probability of model term inclusion of 0.75 or more, or probability of 0.25 or less, would be considered substantial evidence for inclusion or exclusion of that term, respectively [178]. Given the nature of the posterior odds, where a value of 2 indicates that one hypothesis is twice as likely to be true as the other, we also consider posterior probabilities of 0.667 or more, or 0.333 or less, to be notable evidence for inclusion or exclusion of the model term considered. We consider probabilities between 0.333 and 0.667 to

not have been significantly informed by the data.

4.2.8 Bayesian model averaging of parameter sets

Since Multinest returns multiple best-fitting parameter sets, each parameter in a model has a frequency distribution representing the values it takes on over these parameter sets. We thus consider each parameter using a probabilistic representation, per model (posterior marginal distribution) (**Figures 6.5B** and **6.6E**; **Figures C.5B** and **C.4**). Since each candidate model is assigned a posterior probability as in Eq. (4.7), all best-fitting parameter sets for that model can be assigned the same posterior probability. The frequency distribution of one parameter's values across a model's best-fitting parameter sets are thus weighted by its model's posterior probability. Then, the frequency distributions of weighted parameter values per model can be combined, representing the distribution of potential values of a particular parameter, weighted by model posterior probabilities. This way, parameter values in the distribution that come from models with a higher posterior probability (thus higher model evidence) will have more of an effect on the probabilistic representation, since they represent more likely values for the parameter.

To assemble representative fitted parameter sets for each candidate model, we used the first 1000 parameter sets from the Multinest equally weighted posterior samples per model. With up to 44 parameters and up to 5,891 models, the collection has 44 parameter columns and up to 5,891,000 rows representing a parameter vector. The collections were made per dataset.

4.2.9 Comparing parameter distributions

As above, each kinetic parameter has a frequency distribution representing 1000 fitted values per candidate model, meaning up to 5,891,000 fitted values across all models (weighted using Bayesian model averaging, as above). To compare parameter rates across models in the same dataset but with different topologies, we grouped each parameter according to the model topology from which it came. We then sampled 1000 values from the BMA-weighted distribution per kinetic parameter across all models of the same topology. We performed ANOVA followed by Tukey HSD at family-wise error rate (FWER) of 0.01, using the Python module statsmodels. Below an FWER of 0.01, we considered the sampled parameters significantly different across models. We then repeated the sampling, ANOVA, and Tukey HSD for a total of 10 iterations. We then averaged across determinations of significant/non-significant and if a parameter comparison across model topologies was significantly different more often than it was not different, we considered the parameter rates to be different comparing model topologies. The same methodology was used to compare parameter rates across different datasets.

4.2.10 Generating a consolidated model of the SCLC tumor

A hypothesis (model term) whose posterior probability is further from its prior probability indicates more information gained during the nested sampling process – more knowledge provided by the data. Conversely, a posterior probability similar to the corresponding prior probability indicates that the data did not inform our prior knowledge.

To unify the varying models into one view of SCLC biology, we brought together model probabilities from each three-subtype topology per dataset (**Figure 6.6E**). To bring together the results for each three-subtype topology results in the investigation of what appears as a four-subtype topology. In fact, if we are to envision one model that can represent one system that generated all three datasets, it would need to include all four subtypes. We consider this a reasonable practice in that all transition posterior probabilities in the three-topology subtypes either were little informed by the data or had a value indicating that transitions are likely; in addition, all Non-NE effects were either little informed by the data or had a value indicating that these effects are unlikely. Posterior probabilities were not the same between three-subtype topologies, but these trends of likely or unlikely model features generally agreed.

When consolidating models in this way, if model terms were part of multiple topologies (e.g., the A-to-N transition is part of the A, N, and Y topology, best representing the RPM dataset, and the A, N, and A2 topology, best representing the SCLC-A cell line dataset) we took the posterior probability of the model feature closer to 0.5. For example, the posterior probability for the A to N transition in the RPM dataset is 0.709 and the posterior probability for this same transition in the SCLC-A dataset is 0.626. Therefore, in the four-subtype consolidated representation, the posterior probability for the A to N transition is 0.626. This is the most conservative way to represent the knowledge gained by the data from the perspective of the entire SCLC system, allowing for the most uncertainty to remain. We consider this practice as avoiding claiming more certainty about model features than the data may provide.

4.3 Methods: Evaluating small cell lung carcinoma time-course patient-derived xenografts reveals phenotype switching with treatment and chemotherapy-stable subpopulations

Mass cytometry on PDX mice was performed and the resulting data provided to me for this manuscript by Dr. Jonathan Lehman. Methods text related to these two (sections “Mice” and “Mass cytometry”) provided by Dr. Lehman.

4.3.1 Mice

Six- to eight-week-old athymic outbred Foxn1nu/nu (Charles River Labs) mice were injected with 2-5 million cells of the specified patient-derived xenograft cell line 1:1 in Matrigel (BD Biosciences) into unilateral mouse

flank. For the experiments herein, mice were observed for tumoral development with tumor harvesting prior to tumors reaching 2000 mm3 in size for baseline studies. Tumors were harvested for single cell disaggregation as noted and samples were also fixed in 10% buffered neutral formalin for histologic evaluation.

Comparative chemotherapy treated animals received a pre-treatment 18-gauge core biopsy which was disaggregated for analysis and received additional core biopsy prior to each cycle of chemotherapy treatment. Mice were treated with carboplatin IV via tail vein injection on Day 1 of treatment, and etoposide phosphate subcutaneously daily on days 1-3 of each q2week treatment cycle. A noted subset of mice received serial core biopsy within 96 hours of each q2week chemotherapy treatment cycle.

All animals in this study were maintained in AAALAC accredited mouse facilities at Vanderbilt University Medical Center and in accordance with IACUC regulations and protocols at Vanderbilt University Medical Center.

4.3.2 Mass cytometry

Metal-tagged antibodies were used to stain cells in 100 μL cell staining media for 30 minutes at room temperature. After staining, cells were washed once with PBS and once with deionized water, pelleted at 800g, and resuspended in deionized water containing normalization beads (Fluidigm). Standard bead-based normalization was used as previously described [32]. Cells were collected on a CyTOF 2.0 at the Vanderbilt Flow Cytometry Shared Resource. Original data were normalized with MATLAB normalization software prior to further analysis using Cytobank [36], followed by established as well as novel mass cytometry analysis methods as noted in the following sections. The antibody panels used are as noted in **Tables 7.1** and **7.2**.

4.3.3 Visualization of clusters and treatment conditions within tumors

For visualization of unique samples across multiple conditions, (e.g., LX22 comparison between treatment conditions and time), bead-normalized (CytoBank [185]) .fcs files representing each timepoint (pre-treatment, peri-treatment, post-treatment) for one PDX (LU73 or LX22) were transformed into `Pandas` dataframes, (Python) and then concatenated into one dataframe representing all samples, conditions, and timepoints for that PDX. Rows (individual cells read by the CyTOF) were removed if the marker for mouse MHC class I was positive, indicating that that cell was a mouse cell and not a tumor cell from the PDX. Individual cells were also removed if the measure of human histone 3 was less than 10, with 10 indicating positivity, similarly to ensure the cell was of human origin. After these filtering features, 50-85% of the initial rows remained, depending on the dataset (**Figure 7.2A**, left).

CyTOF measurements in the dataframe were then transformed, dividing each value by 15 [186] and then taking the value of the hyperbolic arcsine using the numpy package (np.arcsinh()). The package UMAP-learn

was used to reduce the dimensions of the transformed 44-dimensional dataframe using 1000 neighbors and a `min_dist` of 0.0 (**Figure 7.2A**, middle). The dimension-reduced x- and y-coordinates for each cell, where the cell is represented by the row of the dataframe, were added to that row of the dataframe. In this way, the UMAP coordinates could be permanently associated with the cell and its marker measurements.

Then, Leiden clustering was used to determine the neighborhoods of the dimension-reduced data using the graph object calculated by UMAP-learn (**Figure 7.2A**, right). Labels denoting the sample condition within the dataframe were not used for clustering but were kept for comparison of Leiden clusters to treatment/mock conditions and time. This process identified 14 clusters in the LU73 PDX and 17 in the LX22 PDX.

4.3.4 Comparing samples and clusters across time

For comparing differences in samples across time, the Pandas dataframes per PDX (as constructed above) was used. The density of cells in the dimension-reduced space per sample and condition within a PDX (e.g., LU73 sample 121, pre-treatment) was plotted as a contour map using seaborn's kdeplot function on the UMAP coordinates from a subset of the PDX's dataframe derived from that sample and condition. Density changes across the dimension-reduced space over time per sample could thus be compared graphically (**Figures 7.2B** and **7.2C**, bottom).

The Leiden clusters were used for calculating the proportion of cells in a sample and condition that comprise each cluster. To calculate this, the subset of the PDX's dataframe derived from a particular sample and condition was taken as above, and the number of cells (rows of the dataframe) in that sample and condition used as the total. Within that subset, the number of cells in each Leiden cluster was then divided by the total, resulting in a proportion. In **Figures 7.2B** and **7.2C** barplots, this results in each condition's bars summing to 1 (100%): for example in **Figure 7.2B**, all dark green bars, representing LU73, sample 121, pre-treatment, sum to 1. These proportions may then be used to assess in which samples and conditions a cluster was most occupied, and where cluster occupancy changed over time or with treatment.

4.3.5 Cluster identification

The phenotypic identification of each cluster was assessed via multiple methodologies.

The first methodology is via statistically significant (per the hypergeometric test; see below) GO terms per combination of markers in a cluster, where those markers indicate that the cluster has either higher expression or lower expression of that marker across all clusters. First, to assign markers to clusters, the (transformed, see above) expression values of each marker was averaged across cells in each cluster, resulting in the mean expression of a marker per cluster. Per marker, clusters were then ranked by mean expression, and the 3 highest-ranked clusters for a marker were denoted as "high" in that marker. The 3 lowest-ranked clusters for a

marker were denoted “low” in that marker. Assigning high and low markers per cluster in this way guarantees that 3 clusters at most will be “high” and 3 at most will be “low” for each marker, eliminating the possibility that, for example, all clusters will have the same highest-expressed marker and thus be indistinguishable. However, this does result in the potential for a cluster not to be considered high or low in any marker. After this step, every cluster has a list of markers in which it is high and low, which can be compared to GO terms to investigate biological categories each cluster may be considered high or low in.

However, one marker out of many that a cluster may be high or low in, matching with its corresponding gene in a GO term, may be accountable only to random chance. Therefore, we use the hypergeometric test to assess whether markers assigned to a cluster truly represent a GO term/biological process beyond random chance. In this case, the hypergeometric test determines whether a cluster is enriched in genes representing a particular GO term, using the number of matching markers and corresponding genes, as well as the total number of markers in the cluster’s high or low list and the total number of genes in the GO term list. Since the CyTOF panel used for this analysis includes 42 markers (not counting HH3, mMHC I or Rhodium), while GO terms can use any gene in the genome, all GO term gene lists were filtered such that they included only genes corresponding to the protein markers in the CyTOF panel. Besides assigning GO terms as “high” or “low” in each Leiden cluster if the hypergeometric test returned a p-value of <0.05, we also required that at least 3 markers in the cluster high or low list overlap with genes in a GO term, or more than n-1 markers overlap if n <3, where n is the number of markers in the high or low list. With this process, each cluster is assigned biological processes it is “high” or “low” in, suggesting the phenotypic identity(ies) of each cluster (**Figure 7.3A**; **Figure D.2**; **Tables D.1** and **D.2**).

Marker Enrichment Modeling (MEM) was also used for phenotypic identification of each cluster (**Figure 7.3C**). MEM calculates enrichment scores per marker per cluster [187]. MEM scores are based on median expression levels of each marker per cluster as well as the interquartile range (difference between the 75th percentile of a marker’s values and the 25th percentile of a marker’s values) of that marker.

4.3.6 Cluster density over time

Each sample and condition was subset from the LX22 dataframe and plotted via the seaborn scatterplot() function in gray, serving as the background set of cells that could possibly be part of a cluster in that sample and condition. Then, cells in that sample and condition and in the indicated cluster were plotted using the Python package datashader to color each point by density in the cluster (**Figure 7.3B**).

4.3.7 SPADE

SPADE analysis was performed on LX22 treated vs. untreated tumor data. Each circle in the SPADE diagram (**Figure 7.4C**) represents a population of cells with similar phenotype based on the 35 markers in **Figure 7.4B** and **Table 7.2**. SPADE clusters the cells based on these similar phenotypes and creates a minimum-spanning tree denoting the relationships between cells. The minimum-spanning tree is able to be visualized in 2 dimensions, resulting in the SPADE diagram as seen in the figure. The sizes of the circles (clusters of cell populations) in the diagram are based on the number of cells in each circle's population. While the SPADE minimimum-spanning tree was built using both the untreated and treated LX22 sample, (similar to performing Leiden clustering on all samples and timepoints within a PDX, detailed above) the cluster sizes within the diagram are based on either the untreated (**Figure 7.4C**, top) or treated (**Figure 7.4C**, bottom) sample. In this way we were able to compare the densities of cells within each phenotype and how these densities change with treatment. The heat color for each circle represents the median expression of Oct3/4 in cells within that circle, indicating subpopulations of high Oct3/4 expression in both treated and untreated LX22 tumor cells.

CHAPTER 5

A didactic example to contrast AIC *vs* posterior probability calculated by Bayes-MMI for model selection and multi-model inference [1]

5.1 Using multiple models to evaluate how well a variable informs the observed data: an example

Mathematical models are a useful way to interrogate a biological system. Building a model can help the investigator hypothesize relationships within the system, and more specifically evaluate whether and how these relationships relate to observed phenomena. Each piece of a model represents a distinct hypothesis; thus the model as a whole represents multiple hypotheses about the system of interest.

For example, a farmer might wish to determine which aspects of his farm most affect his monthly income: milk from his cows, eggs from his hens, wool from his sheep, or pony rides offered to the public. He is interested in assessing how he should price the product or service based on how many animals he must keep his income steady. Additionally, if a provided product or service doesn't impact his monthly income, perhaps he will stop providing it. This is a question with one outcome or response variable, the monthly income, and up to four predictor variables, one each for the animals that provide the products or services. We can write a linear equation that describes these interactions as shown below:

$$income = p_{milk} * cows + p_{eggs} * hens + p_{wool} * sheep + p_{rides} * ponies + savings\ interest \tag{5.1}$$

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_0 \tag{5.2}$$

Eq. (5.2) represents a linear regression model corresponding to the more intuitively written Eq. (5.1), where the savings interest represents the "intercept", or a baseline amount of money where the farm budget starts each month. However, Eq. (5.2) is only one of many possible models that the farmer could use to determine how these variables come together to yield his income and perhaps improve on his earnings. A multi-model averaging approach would enable the farmer to explore model hypotheses to explore how much he should charge, β_{1-4}, based on number of working animals that month, x_{1-4}, and his monthly income, y, and, how much each of those animals x_{1-4} contribute to the income y.

Besides using linear regression to determine pricing, the farmer can evaluate all potential models including few, some, or all farm animals to see which animals most directly affect his income. Perhaps the

[1] This chapter appears in an updated version as part of the main text and supplement of 'Unified Tumor Growth Mechanisms from Multimodel Inference and Dataset Integration' published in *PLOS Computational Biology* and has been reproduced with the permission of the publisher and my co-authors, Leonard A. Harris, Michael A. Kochen, Julien Sage, Vito Quaranta, and Carlos F. Lopez.

lower-earning products per month are not needed to maintain the income based on the numbers of animals providing each product. For example, if sheep's wool and pony rides do not bring in as much income as milk or eggs, maybe they do not contribute meaningfully to the income, and the best model would look like this:

$$income = p_{milk} * cows + p_{eggs} * hens + savings\ interest \tag{5.3}$$

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_0 \tag{5.4}$$

With all possible combinations of animals that the farmer might use to support his income, there are 16 different possible models that can help predict his monthly income, below:

$$\begin{array}{lll}
y = \beta_0 & y = \beta_1 x_1 + \beta_2 x_2 + \beta_0 & y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_0 \\
y = \beta_1 x_1 + \beta_0 & y = \beta_1 x_1 + \beta_3 x_3 + \beta_0 & y = \beta_1 x_1 + \beta_2 x_2 + \beta_4 x_4 + \beta_0 \\
y = \beta_2 x_2 + \beta_0 & y = \beta_1 x_1 + \beta_4 x_4 + \beta_0 & y = \beta_1 x_1 + \beta_3 x_3 + \beta_4 x_4 + \beta_0 \\
y = \beta_3 x_3 + \beta_0 & y = \beta_2 x_2 + \beta_3 x_3 + \beta_0 & \\
y = \beta_4 x_4 + \beta_0 & y = \beta_1 x_1 + \beta_4 x_4 + \beta_0 & y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_0 \\
 & y = \beta_3 x_3 + \beta_4 x_4 + \beta_0 &
\end{array} \tag{5.5}$$

The model selection process that the farmer can use involves evaluating how well each candidate model's (from this superset of plausible models) simulations of $pricing * animal$ compared to the observations of his income. This process is often performed for choosing one best-matching model, and whichever model is best is often considered to include only the important model variables (the most important animals from the perspective of the farmer's income. However, we and others argue, one may use model averaging, where parameter values (prices) can be weighted by model probability and then combined into a distribution of likely values, and similarly variable (animal) likelihood can be weighted and summed across all models [24, 188]. Thus, this set of candidate models can provide support for or against *pieces* of models, and therefore support for or against hypotheses.

What is the best framework for which the farmer can perform model selection and model averaging? We argue that a Bayesian framework is optimal, compared in this example to the most common model selection/model averaging framework, the information theoretic approach of Akaike Information Criterion, or AIC [24, 25]. In **Sections 5.2** to **5.6**, we provide the background information that motivate our claim that a Bayesian framework should be optimal. In **Section 5.7**, we provide our Bayesian analysis of a published

linear regression example that investigated AIC model selection and model averaging to compare the two frameworks.

5.2 Marginal likelihood or "evidence" is calculated via model optimization followed by Bayes' Theorem

In a Bayesian statistics context, probability indicates the degree of belief or confidence in an event, or in a proposal that a statement may be true – that is, a hypothesis. This prior degree of belief (hypothesis) is then updated with new data, which can yield evidence that our prior belief must be modified. In this case, data can refer to measurements related to the hypothesis and is synthesized into a likelihood that the hypothesis is true. Prior knowledge combined with data leads to the posterior, updated, probability, defined using Bayes' Theorem as

$$\Pr(H \mid D) = \frac{\Pr(D \mid H)\Pr(H)}{\Pr(D)} = \frac{\Pr(D \mid H)\Pr(H)}{\int \Pr(D \mid H)\Pr(H)\, dH} \tag{5.6}$$

where $\Pr(H \mid D)$ denotes a conditional posterior probability of the hypothesis (H) given the data D, $\Pr(D \mid H)$ denotes the probability of observing the data D if the hypothesis H were true (also called the likelihood), and $\Pr(H)$ indicates prior probability (degree of belief) of the hypothesis before being presented with more data. Calculation of the posterior probability of the hypothesis $\Pr(H \mid D)$ requires dividing the likelihood of one hypothesis times its prior, $\Pr(D \mid H)\Pr(H)$, by $\Pr(D)$ (Eq. (5.6), middle). When comparing many models, $\Pr(D)$ becomes the integral over the likelihood times the prior probability for each hypothesis H within the set of all hypotheses H (Eq. (5.6), right).

When the hypothesis H in Eq. (5.6) represents a candidate model M with n parameters, we can think of model M as the set of parameters, $\vec{\theta_{\mathbf{M}}} = \{\theta_1, \theta_2, ..., \theta_n\}$. Each parameter θ_i itself represents one dimension of the model's probability space. Model optimization to data D assigns a likelihood to each $\vec{\theta_{\mathbf{M}}}$ in probability space, and these values over the entire parameter space represent the numerator in Eq. (5.6) with $H = \vec{\theta_{\mathbf{M}}}$, $\Pr\left(D \mid \vec{\theta_{\mathbf{M}}}\right)\Pr\left(\vec{\theta_{\mathbf{M}}}\right)$, a distribution proportional to the probability distribution of model M. To calculate the probability distribution of M, we must also calculate the denominator of Eq. (5.6), $\int \Pr\left(D \mid \vec{\theta_{\mathbf{M}}}\right)\Pr\left(\vec{\theta_{\mathbf{M}}}\right) d\vec{\theta}$, the so-called marginal likelihood or Bayesian evidence that normalizes the distribution of model M over all n parameters.

The marginal likelihood is often considered the average of the likelihood over the prior space [29, 30, 31]. Consider a multidimensional parameter space for a model M where each point in space has an assigned likelihood. Then, if we add another parameter – another dimension – that is only mildly informative, the likelihood values barely change, and they are now spread across even more space. Thus, less evidence is

present in the overall space, and therefore adding a parameter to a model will result in a lower marginal likelihood, unless that parameter allows the model M to fit the data D significantly better. In this way, the marginal likelihood intrinsically penalizes more complex models, unless a more complex model better matches the data [29, 30, 31].

5.3 AIC is calculated as an estimate of the Kullback-Liebler divergence

The calculation of an information criterion is based on estimating the difference between candidate models and the "true" model (reality). The difference between statistical models or probabilities is measured by Kullback-Leibler (KL) divergence, represented by

$$I(f,g) = \int f(x) \ln \left(\frac{f(x)}{g(x \mid \vec{\theta})} \right) dx \tag{5.7}$$

where f and g represent probability distributions. In a typical KL divergence calculation, $f(x)$ is the reference distribution, while $g(x)$ is the distribution being compared. When using KL divergence in information theoretic model selection, as in Eq. (5.7), we consider $f(x)$ to be "reality," while $g(x \mid \theta)$ is a candidate model parameterized by $\vec{\theta}$, which are estimated from data D (as above). Here, $g(x \mid \vec{\theta})$ is a probability distribution of a model akin to the probability distribution of model M, $\Pr\left(\vec{\theta_{\mathbf{M}}} \mid D\right)$ noted above. I represents information, and because g cannot exactly match reality, $I(f,g)$o represents the information lost when approximating reality f by the model g [25].

The Akaike Information Criterion, or AIC, is an estimate of this KL divergence based on the likelihood function at its maximum point [24]. It also considers the number of parameters in the model $g(x \mid \vec{\theta})$, which reduces bias in the estimate of the KL divergence for that model. Therefore, the AIC is denoted by

$$AIC = -2 \ln \left(L(\vec{\theta_{\mathbf{best}}} \mid D) \right) + 2K = -2 \ln \left(\Pr\left(D \mid \vec{\theta_{\mathbf{best}}} \right) \right) + 2n \tag{5.8}$$

where $L(\vec{\theta_{\mathbf{best}}} \mid D)$ is the likelihood of the best-fitting parameter set and K is the number of parameters in the model (noted as $\Pr(D \mid \theta_{best})$ and n, respectively, to correspond with the Bayesian definitions in **Section 5.2**).

5.4 Notable differences between AIC and Bayesian evidence / posterior probability

While they both aim to rank models from a set of candidates so that one or few can be chosen as the best models, there are several key differences between Bayesian evidence/posterior probability and AIC.

The Bayesian evidence or marginal likelihood is represented by $\int \Pr\left(D \mid \vec{\theta_{\mathbf{M}}}\right) \Pr\left(\vec{\theta_{\mathbf{M}}}\right) d\vec{\theta}$, as noted in **Section 5.2**. Since model optimization to data D assigns a likelihood to each value over the entire parameter

space, and is represented by $\Pr\left(D \mid \vec{\theta_{\mathbf{M}}}\right) \Pr\left(\vec{\theta_{\mathbf{M}}}\right)$, integrating over this thereby integrates over every likelihood value across all evaluated parameter values in each dimension. In this way, the marginal likelihood takes into account every parameter set evaluated during optimization. This represents a more comprehensive score for a model compared to AIC, which as in Eq. (5.8) only considers the best-fitting parameter set $\vec{\theta_{\mathbf{best}}}$. Additionally, as noted in **Section 5.2**, the marginal likelihood penalizes for model complexity (number of parameters). AIC also includes the number of parameters K in its calculation as seen in Eq. (5.8); however, this is again a less comprehensive means of penalizing for number of parameters compared to marginal likelihood.

We then consider the inclusion of prior expectations. At the level of model optimization, both AIC and marginal likelihood can be said to include prior expectations (depending on the algorithm used for optimization). Marginal likelihood includes prior expectations by default, represented by the $\Pr\left(\vec{\theta_{\mathbf{M}}}\right)$ term in the integral above. AIC requires only the best-fitting parameter set, so if optimization was performed via Bayesian principles where prior expectations for parameters were assigned, then these priors could well be said to affect $\vec{\theta_{\mathbf{best}}}$ in the AIC calculation. However, Bayesian principles enable inclusion of prior expectations at the level of model ranking. Often, candidate models are considered equally likely *a priori*, [29, 30, 31] and will thus have a prior probability of $\frac{1}{n}$. When considering H in Eq. (5.6) equal to a particular model M_i, the posterior probability $\Pr(M_i \mid D)$ firstly *requires* a prior probability $\Pr(M_i)$ for its calculation, and once calculated can be compared to that prior probability $\frac{1}{n}$. In this way, Bayesian principles provide a numerical metric to assess how data informed our knowledge about M_i. This plays a role in model averaging as well (see **Section 5.6**). AIC does not have a formal means for including prior probability in its calculation nor in the generation of a posterior probability. Probability in the use of AIC is an interpretation of relative likelihoods across candidate models and is not a true probability (see Eqs. (5.9) and (5.10) in the next section).

Given the incorporation of the full parameter space and prior expectations, as well as the ability to calculate a Bayesian probability that provides insight into knowledge gained from the data, we consider marginal likelihood the optimal means for model ranking and interpretation therein. A specific model selection example where AIC and marginal likelihood/posterior probabilities are compared is detailed in **Section 5.7**.

5.5 Model selection allows us to evaluate which variables or terms have the largest effect on observed data

Model selection is the process of investigating candidate models, each of which may fit the data to a different extent, while prioritizing model simplicity (aka *lex parsimoniae* [189]). That is, the best model must comprise a balance between matching the data exactly and having the fewest variables and parameters required to do so. In linear regression, such as in the following example, this involves investigating which predictor

variables with their fitted coefficients, best match the response variable (**Figure 5.1A**). In kinetic models, such as those with which we aim to capture the behavior of small cell lung cancer (SCLC), this involves investigating which variables and which model terms, with their fitted kinetic parameters, best match the data (**Figures 5.1B** and **5.1C**).

For Bayesian model selection, the hypotheses H in Eq. (5.6) represent candidate models M within the set of all models $\mathscr{M}$, and data D represents the dataset to which each model is fitted. Then, the likelihood of that model M, $\Pr(D \mid M)$, is the marginal likelihood. Calculating the marginal likelihood for all candidate models and normalizing by the sum of this value for all models results in the probability of the model M given the data, $\Pr(M \mid D)$. The higher the probability, the more likely the model best represents the process that generated the data (**Figure 5.1D**, black points).

Models can be ranked by AIC value alone, and in this case a lower AIC value indicates that a model is a better representation of the data. However, those using AIC can scale the models with respect to the minimum AIC value [24, 25]. This results in AIC differences,

$$\Delta_i = AIC_i - min(AIC) \tag{5.9}$$

where $min(AIC)$ is the AIC value of the lowest-scoring (best) model candidate. Using AIC differences, Akaike weights are calculated,

$$w_i = \frac{e^{-\frac{\Delta_i}{2}}}{\sum_{r=1}^{R} e^{-\frac{\Delta_i}{2}}} \tag{5.10}$$

where R is the number of models in the candidate set. Akaike weights represent relative likelihoods of the models given the data, and are interpreted as probabilities [24]. From this perspective, the higher the Akaike weight, the more likely that model best represents the process that generated the data (**Figure 5.1D**, red points).

5.6 Model averaging uses model selection outcomes from all models to demonstrate how the observed data informed the model variables or terms that represent our hypotheses

Probabilities or weights are then used for model averaging. Model averaging is a process used when no clear single best model can be identified after model selection. Model averaging can reveal a model variable or term common to better-performing models, indicating that that the process represented by this variable or term is likely to play a role in the process that generated the data, and allowing the user to identify the parts of a model that best capture the data. In our application, this approach implies a departure from a deterministic single model to a probabilistic understanding of mechanisms, with the probabilities determined from the

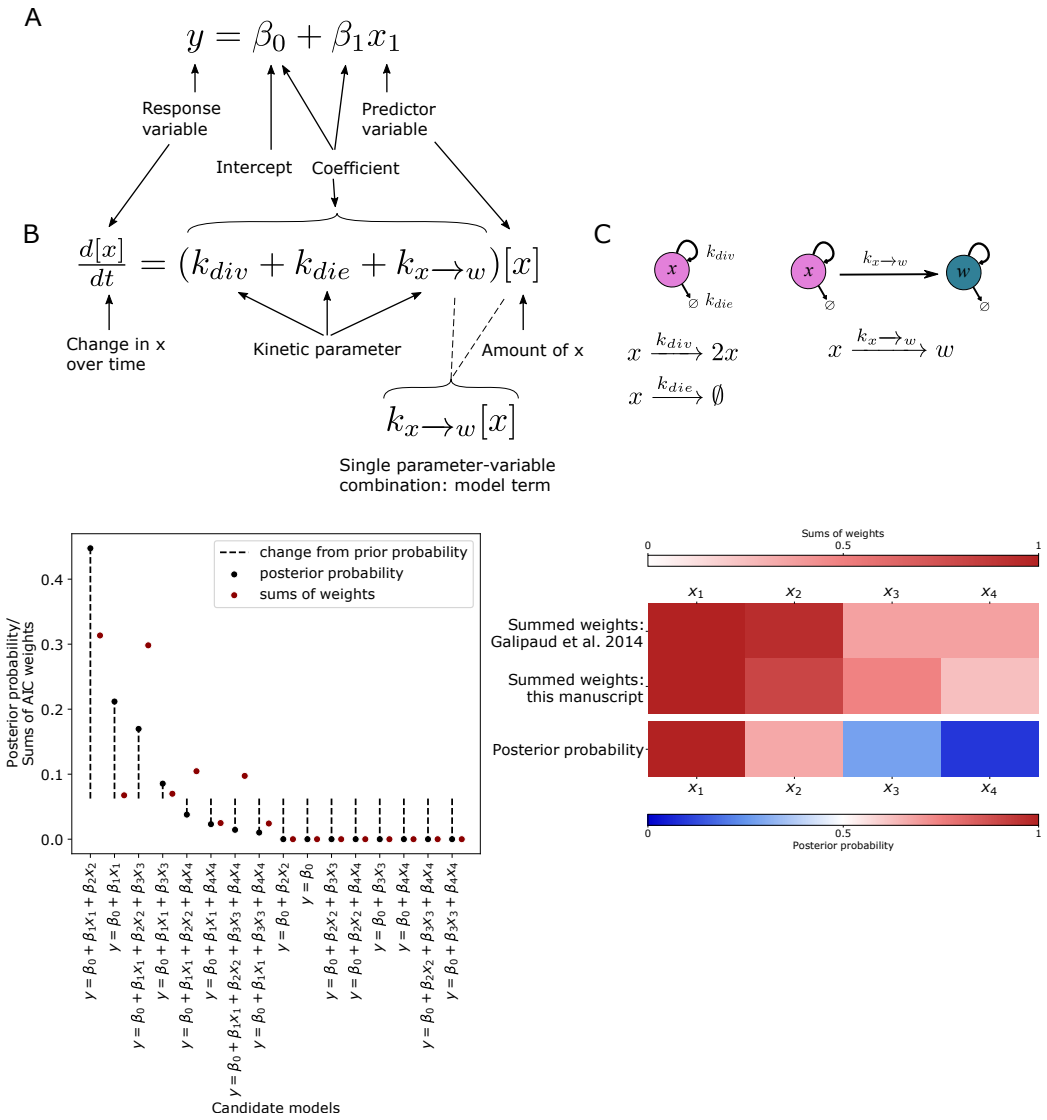


Figure 5.1: Bayesian inference better assesses parameter inclusion in the ”true” model compared to Akaike Information Criterion.

A Aspects of linear regression model assessed by model selection and model averaging.

B Aspects of mass-action kinetics model / ordinary differential equation assessed by model selection and model averaging.

C Schematic representation of the equation in (5.1B).

D Results from our analysis of candidate models using candidate set and data generation code from [1]. Candidate models are arranged along the x-axis by posterior probability. Each posterior probability from our analysis is compared to the AICc weight from our analysis.

E Heatmap for each model averaging analysis based on data generation and candidate models in [1]. Top two rows include summed weights (SW) published in the original [1] manuscript and from our analysis. Top color bar legend corresponds to these two rows, where SW is a value starting at 0 and weights are summed until the final SW value. Bottom row is the posterior probability calculation from our analysis of this data and candidate set. Bottom color bar legend corresponds to this row, where color represents the probability of the variable's inclusion in the ”true” model. Since prior probability starts at 0.5 (white), deeper colors indicate a larger deviation from the prior, with red *vs* blue indicating more likely or less likely probability, respectively.

available data.

In Bayesian model averaging, (BMA) each model is weighted by its marginal likelihood or posterior probability [27] and the likelihood of model with particular term or variable can be used to generate an overall likelihood for that term or variable. Here, the prior probability of a term, which we typically consider to be equally likely vs. unlikely before model optimization, is important. Now, the H of Eq. (5.6) indicates the hypothesis that a particular model term (individual process within the system of interest) is part of the "true" model, and the prior $\Pr(H)$ and likelihood of that term $\Pr(D \mid H)$ can be used to calculate $\Pr(H \mid D)$. Thus we can investigate how the data informed our knowledge about H, from its prior to posterior probability.

To determine whether to predict that a model variable or term is likely to be involved in the system of interest, the Bayes Factor can be used. The Bayes Factor is the ratio of the likelihood of two hypotheses, and when prior probabilities are equal is equivalent to the ratio of the two posterior probabilities:

$$BF = \frac{\Pr(D \mid H_1)}{\Pr(D \mid H_2)} = \frac{\frac{\Pr(H_1|D)}{\Pr(H_1)}}{\frac{\Pr(H_2|D)}{\Pr(H_2)}} \tag{5.11}$$

$$BF = \frac{\Pr(H_1 \mid D)}{\Pr(H_2 \mid D)} \; when \; \Pr(H_1) = \Pr(H_2) \tag{5.12}$$

Here, H_1 may represent the hypothesis that the model variable does play an important role, while H_2 may represent the hypothesis that it does not. The Bayes Factor (BF) between one hypothesis and another can be considered a true difference between the probability of each hypothesis when $10^{-\frac{1}{2}} > BF > 10^{\frac{1}{2}}$. The value of $10^{\frac{1}{2}}$ (≈ 3) is the lowest at which a difference may be determined; or $10^{-\frac{1}{2}}$ ($\approx \frac{1}{3}$) when the higher-probability hypothesis is in the denominator of the BF [178]. In this way, comparing hypotheses via BF can result in one of three outcomes: $\Pr(H_1 \mid D)$ is > 3 times more likely than $\Pr(H_2 \mid D)$, meaning hypothesis H_1 is informed by the data to be significantly more likely than H_2; $\Pr(H_1 \mid D)$ is $< \frac{1}{3}$ times less likely than $\Pr(H_2 \mid D)$, meaning H_1 is informed by the data to be significantly less likely than H_2; or, $\Pr(H_1 \mid D)$ is between $\frac{1}{3}$ less likely and 3 times more likely than $\Pr(H_2 \mid D)$, meaning H_1 is not informed by the data to be significantly different than H_2. We simply denote this last outcome as "H_1 is not informed by the data."

When using information criteria such as AIC, parameter importance analysis is typically performed to investigate how well the data supports a model term, or hypothesis. For each term, the Akaike weights of all models containing that term are summed [24],

$$SW_{H_i} = \sum_k \{w_k \text{ if } M_k \in H_i\} \tag{5.13}$$

where the weight of model M_k, w_k, is summed if M_k is considered part of H_i (contains a model variable

or term that is part of H_i).

The sum of weights (SW) is treated as the probability that a term belongs in the "true" model, thus the biological feature represented by this term plays an important role in the biological system. There does not appear to be an accepted threshold over which the SW for a variable or term indicates it should be included in a model, thus it seems practitioners choose a threshold on an *ad hoc* basis [1].

5.7 Advantages of Bayes-MMI over AIC for model selection and model averaging: continuing example

Despite equivalent goals between BMA and SW, problems have been noted with the latter. To illustrate this point, we consider the example of [1], and compare AIC-based SW to our Bayesian method combining marginal likelihood results of model selection and BMA of terms across candidate models, named Bayes-MMI.

In [1], the investigators generated "ground truth" data using four variables x_{1-4}, assigning differing correlations with the response variable y. x_1 was strongly correlated to response variable y with Pearson correlation coefficient r_{y,x_1} of 0.70, x_2 moderately correlated with r_{y,x_2} 0.20, x_3 weakly correlated with r_{y,x_3} 0.05, and x_4 uncorrelated with $r_{y,x_4} = 0.0$. A variable with no relation to the response variable should not appear in high-scoring models, and should receive a low SW. Galipaud and colleagues performed model selection using all variable combinations including intercept-only, (16 models) calculated AICc, (AIC corrected for small sample sizes) and assessed parameter importance using SW. They found that while x_4 was completely unrelated to the response, it appeared among best-ranked models and had a SW different from zero [1].

To test how Bayes-MMI compares to AIC-based SW methods, we used data simulated via the code provided in [1], and ran a nested sampling analysis via PyMultinest [29, 30, 31] (see **Methods Section 4.1.1**). Galipaud et al. used least squares regression via the R `lm()` function, while for our nested sampling parameter search we used a least squares likelihood function with PyMultinest searching the 5-dimensional parameter space (x_1, x_2, x_3, x_4, and β_0) using uniform prior distributions of 0 to 10 for each response variable x_i and -10 to 10 for the intercept β_0. We used the same 16 candidate models, and obtained similar results to Galipaud and colleagues (**Table B.2**; **Table 5.2** columns "SW: Galipaud" and "SW: this manuscript").

Since nested sampling yields the marginal likelihood, [29, 30, 31], we can leverage Bayesian principles as noted above to explore model hypotheses space and identify the best candidate models. We use candidate model posterior probabilities to perform our own ranking of models (**Figure 5.1D**; **Table 5.1**). We also calculate variable posterior probabilities, (**Figure 5.1E**; **Table 5.2**) using equivalent prior probabilities of variable inclusion: prior probability that the inclusion of a variable in the "true" model is 0.5, or completely unknown (equivalent probability that it is present *vs* that it is not present).

β_0	x_1	x_2	x_3	x_4	k	rank, post. prob	rank, AICc	rank (AICc) in Galipaud et al., 2014	$\ln(Z)$	$\ln(Z)$ error (+/-)	prior prob.	post. prob.
-2.502	0.608	0.139			4	0	0	0	-57.28	0.049	0.062	0.448
-1.219	0.617				3	1	5	4	-58.03	0.043	0.062	0.212
-3.398	0.618	0.129	0.089		5	2	1	1	-58.25	0.054	0.062	0.170
-2.169	0.623		0.090		4	3	4	5	-58.94	0.049	0.062	0.086
-2.729	0.620	0.141		0.007	5	4	2	2	-59.75	0.055	0.062	0.038
-1.462	0.615			0.025	4	5	6	6	-60.24	0.050	0.062	0.023
-3.329	0.619	0.131	0.077	0.002	6	6	3	3	-60.72	0.060	0.062	0.014
-2.300	0.614		0.086	0.026	5	7	7	7	-61.06	0.055	0.062	0.010
3.357		0.164			3	8	8	8	-102.99	0.044	0.062	10^{-21}
4.990					2	9	12	10	-104.79	0.039	0.062	10^{-21}
2.990		0.164	0.038		4	10	9	9	-104.86	0.050	0.062	10^{-22}
3.355		0.164		0.001	4	11	10	11	-105.64	0.051	0.062	10^{-22}
4.602			0.039		3	12	13	12	-106.78	0.045	0.062	10^{-22}
4.901				0.009	3	13	14	14	-107.27	0.046	0.062	10^{-23}
3.043		0.163	0.029	0.004	5	14	11	13	-107.64	0.057	0.062	10^{-23}
4.510			0.041	0.007	4	15	15	15	-109.26	0.052	0.062	10^{-23}

Table 5.1: Summary of nested sampling model selection results on the simulated dataset and model selection problem in Galipaud et al., 2014.

Candidate models are ranked by posterior probability ("post. prob."). Ranking in this Bayesian analysis can be compared to ranking via AICc in this analysis and to the ranking in Galipaud et al., 2014 ("rank, post. prob.", vs. "rank, AICc" vs. "rank (AICc) in Galipaud et al., 2014" columns). First five columns of maximum log-likelihood parameter estimates are part of PyMultinest output. Parameter estimates are reported if present for each of the 16 candidate models. k, total number of estimable parameters; $\ln(Z)$, the natural log of the Bayesian evidence/marginal likelihood (Z), calculated within the prior-bounded parameter space using our least-squares likelihood function; $\ln(Z)$ *error*, the error returned by PyMultinest; prior prob., the prior probability that a model is the "correct" model; post. prob., posterior probability that the model is "correct", calculated as $\frac{Z_i * \Pr(i)}{\sum_j Z_j * \Pr(j)}$.

Firstly, while according to AICc, x_4 can be found in the best-ranked models, when ranking according to posterior probability, x_4 does not appear until the fifth-highest ranked model (**Figure 5.1D**; **Tables 5.1** and **B.2**;[1]). The four highest-ranked models have a cumulative probability of 0.916 (**Figure 5.1D**; **Table 5.1**). Thus, most of the probability that one candidate model is the best model is contained in those models, which more accurately do not include x_4. In fact, x_4 appears only in candidate models whose posterior probability has decreased compared to the prior probability (**Figure 5.1D**, dotted lines above *vs* below posterior probability points; **Table 5.1**). Next, while the SW for x_4 in our AICc analysis is 0.25, the posterior probability that the inclusion of x_4 is supported by the data is 0.09 (**Table 5.2**).

Using Bayesian principles, we consider the prior probability of x_4 inclusion equivalent between presence and absence (probability of 0.5), but after incorporating data, the posterior probability of x_4 inclusion has decreased to a large extent, to a probability of 0.09. We view this as a reasonably low probability, given that x_4 is not at all correlated to y. If H_1 represents "x_4 should be included in the model", and H_2 the opposite, the BF for H_1 *vs* H_2 is $\frac{0.09}{(1-0.09)} = .099 < \frac{1}{3}$, and we determine that x_4 should not be included in the model.

The case of x_3 is also instructive, where the prior probability of inclusion at 0.5 has changed to a posterior probability of 0.28 after incorporation of data ($BF = \frac{0.28}{(1-0.28)} = 0.39$). With the very weak correlation of x_3 to y at 0.05, Bayesian inference has indicated that, at least with the data used, a definitive choice to include or exclude x_3 can *almost*, though not quite, not be made. The SW for x_3 is 0.49, and as such a practitioner certainly might include it in a chosen model, possibly without considering that such a weakly correlated variable could achieve this SW.

Variable	SW: Galipaud et al. 2014	SW: this manuscript	Prior probability	Posterior probability
x_1	1	1	0.5	1
x_2	0.94	0.81	0.5	0.67
x_3	0.37	0.49	0.5	0.28
x_4	0.37	0.25	0.5	0.09

Table 5.2: SW and posterior probability calculations for each model variable.
Prior probability for a variable is set at 0.5, meaning a variable's prior probability for can be calculated per candidate model by dividing 0.5 by the number of models in which the variable appears. Prior probability values only impact posterior probability scores and not SW calculations.

Therefore, we consider using Bayesian analysis based on the marginal likelihood (Bayesian evidence) to be a superior way to perform model selection, and specifically analysis of variable or model term inclusion or exclusion. We thus aim to employ multimodel inference (MMI; model selection and model averaging together) in the context of Bayesian statistics. There have been biological investigations using MMI approaches [190, 191] but, to our knowledge, Bayesian MMI has not previously been applied to our models of interest, cell population dynamics models. Using MMI to assess the posterior probability of different model variables is comparable to Bayesian variable selection, which in biomedicine has been used to determine genetic loci associated with health and disease outcomes in linear models [191]. We find using a Bayesian approach to MMI results in probabilities that biologically relevant model features are (or are not) supported by the data, and is likely relevant to any cancer or developmental biology application and can be used to investigate model variables even in the context of limited or uncertain data.

CHAPTER 6

Unified Tumor Growth Mechanisms from Multimodel Inference and Dataset Integration [1]

6.1 Introduction

A mechanistic understanding of biological processes that explains causal input-output relationships and predicts population behaviors [192] remains a central challenge to all areas of quantitative biology. Mathematical models have become an established practice to specify precise relationships within a biological system, [22] and thereby hypothesize, and subsequently test, the existence of these relationships. For example, multiple mechanistic models of apoptosis execution have been formulated to explore the nature of biochemical interactions that lead to cellular commitment to death, demonstrating that careful model design and suitable data can lead to important biological insights [193, 194, 195]. A more challenging situation emerges when models are formulated for biological processes that are poorly defined or understood, leading to multiple competing, and often juxtaposed mechanistic explanations for a given biological process. For example, in Small Cell Lung Cancer (SCLC), a study of circulating tumor cell-derived xenografts showed that Non-NE-related subtypes could act as a stemlike population ("source"), [196] but archetype analysis of a genetically engineered mouse model tumor showed the Non-NE subtype SCLC-Y acts as an end-state ("sink", rather than source) [197]. Therefore, continued exploration of the hypotheses generated from these works can help elucidate the differences between these and other potential explanations for tumor growth mechanisms.

This phenomenon where multiple mechanistic hypotheses are concurrently proposed but must be assessed with limited data is not restricted to quantitative biology but common to other fields with similar data availability limitations such as ecology [198] climatology [199], and evolutionary biology [200], to name a few. To address this challenge, methods such as model selection and multimodel inference have been proposed using information theoretic scoring techniques such as Akaike Information Criterion, (AIC) with limited success stemming from difficulties with model averaging [1, 26] and the fact that AIC scores do not inherently describe whether a model or features within are informed by the data. More recently, the use of AI and machine learning approaches has given impetus to causal relationship inference [201] but these relationships remain difficult to elucidate, thus underscoring the need for both novel tools for hypothesis exploration, and tools that can be used with rigor in the face of missing data that may not inform all hypotheses.

The model selection process involves candidate model evaluation from a superset of plausible models, relative to a given experimental dataset. A scoring function is then applied to each model usually with

[1] This chapter appears in an updated version as 'Unified Tumor Growth Mechanisms from Multimodel Inference and Dataset Integration' published in *PLOS Computational Biology* and has been reproduced with the permission of the publisher and my co-authors, Leonard A. Harris, Michael A. Kochen, Julien Sage, Vito Quaranta, and Carlos F. Lopez.

the goal to identify the top-scoring model that best represents the process being investigated. A typical approach is to consider a "best" model as comprising the most relevant variables that capture important mechanistic aspects of the explored process, while excluded variables capture process features that are less relevant for the question being explored. However, variables throughout all candidate models can contribute to knowledge about the overall system [24]. In the cases where data is simply less informative for a given set of hypotheses, uncertainty will remain about what constitutes a "best" model, [188], necessitating approaches such as model averaging, where parameter values can be weighted by model probability and then combined into a distribution of likely values. Unfortunately, for information theoretic applications of model averaging, this probability must be weighted and summed across all possible models, which are often not possible to enumerate exhaustively in a biological context.

To address the challenge of employing multimodel inference approaches in the context of biological processes where models cannot be exhaustively enumerated, data may not inform all model evaluation, and model averaging across all models is desired to learn about the system of interest, we introduce Bayesian multimodel inference (Bayes-MMI), a method which combines Bayesian inference with model selection and model averaging. In our approach, Bayesian model selection enables exploration of mechanistic model hypotheses by inclusion or exclusion of species and behaviors that could play a role in each biological process. Application of Bayesian principles in turn reveals the extent to which data informs a given model and its constitutive parameters.

In our system of interest, small cell lung cancer, (SCLC) we integrate the most suitable available datasets and published theories of SCLC cellular biology to enumerate mechanistic hypotheses for SCLC tumor growth. We test the resulting thousands of candidate population dynamics models via nested sampling, comparing candidate model output to tumor steady-state data, applying the principles of model selection and model averaging for a principled and comprehensive assessment of SCLC mechanistic hypotheses. We estimate the probability of each mechanistic hypothesis given the data, generating an interpretation of SCLC tumor growth: highly likely non-hierarchical phenotypic transitions indicating SCLC subtype plasticity, and less likely cell-cell interactions that affect the rate of phenotypic transitions across subtypes. We show how certain aspects of the SCLC model, such as phenotypic transitions and cell-cell interactions related to these, are well informed by the available data, but other aspects, such as tumor initiation and growth rate effects, are not informed. Our approach is generalizable to other biological systems, and as such we suggest a shift away from considering only one "best" model or the variables within and instead propose a move toward Bayesian principles in multimodel inference and a probabilistic understanding of whether each cellular behavior in a model contributes to the behavior of the system under study.

6.2 Existing datasets yield multiple hypotheses in SCLC tumor growth mechanisms

Small cell lung cancer (SCLC) has been denominated a recalcitrant tumor where relapse after treatment is commonplace and the survival prognosis is typically poor. SCLC comprises $\sim$ 15% of all lung cancer cases worldwide and results in $\sim$ 200,000 deaths annually with a 5-year survival rate of less than 10% [7]. Intratumoral heterogeneity is hypothesized to be the main contributor to the natural history of this disease and its morbidity and mortality [7, 9, 202]. SCLC tumors comprise a mix of functionally distinct subtypes of interacting cells, [159, 10, 203], most notably neuroendocrine (NE) and Non-NE. As shown in **Figure 6.1A**, SCLC populations comprise a collection of cellular subtypes within a tumor, identified by differential expression of transcriptional regulators [9].

The overall goal in this work is to computationally explore tumor growth mechanism hypotheses in SCLC. Tumor features that emerge as highly supported by data about the growth mechanism could be used to predict differences in growth across tumors of different genetic backgrounds, responses to *in silico* treatment, or even predict patient-specific tumor behavior after various treatments. Unfortunately, these goals are currently hypothetical, because to build one SCLC model that could be used for these purposes, one would need a unified understanding of the SCLC tumor as a system, and knowledge of SCLC currently exists as nonoverlapping conclusions and hypotheses. We summarize the current knowledge about SCLC tumor growth mechanisms, highlight potential knowledge gaps, and refer interested readers to [9] for a comprehensive review of recent SCLC literature beyond that noted here.

Multiple SCLC subtypes have been identified depending on the experimental model studied, shown in **Figure 6.1A**, as SCLC-A, N, A2, and Y [9]. Other experiments have led to additional proposed phenotypic subtypes, including canonical subtype SCLC-P, but these were not included in our analysis [196, 203, 204, 205, 206]. Our previous work aimed to identify whether all subtypes may be present in a tumor or if only a subset are present, with the result that tumors can be composed of one, multiple, or all subtypes tested [180]. A comprehensive account of initiating SCLC subtype(s) (cell(s) of tumor origin) has not been made, but multiple have been hypothesized in [180, 11, 207].

Studies *in vitro* and *in vivo* have suggested that Non-NE subtype(s) support growth of NE subtypes [10] (**Figure 6.1A**(2)), including vasculogenic mimicking SCLC cells having such supportive effects [203]. The presence of NE subtypes has a dampening effect on Non-NE growth [179]. Recent work has shown that the Hes1-positive (Non-NE) cells supporting NE subtype growth [10] may have upregulated YAP1 [208] (**Figure 6.1A**(2)), and are likely SCLC-Y; otherwise, the referenced studies were completed before the adoption of the canonical subtypes SCLC-A, N, A2, P, and Y, and so it is unclear which of these exactly contribute such effects in each case. NE cells may undergo a transition a toward Hes1$^+$, (likely YAP1$^+$) identity, (**Fig-**

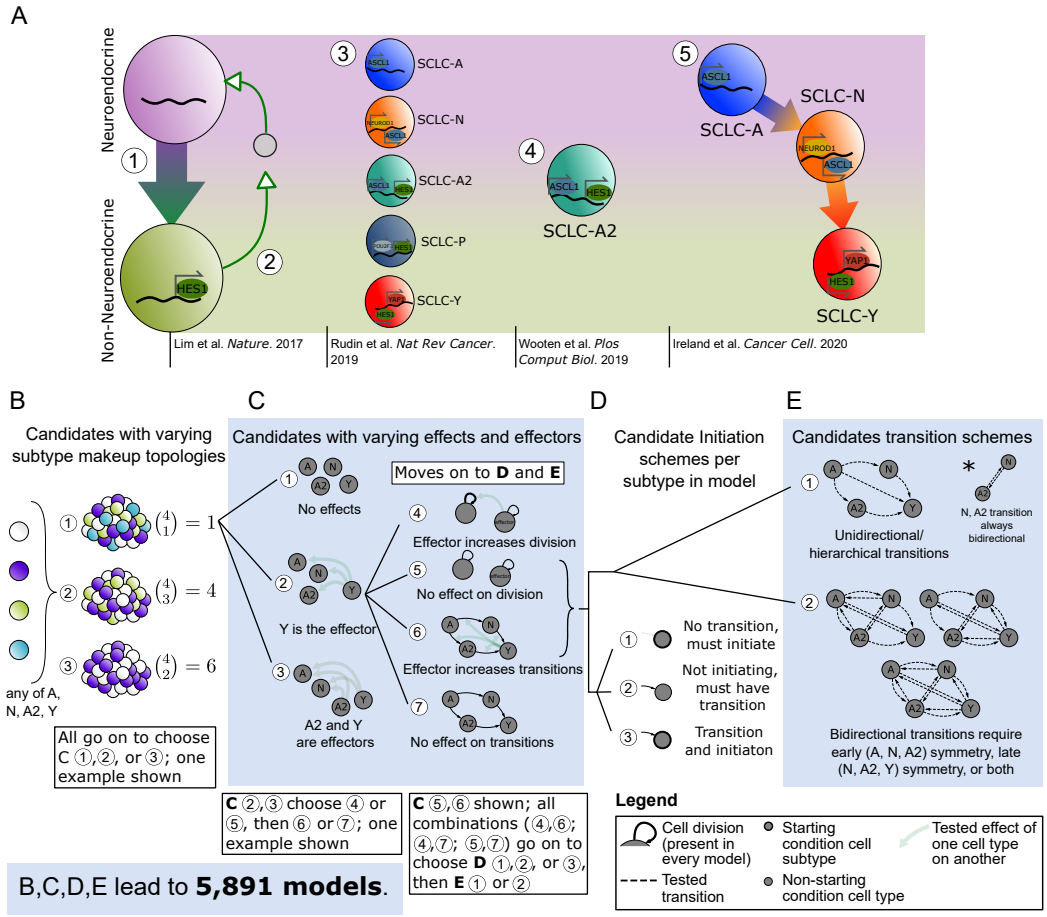


Figure 6.1: Conclusions, hypotheses from literature build mechanistic hypothesis exploration space for tumor growth and development.

A Synthesis of what is currently known about SCLC subtypes, which have been divided into two overall phenotypes, neuroendocrine (NE) and Non-NE, and then further classified into subtypes based on transcription factor expression.

1 NE SCLC cells, which do not express Hes1, transition into Non-NE cells, which do.

2 Hes1$^+$ cells release unidentified factors (gray circle) that support viability and growth of Hes1- cells, and the two Hes1$^+$ and Hes1- populations grow better together rather than separately.

3 Consensus across the field led to labeling SCLC phenotypic subtypes by the dominant transcription factor expressed in that subtype.

4 Subtype with transcriptional signature intermediate between NE and Non-NE, named SCLC-A2.

5 Phenotypic transitions occur in a hierarchical manner from SCLC-A to SCLC-N to SCLC-Y cells.

B Model topologies constructed with 2+ subtypes, with number of combinations per number of subtypes.

C Subtype effect schema, where there are different effectors between candidates and different affected cellular actions.

D Potential initiation schema, where all subtypes in topology must be accessible either as initiating subtypes or via transitions.

E Potential transition schema where, unidirectional transitions are those that follow a hierarchy, and bidirectional transitions must be symmetrical when present.

HES1, Hes Family BHLH Transcription Factor 1; ASCL1, Achaete-scute homolog 1; NEUROD1, neurogenic differentiation factor 1; POU2F3, POU class 2 homeobox 3; YAP1, yes-associated protein.

ure 6.1A(1)) which modulates these Non-NE cells' sensitivity to anticancer drug treatment [10]. Other work found that SCLC-A subtype cells can transition to the SCLC-N subtype and from SCLC-N to SCLC-Y [11] (**Figure 6.1A**(5)). Without the ability to undergo a transition toward a more Non-NE phenotype, tumors were smaller and less aggressive; however, this study did not assess Non-NE or SCLC-Y sensitivity to anticancer drugs [11]. These two landmark studies assessing phenotypic transitions do not assess the same phenotypic transition pathway and thus we cannot compare intermediates, although we hypothesize that the transitions begin with SCLC-A and it seems reasonable to assume the hierarchical pathway ends in SCLC-Y. While our investigations support that SCLC-Y acts as an end state for phenotypic transitions, [197] another study identified that Non-NE subtypes may have stemlike potential, [196] which contrasts with Non-NE or SCLC-Y acting as the end of the hierarchical pathway.

6.3 Multiple mechanistic hypotheses emerge from existing data

Considering the aspects of SCLC tumor growth observed in the previous section, it is clear that no one model exists that could easily recapitulate all datasets. To address this challenge, we explored mechanistic hypotheses in the realm of tumor initiation and composition, phenotypic transitions and their hierarchy, and subtype-to-subtype effects (**Table 6.1**). To select which of these multiple hypotheses to include in a mechanistic model of SCLC without additional findings would introduce bias into the modeling process. Instead, we can address these questions computationally, by including or excluding these behaviors across multiple mechanistic models and evaluating whether model behaviors recapitulate SCLC data; that is, we turn to model selection [24, 27].

For our tumor growth mechanism exploration, we interpret tumor topology, initiation, potential subtype behaviors, etc., as features (model terms) in candidate models (**Table 6.1**). We define model variables as representations of species in the model (e.g., "subtype A"), and model terms as qualitative actions in the model (e.g., "subtype A cell division"), whose rates are denoted by kinetic parameters (e.g., "subtype A division rate") (**Figures 5.1A** to **5.1C**). To fully account for all possible tumor subtype makeup and an exhaustive set of possible explanations, we explored models comprising between two and four subtypes per model (**Figure 6.1B**) and included all possible cell subtype interactions. Growth supportive effects and transition-inducing effects (**Figure 6.1C**) (and growth dampening effects, not shown) are included in some candidate models where, e.g., presence of an effector (supportive cell subtype) increases the rate of growth of a subtype it affects (supports). Subtype A2 has expression features of both NE and Non-NE cells [180], including expression of ASCL1 (seen in NE cells) and HES1 (seen in Non-NE cells) and we therefore assigned A2 NE features in some candidate models and Non-NE in others (**Figure 6.1C**).

To compare a hierarchical system, where a cancer stem cell (CSC) can (re)populate a tumor, and non-

Tumor composition

- Neuroendocrine-classified (NE) subtypes include A (ASCL1$^+$) and N (NEUROD1$^+$), with non-neuroendocrine-classified (Non-NE) subtype Y (YAP1$^+$) and P (POU2F3$^+$). [9]
- Previous work predicted subtype A2, and SCLC-A, SCLC-A2 and SCLC-N have been seen to express ASCL1 [9], [10, 180, 11]. SCLC-A2 expresses Hes1. [180]
- Hes1-positive TKO tumor cells (Non-NE) have YAP1 upregulated. [208]
- It is unclear whether A2 is more NE or Non-NE in character.
- Tumors can be made up of one or more of these subtypes. [9, 180]
- Some subtype combinations have been experimentally verified and others have been predicted using CIBERSORT. [9, 180, 11]
- We do not see SCLC-P in our previous subtype deconvolution. [180]

Evidence of phenotypic transitions

- TKO tumor cells sorted for Hes1-negativity (NE identity) become Hes1 positive (Non-NE identity) when plated with Notch ligand DLL4. [10]
- *Ex vivo* culturing of in situ RPM tumors results in histologic and transcriptional phenotypic changes from NE to Non-NE gene expression over time. [11]
- Transitions between A and A2, A2 and N have not been studied.
- No evidence of SCLC-Y transition to NE identity. [10]

Subtype-to-subtype effects

- Cell viability and division are increased when Hes1-negative cells are plated with Hes1-positive cells, compared to Hes1-negative cells only. [10]
- NE cells suppress Non-NE cell division. [53]
- Application of conditioned media or of isolated exosomes results in a morphological change in TKO-derived NE cell line KP3.*

Table 6.1: Existing data pertaining to SCLC intratumoral heterogeneity and communication.
ASCL1, Achaete-scute homolog 1; NEUROD1, neurogenic differentiation factor 1; POU2F3, POU class 2 homeobox 3; YAP1, yes-associated protein; HES1, Hes Family BHLH Transcription Factor 1. TKO, $p53^{fl/fl};Rb^{fl/fl};p130^{fl/fl}$ tumors [10]; RPM, $Rb1^{fl/fl};Trp53^{fl/fl};Lox-Stop-Lox[LSL]-Myc^{T58A}$ tumors [179]. * *Personal communication (Alissa Weaver, Vanderbilt University)*

hierarchical systems in which phenotypic transitions can occur among multiple or all SCLC subtypes, we include candidate models with several different potential phenotypic transition schemes. Thus, the set of candidate models considered include models without phenotypic transitions, models with transitions that reflect hierarchical transitions observed experimentally [10, 11], and models with reversible transitions, i.e., high plasticity (**Figures 6.1D** and **6.1E**). Unidirectional transitions stemming from one cell subtype indicate a potential CSC, while bidirectional transitions from multiple subtypes indicate phenotypic plasticity. We additionally include tumor initiation from one cell of origin vs. multiple. Thus, candidate models include different numbers of initiating subtypes (**Figure 6.1D**, **Figure C.1G**).

To ensure we built a comprehensive set of candidate models that enable exhaustive exploration of biologically relevant hypothesis space, we combined the potential SCLC behaviors (**Figure 6.1**, **Table 6.1**) with prior knowledge about mechanistic behavior of tumor populations [14, 107, 13, 108, 75, 119]. For example, if there is indeed plasticity in the system, it is likely to be shared among subtypes, leading to symmetrical bidirectional phenotypic transitions across the model (**Figure 6.1E**). We therefore expect that all plausible SCLC tumor growth mechanisms are represented in our candidate model hypothesis space to the best of our knowledge. Accounting for all these different possibilities led to a set of 5,891 unique candidate models, each representing a possible SCLC tumor growth mechanistic hypothesis.

6.4 Bayesian exploration of candidate population dynamics models using experimental data

We use multiple datasets to identify consensus behavior of SCLC and provide a unifying model of tumor growth mechanisms broadly supported by available data (**Figure 6.2A**). These datasets include the triple-knockout (TKO) model ($p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$) of genetically-engineered mouse tumors [10]; equivalent to the RPR2 GEMM [11, 179, 209]), and the RPM model ($Rb1^{fl/fl}$;$Trp53^{fl/fl}$;$Lox-Stop-Lox[LSL]-Myc^{T58A}$ [179]), and cell lines from the Cancer Cell Line Encyclopedia (CCLE) [210] made up largely of the SCLC-A subtype determined in [180] (**Figure 6.2A**). This data provided SCLC subtype-assigned proportions of tumor samples, using the same gene signatures across samples, which had been automatically determined by CIBERSORT from samples of CCLE SCLC cell lines [210] and their consensus clustering class labels [180]. We consider this preferable to our own ad hoc decisions of individual cell subtype identity necessary to assign the required subtype proportions of tumors had we used newer, available single-cell RNA sequencing data. The different datasets represented in **Figure 6.2A** demonstrate differing SCLC tumor makeup, according to the experimental model employed in the study.

To explore the roles of phenotypic heterogeneity and cellular behaviors (cell-cell interactions, phenotypic transitions) on SCLC tumor growth dynamics, we used population dynamics modeling, building on our previous work [211]. Population dynamics models employ a mathematical description of the dynamics within and between heterogeneous subpopulations in an overall population [71, 73]. With such models, researchers can mathematically simulate population growth over time and investigate growth dynamics inherent in the simulations (6.3).

With 5,891 candidate models, (**Figures 6.1B** to **6.1E**) we aimed to determine which one(s) could best represent the SCLC system – in an ideal world, we would determine which model represents reality. However, because we will never know the "true" model, we must evaluate each model probabilistically – which model is most likely to explain the data? Before comparing our 5,891 models to data to determine which is best, we consider each to have the same probability of explaining the data. That is, we assign all models an equal

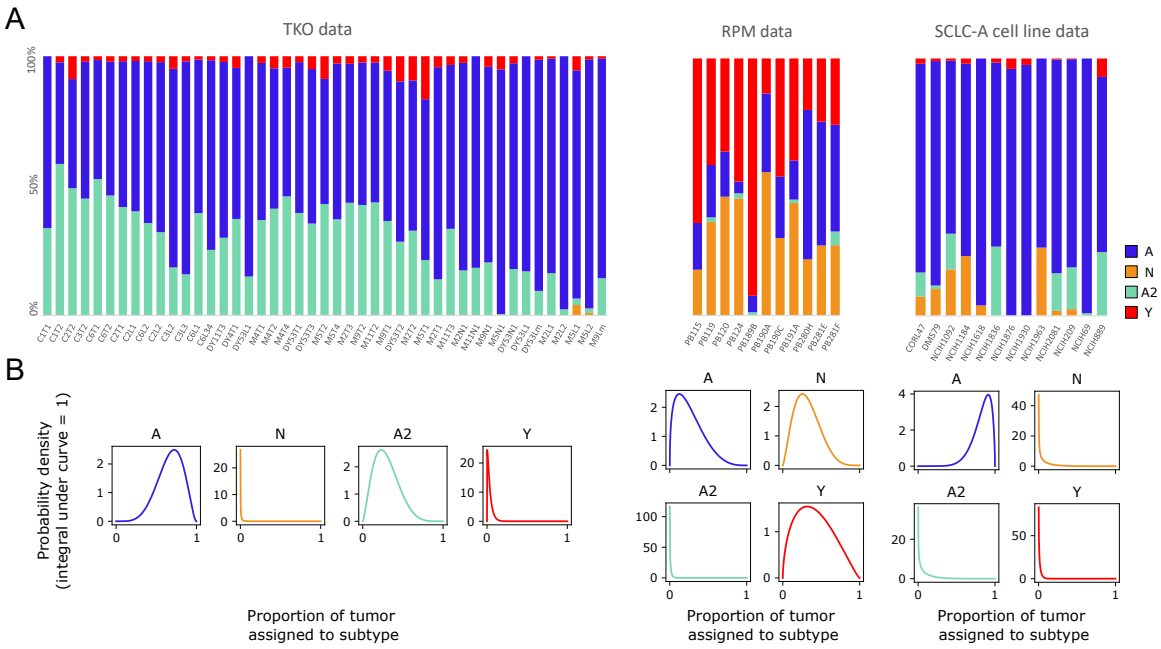


Figure 6.2: Population composition data and probabilistic representation.

A CIBERSORT deconvolution of TKO and RPM genetically engineered mouse model (GEMM) samples (previously published) as well as SCLC-A cell line samples. CIBERSORT was performed on bulk RNA-sequencing data.

B Probabilistic representation of tumor proportion based on mean and standard deviation of proportions across samples within an experimental model; these distributions were then used for fitting models to data.

TKO, $p53^{fl/fl};Rb^{fl/fl};p130^{fl/fl}$ tumors [10]; RPM, $Rb1^{fl/fl};Trp53^{fl/fl};Lox-Stop-Lox[LSL]-Myc^{T58A}$ tumors [179]; SCLC-A cell lines, a subset of SCLC cell lines from the CCLE [210] that we previously assigned as representative of tumors made up largely of the SCLC-A subtype [180].

probability, as this way, we will not bias the exploration of the data toward any particular model or group of models. This is the reason we built a set of models that enables full exploration of the hypothesis space, (detailed above) so that with the addition of data, we can ensure we will study all possibilities and find the best model(s) within that hypothesis space. We compare each candidate model to the data discussed above, and with model optimization, we assess how well its simulations match what we see in that data. After this, each model's probability will be updated, with a model that explains the data better achieving a higher probability than one that does not explain the data as well.

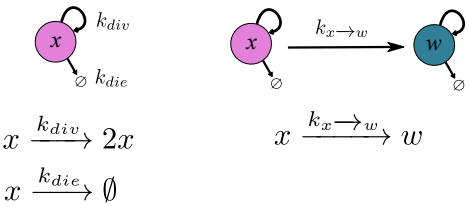
The principles of model selection enable us to assess which candidates are best supported by experimental data, while prioritizing model simplicity [24], [25]. Information theoretic approaches for model selection using the Akaike Information Criterion (AIC) have been used in prior work but do not yield a much-needed statistical understanding of the data. For a didactic demonstration of this, we refer the reader to **Chapter 5**. We therefore employed the marginal likelihood as a more principled means for model ranking and model averaging.

Prior work has estimated the marginal likelihood for kinetic model fitting using thermodynamic integra-

Box 1. Population dynamics modeling and inter-subtype effects

A population dynamics model represents behaviors of species over time, tracking the size of a cell population, rather than tracking individual cells. However, such a model can also include signaling and dependence beteween species (Harris et al., 2019). Cell population changes are represented as reactions, where subpopulation abundances increase or decrease due to varying events (division, death, phenotypic transitions), and the rates of increase or decrease can be affected by the presence of other subpopulations. Here, cell types *x* and *w* are able to undergo division, death, and phenotypic transition at rates k_{div}, k_{die}, and $k_{x\text{-}w}$, respectively (A).

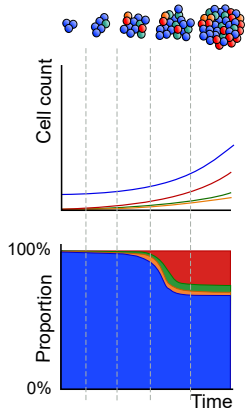
A



$$x \xrightarrow{k_{div}} 2x \qquad x \xrightarrow{k_{x\to w}} w$$

$$x \xrightarrow{k_{die}} \emptyset$$

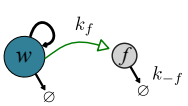
B



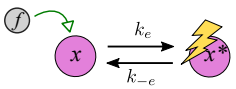
In our case, where species are cells of an SCLC tumor population, each is assigned a subtype identity (A, N, A2, or Y). The model simulates tumor growth over time, (B, top) calculating the increase in subpopulation amount (B, middle), from which tumor subtype proportion may be calculated (B, bottom) and compared at steady state to tumor proportion data. We define steady state as the composition of a tumor based on the relative abundance of each cell subtype in the tumor, without external perturbations. Though tumor growth may continue exponentially in a steady state, the proportion of subpopulations within remains constant (Harris et al., 2019).

Non-spatial cell-cell interactions may also be modeled, by changing the rates of reactions for one subpopulation according to the amount of another subpopulation present. Here, using the hypothesized biology where subtype *w* produces a secreted signal *f* that affects subtype *x*, we calculate the change in reaction rates. In (C), *w* secretes the unknown signaling factor *f* at a rate of k_f, while *f* is degraded at rate k_{-f}. Factor *f* may affect subtype *x* at a rate of k_e, where *x* becomes *x** ("*x* under effect of *f*"). Subtype *x* can then revert back to its unaffected form at a rate of k_{-e}, indicating the rate at which the signaling is completed.

C



$$w \xrightarrow{k_f} w + f \qquad f \xrightarrow{k_{-f}} \varnothing$$



$$x + f \xrightarrow{k_e} x^* + f \qquad x^* \xrightarrow{k_{-e}} x$$

D

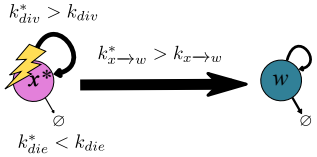
$$\frac{k_{-f}}{k_f} \equiv K_f^{eq} \qquad \frac{k_{-e}}{k_e} \equiv K_D$$

$$\nu_{fate} = \left(\frac{k_{fate} K_D K_f^{eq} + k^*_{fate}[w]}{K_D K_f^{eq} + [w]}\right)[x]_T$$

With k_f the production rate constant and k_{-f} the degradation rate constant, K_f^{eq} is the equilibrium constant for the amount of factor *f* in the system (D). Similarly, K_D is the equilibrium constant related to the on-effect rate constant k_e and the off-effect rate constant k_{-e}. Given these, the rate of a cell fate (division, death, or phenotypic transition) for *x* (ν_{fate}) can be calculated as a function of the population size of the effector cell *w* (D) (Harris et al., 2019) (see Note S1 for calculations).

By assigning the value of k^*_{fate} as more or less than k_{fate} for a particular cell fate, the presence of the effector cell subpopulation can increase or decrease, respectively, the rate of the cell fate for subpopulation *x*. In our population dynamics model, typically effector cells increase division and transition rates and decrease death rates, as shown in (E).

E



$$x \xrightarrow{k^*_{div}} 2x$$

$$x \xrightarrow{k^*_{die}} \emptyset$$

$$x \xrightarrow{k^*_{x\to w}} w$$

Figure 6.3: Box: Population dynamics modeling and inter-subtype effects

tion [176]. In this work we instead use nested sampling, [28], which is computationally more efficient and has fewer limitations with regard to the shape of the probability space traversed during evidence calculation [29, 30, 31]. The nested sampling method was run once for each of the 5,891 candidate models on each of the three experimental datasets, amounting to 17,484 potential interpretations of tumor growth mechanisms. The average fitting time for each model was $\sim$ 19 wall-clock hours, thus necessitating high-performance computing for a complete parameter space exploration of the candidate models.

Each model is thus optimized to our datasets via nested sampling, which explores the full volume of the likely parameter space. Each point in parameter space represents a set of possible parameter values (**Figure C.2**). At each of these points, nested sampling assigns a likelihood value for how well that set of parameter values fits the data. On completion of the algorithm, the output includes the highest-likelihood parameter values. Since each tested point in parameter space is a set of parameter values, the highest-likelihood values for a model are returned by the algorithm as a list of parameter sets (**Figure C.3**). Returning a list of parameter sets rather than one top-scoring set already incorporates Bayesian methodology into the process – each individual parameter has multiple best-fitting values, which can be interpreted as a distribution of parameter values [176] – but with nested sampling we add yet more Bayesian methodology. Having assigned a likelihood to every point in parameter space, nested sampling uses these to calculate one overall likelihood per model, the marginal likelihood, which takes into account parameter fit as well as model simplicity (number of parameters). For more detail on how the marginal likelihood is calculated to incorporate both model fit and size, see **Chapter 5** and **Methods Section 4.1.1**. Finally, with marginal likelihood values for each model in the candidate set, and the candidate model set representing the full hypothesis space with all potential SCLC population dynamics models, we can calculate a probability. Summing the marginal likelihood values, and dividing each individual marginal likelihood by these, results in a model posterior probability, representing a change in probability from pre-model fitting (all models with equal prior probability) to post-model fitting (see **Chapter 5**). We are then able to compare model probabilities and additionally perform model averaging to evaluate kinetic parameter value distributions and probabilities of model variables and terms.

6.5 A small subset of candidate tumor growth models is supported by experimental data

In a Bayesian model selection approach, a more likely model comprises a higher proportion of the probability of the candidate model space (**Figures 6.4A** to **6.4C**). After nested sampling, our results indicate the highest-scoring model for each dataset is $\sim 10^{19}$ times more likely than the lowest-scoring model, and $\sim 10^{3}$ times more likely than the median scoring model. For reference, the smallest comparison between models that is considered significant is $10^{\frac{1}{2}}$ [178] (**Chapter 5**). Therefore, our results indicate that the data used for model fitting has informed our knowledge about the system, because before nested sampling, all models are equally

likely.

Performing nested sampling on all candidate models did not yield a unique best-fitting model for any dataset (**Figures 6.4A** to **6.4C**). We therefore leveraged a multi-model inference approach and calculated a confidence interval (CI) representing a set of best-fitting models per dataset. While a 95% CI is a traditional approach, (**Figures 6.4A** to **6.4C**, orange) we also calculate a "relative likelihood confidence interval," as discussed in [24] (see **Methods Section 4.2.5**). For this relative likelihood CI, we calculate the Bayes Factor (BF) between the highest-scoring model and every other model, using the least strict cutoff of BF $>10^{\frac{1}{2}}$ (as above, and in **Chapter 5**). Even with this permissive cutoff, the relative likelihood CI includes only tens of models per dataset, a large decrease from the initial number of candidates ($\sim 1\%$ or less, **Figure 6.4D**).

In summary, we can determine a subset of candidate models that adequately represent the data, conditional on the fitted parameter sets resulting from the model optimization in nested sampling. Investigating these parameter sets can provide more insight into the similarities and differences between candidate models and their fits within and between datasets. Moving beyond the parameter values assigned to each model term, we wanted to investigate how the data available can inform model terms. If data does not inform model terms and variables and the corresponding fitted parameter rates, it indicates that the mechanistic conclusions we desire to draw from this data using mathematical modeling may require additional or different data.

6.6 High-likelihood model topologies are nonoverlapping between datasets

Given our observation that no one candidate model stands out among other models to explain the experimental data, we employed the multi-model inference technique of Bayesian model averaging (BMA). Briefly, the reasoning behind BMA is that a combination of candidate models will perform better in explaining the data than a single model [188]. In BMA, each model is weighted by its posterior model evidence [27] and the model terms within each model receive an averaged likelihood [191] (**Chapter 5**).

However, before we could investigate model-averaged probabilities, we found that the fitted parameter distribution outcome was dependent on choice of initiating subtype (**Figure C.4**). While we aimed to evaluate initiating subtype possibilities with our Bayes-MMI approach, we cannot do this in a unifying fashion as this choice affects fitted parameters and thus model behavior. Therefore, we turned to the literature to impose stricter constraints about initial subtype conditions, to investigate model variable and term probabilities when kinetic parameter posterior distributions were constrained. As mentioned previously, reports link NE SCLC subtypes and long-term tumor propagation [10, 53] and, in particular, cells of subtype A [11]. We thus used only candidate models with an initiating subtype of A, with or without other initiating subtypes. Since we required that subtype A be an initiating subtype, model structures that do not include subtype A received zero posterior probability (models 3 and 8–10 in **Figure 6.5A**; model topology probabilities without filtering by

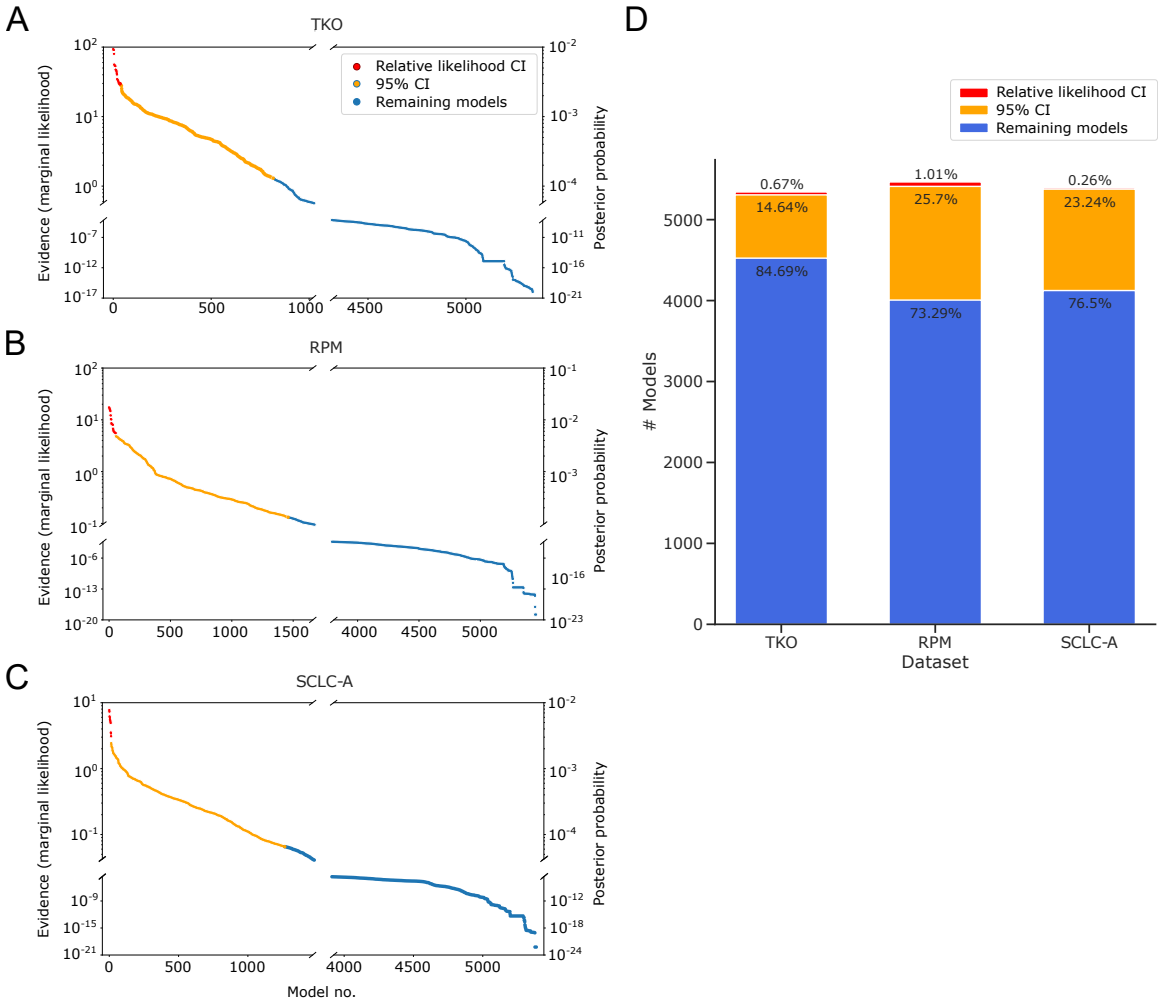


Figure 6.4: Fitting to data and assigning Bayesian evidence separates candidate models into more and less likely.

(6.4A - 6.4C) Evidence values (left y-axis) and posterior probability values (right y-axis) from nested sampling, one point per model, ordered from model with greatest evidence to model with least evidence. Models whose evidence value are within $10^{\frac{1}{2}}$ of the greatest evidence value, the "relative likelihood confidence interval," are colored in red. Nested sampling and evidence calculation is performed per dataset.

A TKO dataset

B RPM dataset

C SCLC-A cell line dataset

D Numbers and percentages of models in the relative likelihood confidence interval, 95% confidence interval, and remaining non-confidence interval models.

TKO, $p53^{fl/fl};Rb^{fl/fl};p130^{fl/fl}$ tumors [10]; RPM, $Rb1^{fl/fl};Trp53^{fl/fl};Lox-Stop-Lox[LSL]-Myc^{T58A}$ tumors [179]; SCLC-A cell lines, a subset of SCLC cell lines from the CCLE [210] that we previously assigned as representative of tumors made up largely of the SCLC-A subtype [180].

initiating subtype are shown in **Figure C.5A**).

We perform BMA across all models for each dataset. As shown in **Figure 6.5A**, all datasets (TKO, RPM, and SCLC-A cell lines) support both two- and three-subtype topologies. Higher probabilities for two-subtype topologies are expected given that nested sampling prioritizes model simplicity and goodness of fit [28]. Statistically, this result suggests that a two-subtype model could be used to interpret the data reasonably well, but it also shows that topologies comprising three subtypes cannot be excluded. Approximately 10% of the probability for the GEMM datasets (TKO and RPM) fall in the three-subtype topology that encompasses the high-probability two-subtype topologies (model 1 for RPM and model 2 for TKO in **Figure 6.5A**). For the models fit to SCLC-A cell line data, most of the probability occurs in the topologies with higher probabilities for the GEMM data. This is reasonable, given that SCLC-A cell line data appears as an intermediate between the GEMMs (**Figure 6.2A**). However, the SCLC-A cell line data also has probability that falls in the A, N, and A2 topology (model 4 in **Figure 6.5A**) – this is the only topology at all likely to represent the SCLC-A cell line data but not at all likely to represent the other two datasets.

We interpret the spread of probabilities across multiple topologies, and that most topologies either are probable as representing either TKO or RPM data but not both, to mean that data coverage from these datasets is not sufficient to support one unifying topology. Therefore, each dataset supports a different representation of SCLC tumor growth given its particular (epi)genetic background and environment. This does not mean that a unifying topology or unifying model of SCLC growth cannot exist, but that the biases underlying the experimental data result in different explanations for tumor growth mechanisms.

6.7 All datasets support alteration of phenotypic transition rates in the presence of N or A2 subtypes

Having established that multiple model topologies can explain tumor growth mechanisms, we explored whether the rates of distinct cellular subtype behaviors were characteristic for given model topologies within each dataset. We therefore compared kinetic parameters across models to learn about dynamic variation between model topologies. We again used BMA to attain this goal, applying the approach to fitted kinetic parameter distributions from nested sampling. In this setting, parameter values from more likely models are assigned higher weights and corresponding parameter distributions are weighted accordingly.

The highest likelihood model topologies (**Figure 6.5A**, blue) for the TKO GEMM data, along with the four-subtype topology, are compared in **Figure 6.5B** (left). Three model terms have significantly different parameter rates across model topologies, all of which are discussed in Section C.3. Here we highlight differences in the A-to-Y and A-to-A2 transitions across model topologies in the TKO dataset: the A-to-Y transition has a slower rate if A2 is present in the population, and only Y affects the A-to-A2 transition, increasing its rate. The mechanistic implication of these observations are as such: A2 may represent an inter-

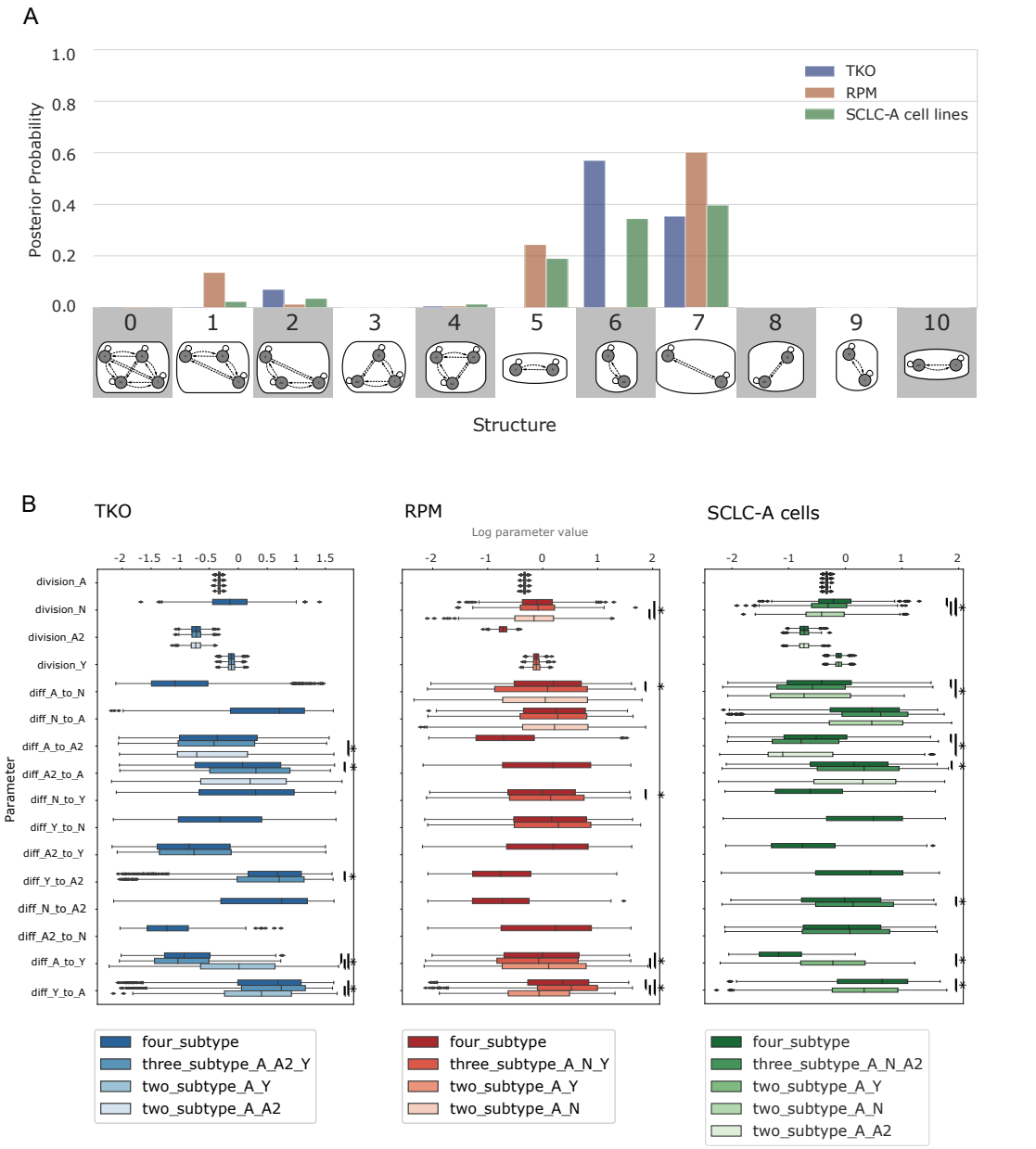


Figure 6.5: Likely model topologies vary across datasets; transition rates vary according to subtype presence in similar ways.

A Hypothesis assessment of model topologies, per dataset. Probability indicates the result of Bayes theorem using equivalent prior probabilities per topology (e.g., 9% probability that one of the topologies in the x-axis best represents a dataset) and evidence values (marginal likelihoods) summed per topology. Model topologies represented by images and corresponding numbers along the x-axis. Posterior probability based on marginal likelihoods of all candidate models that include A as an initiating subtype.

B Division and phenotypic transition parameters for TKO, RPM, and SCLC-A cell line datasets, comparing between higher-probability topologies 6.5A and four-subtype topology per dataset. *(*) indicates significance between samples from BMA parameter distributions at family-wise error rate (FWER) = 0.01, averaged over ten sampling iterations using one-way ANOVA plus Tukey HSD.*

TKO, $p53^{fl/fl};Rb^{fl/fl};p130^{fl/fl}$ tumors [10]; RPM, $Rb1^{fl/fl};Trp53^{fl/fl};Lox-Stop-Lox[LSL]-Myc^{T58A}$ tumors [179]; SCLC-A cell lines, a subset of SCLC cell lines from the CCLE [210] that we previously assigned as representative of tumors made up largely of the SCLC-A subtype [180].

mediate subpopulation in the tumor that is longer-lived, and will only slowly transition to Y. In the topology with A, A2, and Y (**Figure 6.5A**, structure 2; **Figure 6.5B** left, third-darkest blue), the A-to-A2 transition takes up more of the flux in the network. Additionally, the N-to-Y transition is faster relative to the A2-to-Y transition (**Figure C.5B**), suggesting that N is a shorter-lived intermediate in the A-to-N-to-Y transition. This result aligns with previous experiments [11] where N was identified as a short-lived state in the A-to-N-to-Y transition. We therefore predict that A2 and N are involved in regulating the relative abundance of, and flux between, A and Y in the tumor.

We also compared the highest likelihood model topologies (**Figure 6.5A**, red) for RPM-fitted models, as well as the four-subtype topology (**Figure 6.5B**, middle). Five parameter rates are significantly different across model topologies, and we highlight again the A-to-Y transition and here the A-to-N transition (see Section C.3 for discussion of the remaining significantly different parameter rates). The same A-to-Y transition affected in the TKO-fitted models is affected in the RPM model in the same way (reduced rate via an intermediate, in this case N), and the A-to-N transition is affected by Y similarly to the A-to-A2 transition affected by Y in the TKO data. These similar effects occur despite the experimental data used for BMA being different. We thus predict that N and A2 are modulating the transition between, and relative abundance of, A and Y. Unlike in the TKO data, when A2 is present the flux through the system spends more time in the N subtype, with more frequent transitions to N and less frequent transitions to Y. We predict that while N may be a shorter-lived intermediate than A2, A2 regulates the flux from A-to-N-to-Y.

Next, we compared the highest likelihood model topologies (**Figure 6.5A**, green) for the SCLC-A cell line data and the four-subtype topology (**Figure 6.5B**, right). Six parameter rates are significantly different across model topologies, five of which recapitulate rate alterations based on the presence or absence of different subtypes in TKO or RPM datasets, including the rate alterations discussed above (see Section C.3 for more detail).

In summary, BMA enabled us to determine that the A-to-Y transition is regulated in a similar manner for the RPM, TKO, and SCLC-A datasets and that the A-to-N and A-to-A2 transitions are regulated similarly in each dataset. Using the higher likelihood model topologies and model-averaged parameter sets, we can infer features of the SCLC tumor generally, despite disparate datasets. Finding the same or similar effects on kinetic parameter rates across independent datasets lends more weight to these predictions about how the N and A2 subtypes may regulate the system flux from A to Y through intermediates and is an advantage of our methodology using Bayes-MMI to work toward a unifying model of SCLC tumor growth based on multiple datasets.

6.8 Model analysis supports a non-hierarchical differentiation scheme among SCLC subtypes

We have considered candidate models (**Figure 6.4**), model topologies (**Figure 6.5A**), and kinetic parameters (**Figure 6.5B**) to explore tumor growth mechanisms in SCLC. There is compelling experimental evidence for multi-subtype tumor composition, which implies multiple potential growth mechanisms [196, 10, 11, 179, 212]. We therefore focused on model topologies 1, 2, and 4, which are three-subtype topologies with detectable probability (>1%) (**Figure 6.5A**), along with the four-subtype model. Using these, we integrate candidate models, topologies, and kinetic parameters, investigating phenotypic transitions between subtypes, whether the presence of certain subtypes affects the behaviors of other subtypes, and if so, which subtypes bring about the effects (**Table 6.2**). We conclude by proposing a unifying four-subtype model of tumor growth in SCLC, aiming to represent with one model the varying growth mechanisms accessible across datasets.

We investigate the posterior probabilities, and therefore posterior odds, of each model term (see **Methods Section 4.2.7**). Despite different posterior probability values (**Figure 6.6A**), the probabilities of model terms across datasets were similar in their trends: across all three-subtype topologies, phenotypic transition probabilities were all more than $\frac{1}{2}$ (**Figure 6.6A**, red squares). While some probability values were poorly informed (light red), (probability between $\frac{1}{2}$ and $\frac{2}{3}$), more were informed by the data (deep red) ($\frac{2}{3}$ or more). Conversely, probabilities of Non-NE effects on the growth or transitions were all less than $\frac{1}{2}$ (**Figure 6.6A**, blue squares). Some probability values were poorly informed, (light blue) (between $\frac{1}{3}$ and $\frac{1}{2}$) and others were informed (deep blue) ($\frac{1}{3}$ or less) with the addition of data.

Overall, the data suggests that Non-NE effects on transition rates of N-to-Y, or A2-to-Y, are unlikely, (**Figures 6.6B** to **6.6D**, deep blue) regardless of whether "Non-NE" defines only the Y subtype, or both A2 and Y are Non-NE (**Figure 6.1C**). Inter-subtype effects on SCLC phenotypic transition rates have not previously been studied and our analysis predicts that at least effects on "late transition" (**Figure 6.6A**), those interactions affecting N-to-Y or A2-to-Y, are unlikely to exist. By contrast, transitions involving A-to-N, N-to-A, A2-to-A, N-to-Y, Y-to-N, A2-to-Y, Y-to-A2, N-to-A2, A2-to-N, A-to-Y, and Y-to-A had posterior probabilities informed by the data (**Figures 6.6B** to **6.6D**, deeper red). We interpret these results as transitions being likely, i.e., our degree of belief in these transitions has increased. Investigating initiating events via one or multiple cells of origin across the candidate models, we find that from equal prior probabilities of 6.67% per initiating subtype(s) (**Figure C.1G**) the posterior probabilities are not significantly altered, being between 1% and 15% (not shown). Thus, initiating subtype events were poorly informed by the data. Additionally, analyzing specific model terms, inter-subtype effects on NE subtype growth, inter-subtype effects on transition rates between A and N, or A and A2, and the A-to-A2 transition, were also poorly informed by the data (**Figures 6.6B** to **6.6D**, light blue, light red).

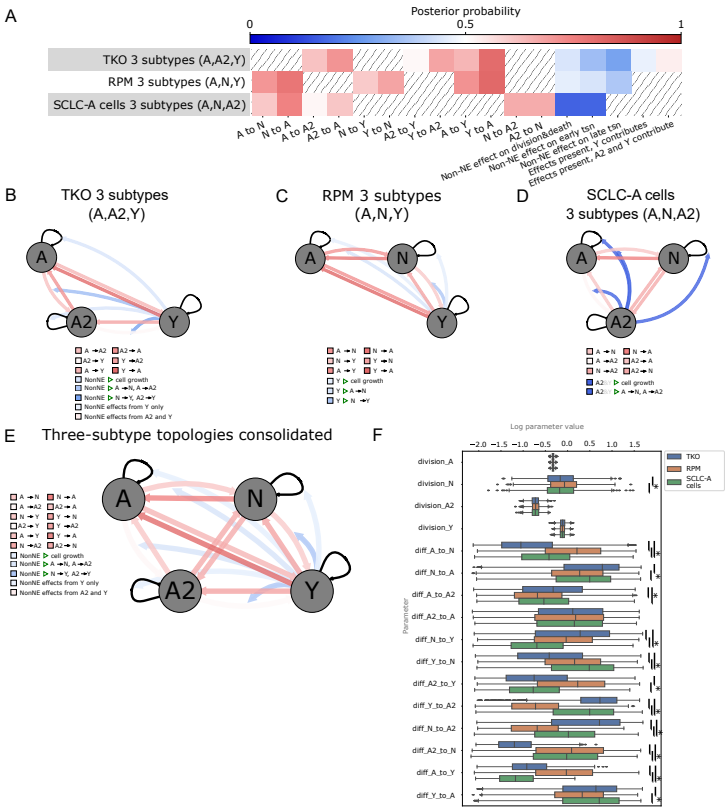


Figure 6.6: Across datasets, multimodel inference indicates likely bidirectional phentoypic transitions, suggesting high SCLC phenotypic plasticity.

A Heatmap for high probability three-subtype topologies for each dataset (rows), all models initiated by A +/- other subtypes. Color represents the probability of each cellular behavior (column). Since prior probability starts at 0.5 (white), deeper colors indicate a larger deviation from the prior, with red vs blue indicating more likely or less likely, respectively.

B - 6.6D. Model schematics with each cellular behavior represented by edges coming from or moving toward each cell subtype, (gray circles) growth rates, (self-arrows) or transitions (arrows between gray circles). Edge colors correspond to colors for that behavior in the heatmap in (6.6A). 6.6B, Top-scoring three-state topology for TKO dataset

C Top-scoring three-state topology for RPM dataset

D Top-scoring three-state topology for SCLC-A cell line dataset

E Schematic of consolidated model behaviors, drawn from each dataset's high-probability three-subtype topology results (6.6B-6.6D). When multiple dataset results included different posterior probabilities for a model feature, the one closest to 0.5 was chosen (most conservative). Edge colors correspond to posterior probabilities, with intensity of colors representing information gained from data, as in (6.6A-6.6E).

F Parameter fitting results (part of the nested sampling algorithm) for four-subtype topology models initiated by A +/- other subtypes, across datasets.

tsn, transition (e.g., subtype transition). TKO, $p53^{fl/fl};Rb^{fl/fl};p130^{fl/fl}$ tumors [10]; RPM, $Rb1^{fl/fl};Trp53^{fl/fl};Lox-Stop-Lox[LSL]-Myc^{T58A}$ tumors [179]; SCLC-A cell lines, a subset of SCLC cell lines from the CCLE [210] that we previously assigned as representative of tumors made up largely of the SCLC-A subtype [180]. (*) indicates significance between samples from BMA parameter distributions at family-wise error rate (FWER) = 0.01, averaged over ten sampling iterations using one-way ANOVA plus Tukey HSD.

Informed high posterior probabilities

- Simulated tumors appeared more likely to be made up of fewer than four subtypes, indicating the model selection algorithm's preference for parsimony (fewer subtypes to explain the same data).
- Phenotypic transitions A-to-N, N-to-Y, A-to-Y had posterior probabilities between 60% and 71%
  - Posterior odds for these are between 1.5 and 2.5.
- Transitions are bidirectional: phenotypic transitions N-to-A, A2-to-A, Y-to-N, Y-to-A2, and Y-to-A, had posterior probabilities between 61% and 81%
  - Posterior odds: $\sim$ 1.5 to 4.
- Transitions between N and A2 (N-to-A2, A2-to-N) had posterior probability 66%
  - Posterior odds: $\sim$ 2

Informed low posterior probabilities

- Low probability of effects that lead to more/quicker phenotypic transitions from NE to Non-NE subtypes, posterior probabilities between 17% and 43% (average 32%)
  - Posterior odds: $\sim$ 0.5
- In SCLC-A cell line datasets, trophic effects, where Non-NE subtypes increase NE division and decrease NE death, had posterior probability 16%
  - Posterior odds: 0.19

Uninformed posterior probabilities

- Initiating / early post-initiation number of subtypes: out of 15 model initiation options, each probability was between 1% and 19%
- Phenotypic transitions A-to-A2 and A2-to-Y had posterior probabilities between 52% and 62%
  - Posterior odds for these are between 1.08 and 1.6.
- In TKO and RPM datasets, trophic effects, where Non-NE subtypes increase NE division and decrease NE death, had posterior probabilities between 45% and 46%
  - Posterior odds: 0.79 to 0.85.
- In TKO three-subtype models (only three-subtype model with both types of effects) A2 and Y effects are 53% probable vs. Y only effects at 47% probable
  - Posterior odds for A2 and Y effects is 1.13, posterior odds for Y only effects is 0.89.

Table 6.2: Probabilities after hypothesis exploration using Bayesian multimodel inference.

Finally, to consolidate phenotypic transitions and cell-cell interactions into a unifying mechanism for SCLC tumor growth, we integrated model probabilities from each of the three-subtype topologies for each dataset into one model (**Figure 6.6E**). Briefly, phenotypic transition probabilities were chosen from the models least informed by the data in an attempt to make conservative predictions (see **Methods Section 4.2.10**). Model-averaged parameter rates were visually compared (**Figure 6.6F**) to ensure that they were within reasonable bounds and that transition rates relate to each other between datasets similarly to our analyses using

high-probability topologies (**Figure 6.5B**). Values from the consolidated probabilities (**Figure 6.6E**) are those reported in **Table 6.2**, along with posterior odds to compare one hypothesis to its opposite.

Taken together, these results provide insight not only into what model terms and variables the data is able to inform, but SCLC tumor behavior as well. Knowledge of trophic effects provided by Non-NE cells to the benefit of NE cells was not provided by this particular data; therefore, we cannot use it to understand this behavior. However, we were able to gain knowledge about the likelihood of phenotypic transitions, in fact indicating that nearly all options for phenotypic transitions are likely to exist. We interpret this as high SCLC plasticity, supporting a non-hierarchical differentiation scheme where tumor population equilibrium is achieved through any phenotypic transitions (**Figure 6.6E**). It is also clear that consolidating the results across different tumor types is an important step in order to achieve a broader view of the SCLC tumor as a system rather than as one particular experimental model.

6.9 Discussion

The experimental data used for this analysis favors two-subtype topologies as higher-probability candidates. This is not surprising, because nested sampling prioritizes simpler models. Despite this finding, mounting experimental evidence supports multi-subtype tumor composition and these data have been interpreted in the context of multiple existing phenotypes [196, 10, 11, 179, 212]. In fact, previous work from our labs suggests that the tumor genetic background dictates possible phenotypic subtypes within a tumor population, and that phenotypic transitions likely mediate the ability of tumor cells to achieve these phenotypes [211, 213, 20]. We hypothesize that with additional timecourse and perturbation data, the topology likelihood for tumor growth mechanism could be shifted toward three- or four-subtype models. This is consistent with studies of tumor dynamics from cancer broadly, both with regard to phenotypic transitions toward or away from a rarer subpopulation [13], [44, 15, 45] and changes in the proportion of phenotypic subpopulations after a perturbation. Often this perturbation is the application of drug treatment [64, 69, 68, 60, 214, 67], but may also be changes in microenvironment or related factors [11, 215].

The results presented here also provide strong evidence for phenotypic plasticity in SCLC tumors, based on the higher likelihood for most phenotypic transitions tested, regardless of differentiation hierarchy. With a more plastic and less stem-cell based phenotypic equilibrium, instead of rare remaining stem-like populations leading to regenerate a tumor after treatment, we hypothesize that any SCLC subtype that remains post-treatment can lead to tumor regeneration and subsequent treatment resistance, patient morbidity and mortality. While hierarchical phenotypic heterogeneity vs. phenotypic plasticity has not been experimentally studied in SCLC, we predict that plasticity is highly likely, and extremely important in the growth and evolution of the SCLC tumor. It is of particular interest to compare phenotypic plasticity and the prevalence of non-

hierarchical transitions in treated vs. untreated tumor samples, as treatment is likely to alter the mechanisms by which tumor population equilibrium is maintained. Time-course experiments with surface marker labeling or live-reporter imaging can resolve and provide confirmation for bidirectional phenotypic transitions, which are crucial to understand in order to battle SCLC treatment resistance.

The invasive or metastatic potential of the SCLC tumor is known to be increased by Non-NE subtypes [203, 216]. It is unclear whether the conclusions and predictions presented here apply to SCLC in the invasive or metastatic setting, but future work will include model additions to place the tumor growth in a physiologic context that includes both the tumor in situ and during invasion.

We believe a shift from information theoretic multimodel inference toward a Bayesian approach, enabling investigation of optimal model(s) with identifiable parameters for a particular dataset, will benefit modeling in systems biology. The methodology employed herein incorporates model selection and model averaging into a multimodel inference framework, followed by Bayesian analysis to identify not only whether a hypothesis investigated via mechanistic modeling is or is not likely, but how likely (and thus how informed by the data) that hypothesis is. Understanding which hypotheses are informed by the data is especially important given variability between data in investigations of the same systems, such as a particular tumor type. It is difficult to attain a consensus model since investigators use varying experimental models within the same physiologic or disease process and thus may draw nonoverlapping conclusions, building parts of a picture but not a whole. Striving for the whole picture, via principled statistical analysis, to be followed by experiments based on informed model predictions, will advance cancer research and lead to better treatments.

CHAPTER 7

Evaluating small cell lung carcinoma time-course patient-derived xenografts reveals phenotype switching with treatment and chemotherapy-stable subpopulations [1]

7.1 Introduction

Small cell lung cancer (SCLC) is the most lethal and aggressive subtype of lung carcinoma. Taken by itself it is the 8th leading cause of cancer death [2]. Unlike many aggressive solid carcinomas, however, SCLC has a very high initial overall response rate (ORR) of up to 80%, even in metastatic cases, on initial therapy. Up to 10% of patients will experience a complete response with no radiologic signs of carcinoma. However, 95% of cancers in these patients still recur, and despite recent advances the median survival of patients with SCLC remains less than one year [12].

The natural history of this disease is likely is related to a high degree of functional heterogeneity in SCLC [7]. Various cell subtypes have been found to coexist within the tumor, providing trophic support, signaling factors that may increase the ability to undergo phenotypic transitions, and increasing metastasis [10, 53, 159]. Molecular pathways altered in SCLC are known to affect factors related to stem cell biology, cell fate decisions, and lineage plasticity [7], and it has been shown that phenotypic plasticity plays a key role in SCLC tumor behavior [197, 209, 196]. Examinations of SCLC to date have suggested different predominant subtypes of SCLC, and these are largely based on transcription factor expression revealed by bulk or single-cell sequencing [10, 197, 180, 196, 9, 208, 203, 204, 205, 206, 11, 212]. Existing work on SCLC heterogeneity has suggested that multiple subsets and mechanisms of resistance may be present simultaneously [12, 217].

To dig deeper into these phenotypic subsets and mechanisms of resistance, here we move away from the transcriptional landscape into protein expression and modification in SCLC PDX tumors. We use data generated by Dr. Jonathan Lehman, using cytometry time-of-flight (CyTOF) with a specifically chosen panel of antibodies for SCLC, applied to PDX tumors receiving subtotal chemotherapy treatment followed by serial core sampling. This enables the tracing of tumoral evolution of subpopulations over time. Analysis of this data has enabled the identification of multiple subpopulations with variable frequency based on chemotherapy treatment and include subpopulations which are chemotherapy stable, potentially contributing to observed tumoral recurrence via a stem cell-like mediated process similar to that observed in lung wound healing.

[1] This chapter includes work contributed by, and is included with permission of, Jeremy Staub and Dr. Jonathan Lehman.

7.2 Experimental background: mass cytometry data from two PDX time-courses

The data provided for these analyses is from mass cytometry, or Cytometry by Time-Of-Flight (CyTOF). In mass cytometry, antibodies have heavy-metal isotopes as reporters, and can be measured by the mass cytometer via their mass-to-charge ratio similar to mass spectrometry [218]. We are interested in mass cytometry rather than single-cell RNA sequencing (scRNAseq) for several reasons. While scRNAseq is relatively unbiased, enabling the measurement of a wider range of markers, (RNA transcripts) it is a relatively shallow and sparse measurement (**Figure 7.1B**). Few transcripts may be captured per marker, and the process for single-cell isolation and library preparation can result in dropouts that artificially decrease transcript counts [219]. Mass cytometry represents a higher-throughput method for single cell measurements without the aforementioned decreased marker counts, enabling increased depth of measurement [220]. However, markers must be specifically selected for target detection, so measurements cannot be as broad as scRNAseq (**Figure 7.1B**). Provided mass cytometry data measured 43 markers, whose names and description can be found in **Table 7.1**.

Marker	**What it may indicate**	**Marker**	**What it may indicate**
CD45	Immune cells	ASCL1	Neuron development, SCLC-A subtype
Rhodium	Positivity means dead cells	PDGFRβ	Endothelial cells, fibroblasts
CD66b	Granulocytes	p-STAT3	Self-renewal of embryonic stem cells
CD16A/B	A: Mast cells, macrophages B: Neutrophils	BMX	(A BTK) related to actin reorganization, migration, adhesion
CD8	Cytotoxic T cells	YAP1	Hippo pathway; proliferation
CD14	Macrophages	CD66e	Epithelial cells
CD4	Helper T cells	S100B	Glial-specific, expressed by astrocytes
CD19	B cells	CD49F	An integrin; adhesion, surface signaling
EPCAM	Epithelia, epithelial neoplasms	Oct3-4	Embryonic stem cells
β-CATENIN	Wnt signaling; adhesion; transcription	CD44	Cancer stem cell marker in some cancers; homing, metastasis
CK8/18	Secretory epithelia, reacts w/ NE tumors	MET (Hepat. GFR)	Organogenesis, wound healing – seen in metastasis, angiogenesis in cancer
NEUROD1	Neurogenesis, SCLC-N subtype	KI67	Active cell cycle (proliferation)
HES1	Maintains neural stem cells	CD24	B cells, neuroblasts
c-CASP3	Apoptotic signaling	GFAP	Astrocytes
SYP	NE cell / NE neoplasm marker	p-ERK	Growth / cell cycle
p-STAT6	IL-4, IL-13 signaling (inflammation, effector T cells)	NKX2-1	Developing thyroid, respiratory epithelium, diencephalon
SOX2	Stemness, pluripotency	CD90	Axonal processes
p-S6	Ribosomal protein; indicates growth	HLA-DR	MHC class II surface receptor
p-AKT	Survival / growth signaling	CD298	N/K transport ATPase
SOX1	Maintains neural progenitors	c-MYC	Cell cycle progression
CD56	Neural cell adhesion	CD11b	An integrin; innate immune marker
h-H3	Denotes human cell (from PDX tumor)	Mouse MHC class I	Denotes a mouse cell (the host for the PDX tumor)

Table 7.1: Mass cytometry panel for **Figures 7.2** and **7.3**

p-, phosphorylated; c-, cleaved; Hepat., hepatocyte; GFR, growth factor receptor; BTK, Bruton Tyrosine Kinase; h-H3, human histone 3

Two PDXs were used to generate the provided data, which includes four mouse tumors, two from each PDX. There are 3 conditions for each tumor: a pre-treatment tumor core sample, peri-treatment tumor core sample, and the terminal harvest entire tumor (**Figure 7.1A**). Two tumors are patient-derived xenografts of SCLC tumor LX22, [221] a "variant" SCLC tumor considered mainly SCLC-N, or having high NEUROD1 transcript expression [222]. One of these tumors, LX22 sample 104, was untreated, and thus each timed sample represents the evolution of an SCLC variant tumor over time with no perturbation. The other, LX22 sample 113, was treated with standard-of-care (SOC) cisplatin and etoposide, and thus each timed sample represents the evolution of an SCLC variant tumor during treatment. LX22 tumor sample 113 had gone through four cycles of chemotherapy at the time of terminal harvest. The remaining two tumors are patient-derived xenografts of SCLC tumor LU73, a "classical" SCLC neuroendocrine tumor [2]. Each LU73 sample (121 and 368) was treated with SOC cisplatin and etoposide, and thus their data can be considered as replicates of a classical SCLC tumor receiving treatment. LU73 samples 121 and 368 also underwent four cycles of chemotherapy treatment by the time of terminal harvest.

Given these four samples, we can compare SCLC tumors over time in different ways: one comparison with and without treatment, and another comparison between treated replicates over time (**Figure 7.1A**). This is a novel means of evaluating SCLC, where many studies compare untreated to post-treatment samples [10, 53, 196, 205, 212, 223]. Here, we not only include a time course of an untreated tumor, (LX22 tumor 104) but include three timepoints instead of simply "before" and "after" treatment. Treatment was also performed as it would be for an SCLC patient, with the chemotherapeutic dose for the mouse at a level corresponding to the level a patient would receive. This is in contrast to other cancer research in general, where very high doses are given in PDX treatment. With such data, we can evaluate the evolution of each tumor over time in the mouse with therapy at a human-relevant level, approximating how a human tumor would develop over time and with treatment.

7.3 UMAP and Leiden clustering applied to mass cytometry data shows density changes over time, and which clusters increase or decrease with treatment

We were first interested to assess the "landscape" of each tumor as well as general changes over the time course. From the mass cytometry raw data, each sample was bead-normalized to account for signal variation over time [224], and then filtered to include only reads that are human (positive for human histone 3), not mouse (negative for mouse MHC class I), and complete cells (not fragments) (**Figure 7.2A**, left). We then used UMAP dimensionality reduction and Leiden clustering to partition the data into well-connected clusters [225] (**Figure 7.2A**, middle and right). We performed dimensionality reduction and clustering per PDX, so

[2] Personal communication between Dr. Jonathan Lehman and Stemcentrx

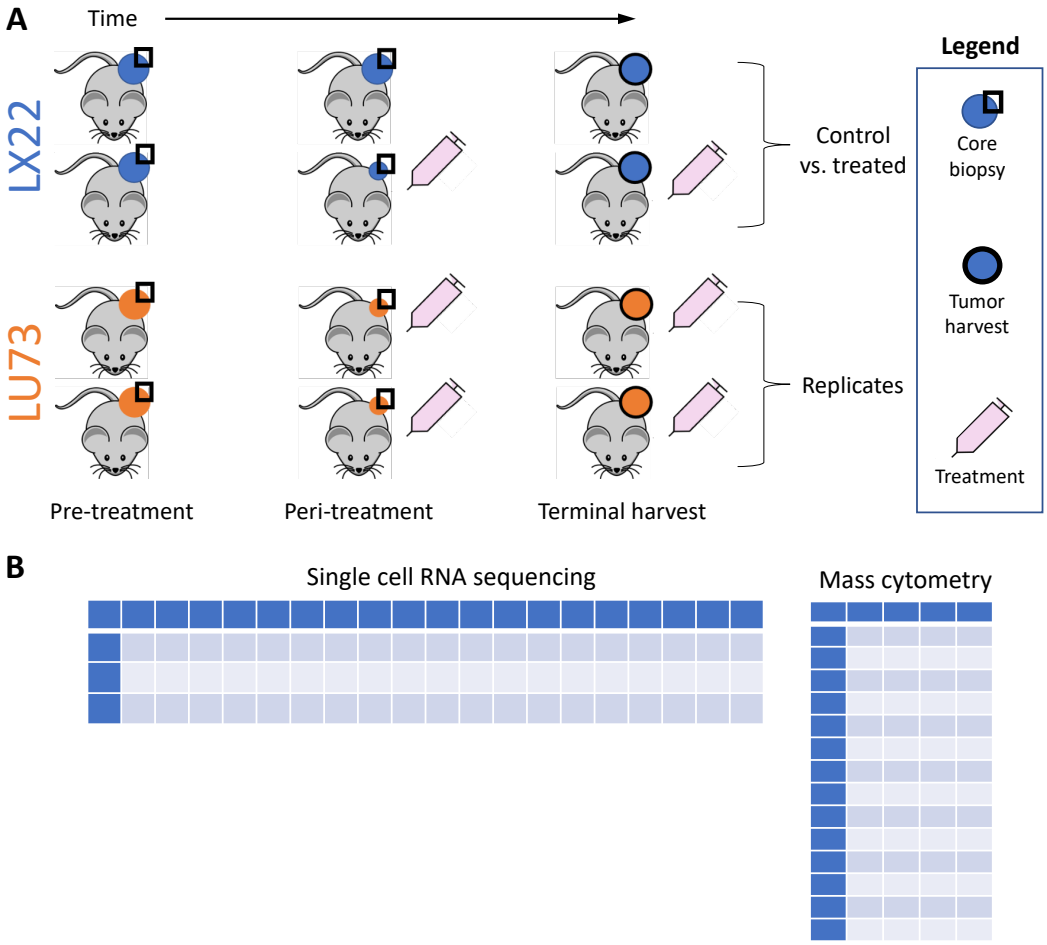


Figure 7.1: Experimental setup that generated the data used in this manuscript.

A Two PDXs were used, LX22 (SCLC-N variant) and LU73 (classical neuroendocrine SCLC tumor). Two tumors per PDX were provided for this manuscript, with three timepoints each. The first two timepoints data come from tumor core samples and the last timepoint data comes from tumor harvest. The pair of LX22 tumors represents untreated (top LX22 timecourse) vs treated (bottom LX22 timecourse); the pair of LU73 tumors were both treated and thus represent replicates of this tumor's behavior and composition over treatment.

B Single-cell RNA sequencing (scRNAseq) breadth and depth vs mass cytometry breadth and depth. scRNAseq measures a broader range of markers (transcripts, represented by the top blue row in the "Single cell RNA sequencing" array) but fewer events (cells, represented by the leftmost blue column in the array). Mass cytometry (CyTOF) measures fewer, pre-selected, markers (modified or unmodified proteins, top blue row in the "Mass cytometry" array) but more powerfully measures 10-1000 times more events (cells, leftmost blue column in the array).

both LX22 samples and all conditions (6 total) were dimension-reduced and clustered together, and the same for LU73 samples / conditions. Because the UMAP and clustering is based on the measurement of each protein in our panel, we consider each location in the UMAP to represent gradations across phenotypes that a cell may take on, and each cluster to represent a group of similar-phenotype cells. Assessing the proportions of the cells in each sample/condition occupying each cluster enabled us to evaluate the cluster sizes over time and across samples. This gives an idea of how cellular identities and phenotypes may change over time in the

tumor.

For PDX LU73, combining each sample (121 and 368) pre-, peri-treatment, and terminal harvest samples followed by Leiden partitioning resulted in 14 clusters (**Figure 7.2B**, top). Since the LU73 samples represent two replicates of a treated PDX, we can evaluate the similarities and differences between changes in the LU73 tumor over time. As an example, in cluster 0, cells seem to have left the cluster over time in one replicate (LU73 121) while in the other replicate, cells remain in cluster 0 across treatment. By contrast, in cluster 5, cells leave over time in both replicates. In general, in dimension-reduced space, cells over time in PDX 121 move "down" the y-axis (UMAP1) and "rightward" on the x-axis (UMAP0), while cells over time in PDX 368 move mostly "leftward" on the x-axis (UMAP0) and slightly down (UMAP1) (**Figure 7.2B**, bottom).

For PDX LX22, combining each sample (104 and 113) pre-, peri-treatment, and terminal harvest samples followed by Leiden partitioning resulted in 17 clusters (**Figure 7.2C**, top). The two samples start out relatively similar, with cells in clusters 6 and 10 (**Figure 7.2C**, top). However, comparing the timecourse with and without treatment, cells from clusters 6 and 10 in the LX22 untreated condition (tumor 104) move to clusters 0 and 1, while those cells in the LX22 treated condition (tumor 113) move to clusters 3 and 5 and then finish in clusters 4 and 7 (**Figure 7.2C**, top). In dimension-reduced space, we can see that untreated cells over time generally stay in the same location (**Figure 7.2C**, bottom, top row) while treated cells over time move "down" the y-axis (UMAP1) and then move "leftward" on the x-axis (UMAP0) (**Figure 7.2C**, bottom, bottom row).

As we expected, there were more differences between treated vs untreated PDX samples than between two replicates of treated PDX samples, when it comes to cell "movement" in phenotypic space over time (**Figure 7.2B**, top). Interestingly, however, in only the LX22 tumor type, cells "moved" in phenotypic space from one location to a completely different location with treatment. This is most easily appreciated in **Figure 7.2C**: LX22 104 vehicle-treated and LX22 113 pre-treatment exist in approximately the same location in dimension-reduced space, while LX22 113 peri-treatment has partly moved into a region of the UMAP where cells of the other conditions do not exist, and LX22 113 terminal harvest is in a completely different location. In the LU73 tumor type, cells "move" over the treatment course to change regions of highest density within the dimension-reduced space, but never to a completely new region of the UMAP (except a small one that represents cluster 11).

7.4 LX22 tumor clustering over time reveals treatment-induced subtype switching of this tumor

We were then interested in identifying phenotypes that correspond to the Leiden clusters for each tumor. This could enable us to draw conclusions about phenotypic changes over time, or across treatment conditions. We used two means for identifying cluster phenotypes. Briefly, the first process involved ranking each cluster by marker expression, per marker, and assigning highest- and lowest-ranked clusters per marker, then using

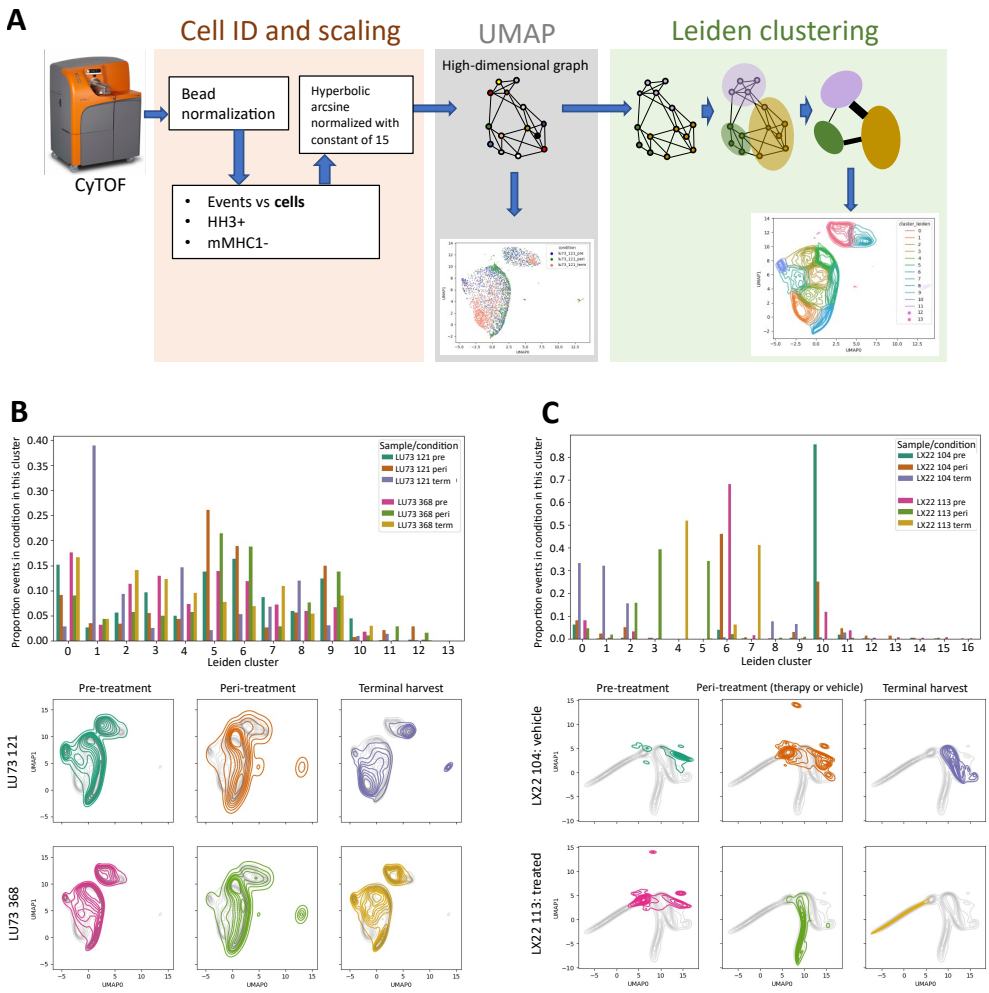


Figure 7.2: Using UMAP and Leiden clustering on mass cytometry data enables assessment of density changes over time and mapping which clusters increase or decrease with treatment

A When data comes from the mass cytometry machine, preprocessing includes bead normalization to account for changes in detection over multiple runs. Preprocessing continues with delineation of cells (to be kept) vs events (which may be cell fragments or other noise). Cells are then filtered to include only those positive for human histone 3 (HH3) and negative for mouse MHC class I (mMHC1). Preprocessed data is then transformed via hyperbolic arcsine with a normalization constant of 15. This preprocessed and normalized data is then dimension-reduced via UMAP. UMAP enables scatter or density plotting of the data in 2 dimensions. Leiden clustering is then performed on the high-dimensional graph generated by the UMAP algorithm, to ensure clustering in a more accurate dimensional space. Here, the LU73 data (sample 121) is shown as an example of a UMAP scatterplot (middle) and density plotting of the Leiden clusters (right).

(7.2B, 7.2C) Comparing Leiden clusters over conditions and times. Above, bar plots for each tumor, with all conditions and times showing the proportion of the cells in the sample/condition occupying a particular cluster over time. Below, density plots for each tumor based on condition and timepoint. Gray contour lines indicate the density plot for all tumor samples and conditions for that tumor, with the colored density plot for that particular condition or time overlaid. Colors used in the density plot to denote condition/time correspond to the bars in each barplot denoting the same condition/time.

B LU73

C LX22

Note the subtype switching over time in the LX22 treated sample (sample 113) where, for example, cluster 4 is composed entirely of cells from the terminal harvest (and no cells from any other condition).

each cluster's list of markers it was highest and lowest in as genes to evaluate via Gene Ontology (GO) and other databases (see **Figure D.2** for a more in-depth step-by-step description). We combined general GO terms from all clusters' results and noted which clusters were highest-ranked for each general GO term, and which were lowest-ranked (**Figure 7.3A**, **Tables D.1** and **D.2**). We thus had an approximation of phenotypic identities or activities that each cluster was highest- or lowest-ranked in. Additionally, we used Marker Enrichment Modeling (MEM), which calculates enrichment scores per protein marker per cluster [187]. This enabled comparison between Leiden clusters for expression of each marker, as well as grouping markers by similar expression profiles across the dataset and grouping Leiden clusters by similarity of marker expression (**Figure 7.3C**).

With general phenotypic identities across clusters, we investigated LX22, as it had the most striking changes in cell population within clusters over time and with treatment. In **Figure 7.3**, we highlight clusters 4, 6, and 11, for their varying behaviors over time and treatment. Cluster 4 (**Figure 7.3A**, first column; **Figure 7.3B**, top) appears to grow out during treatment and represent a majority of the cells and cell density in the terminal-harvest, chemoresistant condition of treated LX22. Per both cluster-identification methods, cluster 4 was one of the higher clusters in developmental markers, regulation of apoptosis, and signaling (see for example the high expression of YAP1, Oct3-4, and SOX1, and high expression of phosphorylated AKT, STAT3, and ERK by MEM in **Figure 7.2C**).

Cluster 6 decreases over time in both treated and untreated conditions, though in the untreated condition it first increases at the peri-treatment measurement before decreasing (**Figure 7.2C**, top). Cluster 6 is one of the clusters that has the lowest expression of many markers, as per **Figures 7.3A** and **7.3C**; according to GO terms it is one of the lowest-expressing clusters for apoptosis regulation, neural development, and migration (among others). We cannot say whether cluster 6 cells don't express many proteins in general, or whether they express specific proteins that were not part of our mass cytometry panel (**Table 7.1**). Since mass cytometry measurements must be made by evaluating specific markers, it is always possible that informative markers may be missed if not included.

Cluster 11 remains at a somewhat constant cell density, despite tumor evolution over time and differing treatment conditions. This cluster was a higher-expressing cluster for markers related to cell-surface interactions, and one of the lower-expressing clusters for stem cell pluripotency (**Figure 7.3A**), and indeed has lower stemness marker expression via MEM (**Figure 7.3C**).

Based especially on the decrease in cell density in cluster 6 and increase in cluster 4, it seems that treatment induces subtype switching in LX22 toward a more stem-like phenotype. This is consistent with the type of phenotypic transitions that would help a tumor withstand and develop resistance to treatment. An increase in stemness could help the tumor withstand treatment in that stem and stem-like cells are known to divide

A

LX22 clusters	4	6	11
Apoptosis regulation	Hi	Lo	
Autophagy			
Axon guidance		Lo	
Cell adhesion			
Cell surface interactions		Lo	Hi
Cytokine signaling			
Cytoskeleton regulation	Lo	Lo	
Development	Hi (neural)	Lo[3]	
ECM organization		Lo	
Immune signaling	Hi[1]	Lo[4]	
Immune system	Lo		
MicroRNAs in cancer			
Migration		Lo	
Neutrophils			
Proliferation			
Proteoglycans in cancer	Hi		
Signaling	Hi[2]	Lo[5]	
Stem cell pluripotency			Lo
Stem cell proliferation			
Transcription	Hi		

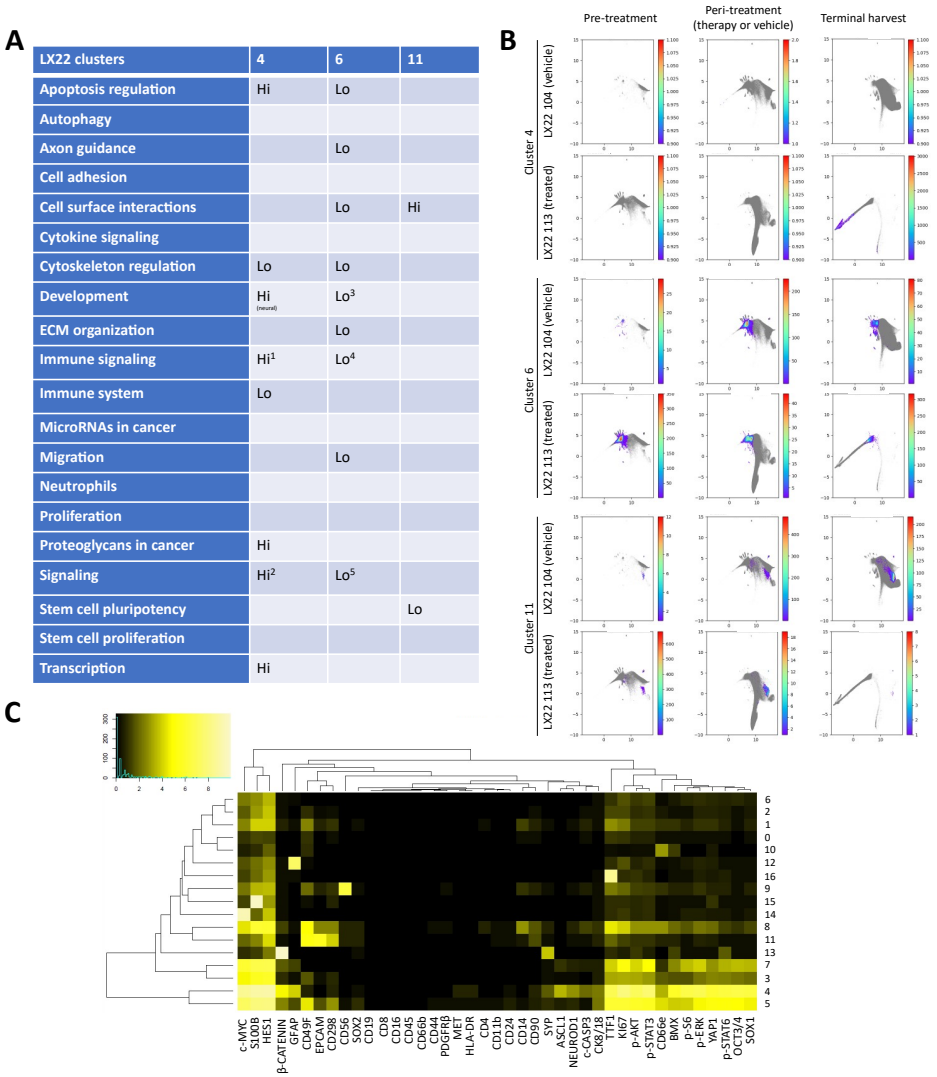


Figure 7.3: LX22 tumor clustering over time reveals treatment-induced subtype switching of this tumor

A Select LX22 clusters via Leiden clustering denoted in this table. All clusters are annotated with high ("hi") or low ("lo") expression of each protein marker, and high- and low-groups of markers were given gene ontology (GO) categories with significant enrichment of multiple markers in the group. The GO categories are summarized in this table, and where more specific information about a category is available for a particular cluster, (for example, "MAPK signaling" rather than simply "signaling") it is noted in the table or legend. The full table of which this is a subset is **Table D.2**. In (7.3A), select LX22 clusters and the GO categories for which they have high or low expression are noted.

B LX22 clusters over time. The same clusters (4, 6, and 11) shown in the table in (7.3A) are noted here in (7.3B) with density of cells per cluster in the condition at the particular timepoint noted (density scale to the right). Scatterplots are based on UMAP dimension reduction, where gray dots indicate all cells in the sample at the indicated time (indicates background sampling amount). Cells in the cluster are purple through green and yellow to red based on density. Note that cluster 4, which includes stemness/development markers and apoptosis regulation markers per (7.3A) and (7.3C), is found only at terminal harvest in treated samples.

C Marker enrichment modeling (MEM) plot for all LX22 conditions and clusters, with black to yellow hue based on expression level per marker per cluster. Markers have been clustered based on expression level correlation between clusters (top dendrogram), and Leiden clusters of the combined dataset have been clustered based on similarity across markers (left dendrogram).

1: IL-2, 4; TGFβ; 2: Jak-STAT, Rac1/MMP2 pathway, prolactin signaling, TSLP, VEGF; 5: hematopoietic, cardiac progenitor; 6: "interleukins", Fc epsilon receptor signaling; 7: PKB, MAPK cascade, Rap1, DAP12, EGFR, FGFR, NGF

more slowly [226], which is especially useful against chemotherapy such as cisplatin and etoposide that kill quickly-dividing cells [227]. The stemness increase could also help the tumor explore other phenotypes that would be less susceptible to treatment [228]: more stem-like cells means more accessible phenotypes for the tumor or a subpopulation within the tumor.

Evaluating LU73 via GO terms and MEM, in the replicate LU73 121, clusters 0 and 10 decrease in density, and clusters 8 and 1 increase in density (**Figure 7.2B**; **Figures D.1A** and **D.1B**). Cluster 0 is one of the higher-expressing clusters for markers related to apoptosis/autophagy regulation, and development (including, more specifically, markers related to axon guidance), and cluster 10 is one of the higher-expressing clusters for developmental markers as well, along with migration and signaling markers. Cluster 1 was one of the lower-expressing clusters in general, and specifically low-expressing with regard to signaling (including immune-specific signaling) and the immune system in general. Cluster 8 is high-expressing for markers related to axon guidance like cluster 0, but also high-expressing for markers related to cell adhesion and immune markers (**Table D.1**). Considering these changes in cell density, tumor LU73 121 seems to decrease in developmental and motility markers and increase in immune markers, while remaining neuronal in nature over time. By contrast, in LU73 368, cluster 0 remains at an approximately constant density over treatment and cluster 10 increases in density over treatment (**Figure 7.2B**; **Figure D.1B**). This indicates that LU73 tumor 368 may increase in developmental and motility markers over treatment, in the opposite fashion from LU73 tumor 121. This indicates overall that both LU73 tumors retain the neuronal characteristics of "classical" neuroendocrine SCLC over the treatment course, but one increases in stemness and motility similar to treated LX22 (tumor 113) while the other increases in markers related to motility and immune signaling.

7.5 Tracking stem-like subpopulations through chemotherapy treatment

Interested in the stemness features we noted in the treated LX22 tumor, we delved deeper into evaluating this population within the LX22 PDX. Data includes a separate set of tumor samples, one vehicle-treated sample and a sample that underwent four cycles of chemotherapy treatment. To begin this evaluation, we wanted to use MEM in a more classical way, evaluating selections, or gates, of cell subpopulations based on mass cytometry markers [186]. After bead normalization, cell-vs-fragment selection, and filtering for human cells (similar to **Figure 7.2A**, left) we envisioned the LX22 PDX in a dimension-reduced space, using visualization of t-Stochastic Neighbor Embedding (visNE) [229]. The visNE plot represents the 26-dimensional space built by the measurements for an alternate antibody panel we used (**Table 7.2**), reduced into two dimensions for visualization.

We determined gates from this visNE plot, as different areas in the two-dimensional space represent

Marker	What it may indicate	Marker	What it may indicate
CD45	Immune cells	CD56	Neural cell adhesion
Rhodium	Positivity means dead cells	PDGFRα	Endothelial cells, fibroblasts
CD151	cell adhesion; integrin trafficking/function	MET (Hepat. GFR)	Organogenesis, wound healing – seen in metastasis, angiogenesis in cancer
CD274 / PD-L1	Tumor suppression of immune system	CD44	Cancer stem cell marker in some cancers; homing, metastasis
CD68	Macrophages	p-S6	Ribosomal protein; indicates growth
Total EGFR	Tumor marker	p-AKT	Survival / growth signaling
p-EGFR	Activated EGFR	HLA-DR	MHC class II surface receptor
Vimentin	Mesenchymal cells; EMT marker	c-MYC	Cell cycle progression
EPCAM	Epithelia, epithelial neoplasms	Oct3-4	Embryonic stem cells
SYP	NE cell / NE neoplasm marker	SOX2	Stemness, pluripotency
c-CASP3	Apoptotic signaling	p-STAT3	Self-renewal of embryonic stem cells
CD24	B cells, neuroblast	KI67	Active cell cycle (proliferation)
h-H3	Denotes a human cell (from PDX tumor)	Mouse MHC class I	Denotes a mouse cell (the host for the PDX tumor)

Table 7.2: Mass cytometry panel for **Figure 7.4**

p-, phosphorylated; c-, cleaved; Hepat., hepatocyte; GFR, growth factor receptor; BTK, Bruton Tyrosine Kinase; h-H3, human histone 3

changes in combinations of markers measured by the mass cytometer. We gated the data using all timepoints of both LX22 treated and untreated tumors together, resulting in 18 gates (**Figures 7.4A** and **7.4B**). Comparing the untreated sample to the treated sample (**Figure 7.4A**), we see that just as in the UMAP and Leiden clustering approach, the treated sample is most densely present in a different location in the dimension-reduced space on the visNE plot, indicating subtype switching.

Using these gates, we performed MEM, and were thus able to identify more- and less-expressed markers in our panel per gate (**Figure 7.4B**). We were particular interested to see that Gate 8 had relatively high expression of SOX2 and Oct3/4, and Gate 11 also had relatively high expression of SOX2, indicating that the cells captured by one or both of these gates may represent a stemlike population. We then performed SPADE, to envision the LX22 tumor cells as subpopulations within the tumor and to see any hierarchical connections between them [230]. Comparing untreated LX22 to treated, we see again that density within the population (representing areas of more cells with a particular phenotype in the tumor) changes with treatment (**Figure 7.4C**). We then evaluated the expression of Oct3-4 within the SPADE tree, and found that the same groups of cells between untreated and treated had high expression of Oct3-4, indicating that this relatively smaller subpopulation represents a chemotherapy-stable subset of stem or stem-like cells within the tumor (**Figure 7.4C**, red).

While this SPADE expression analysis indicates approximately equal expression of Oct3-4 pre- and post-treatment, in our earlier analysis of LX22 we do not see as high expression of stemness markers pre-treatment as we do post-treatment. However, using **Figures 7.2B** and **7.2C** to evaluate clusters 7 and 8, which represent

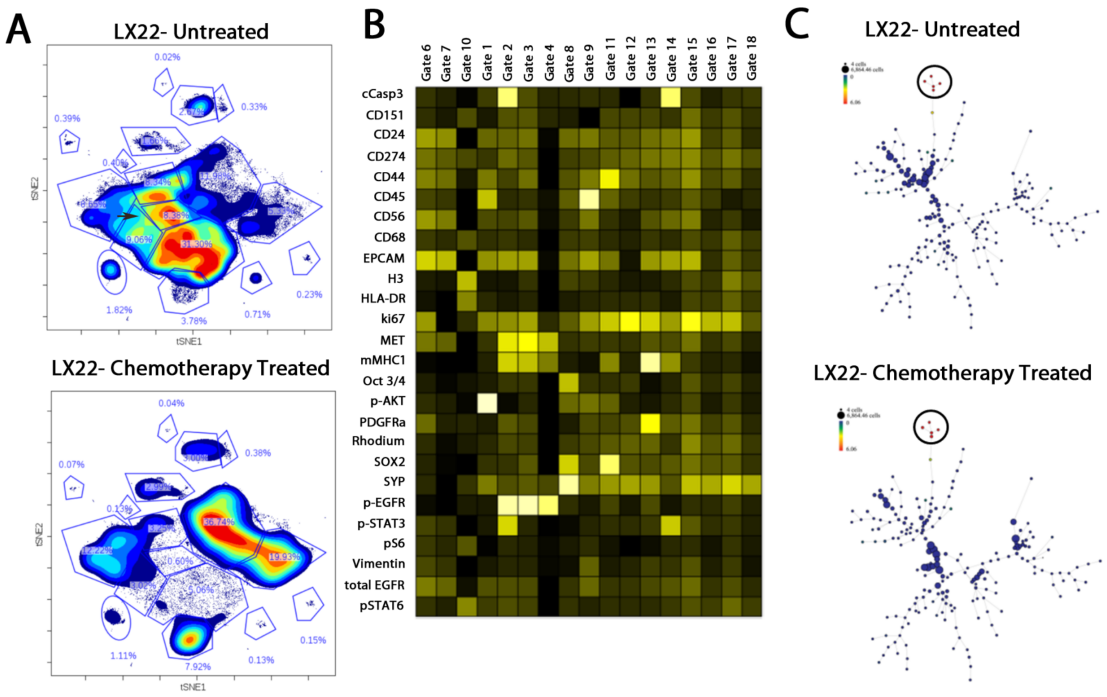


Figure 7.4: Tracking stem-like populations through chemotherapy treatment

A 35-marker panel Mass Cytometry viSNE plot with viSNE assisted gating. visNE performed on combined treated and untreated data, then plotted in two dimensions using color to indicate cell density (red, most dense; blue, least dense). Gates were drawn manually based on this visNE, with the varying densities enabling determination of most similar cells (in areas of higher density separated from other cells by areas of lower density).

B Marker enrichment modeling (MEM) plot for both LX22 conditions using visNE-assisted manual gating, with black to yellow hue based on expression level per marker per cluster. Note Oct3/4 positivity in Gate 8, and SOX2 positivity in Gates 8 and 11.

C SPADE analysis of LX22 treated and untreated tumor data, arranged based on similar marker expression and connected using a minimum spanning tree. Each circle represents a population of cells with similar phenotype based on the 35 markers in (7.4B), and the sizes of the circle are based on the number of cells in each circle's population split across untreated (top) vs treated (bottom), to envision cell phenotype density changes without and with treatment. The heat color for each circle represents the median expression of Oct3/4 in cells within that circle. Large black circle indicates the SPADE-assessed subpopulations of high Oct3/4 expression in both treated and untreated LX22 tumor cells.

cell density in LX22 treated pre- and post-treatment samples, and LX22 untreated terminal sample, respectively, we see via MEM that these do have expression of Oct3-4 (**Figure 7.3C**). Thus, both analyses indicate that a subpopulation within pre-treated, on-treatment, and post-treatment LX22 expresses Oct3-4. This may represent chemotherapy-stable cells and could be a subpopulation that enables the growth of the treatment-resistant, stem-marker-expressing cells seen in the treatment-resistant LX22 terminal harvest sample, made up for the most part by clusters 4 and 7 (**Figure 7.2C**).

Working to confirm this via immunohistochemistry, we see that the LX22 PDX tumor includes rare cells positive for Oct3-4 (see **Figure 7.5C**). The Oct3-4 positive cells do not appear to be clustered together in a

particular location in the tumor, but distributed throughout.

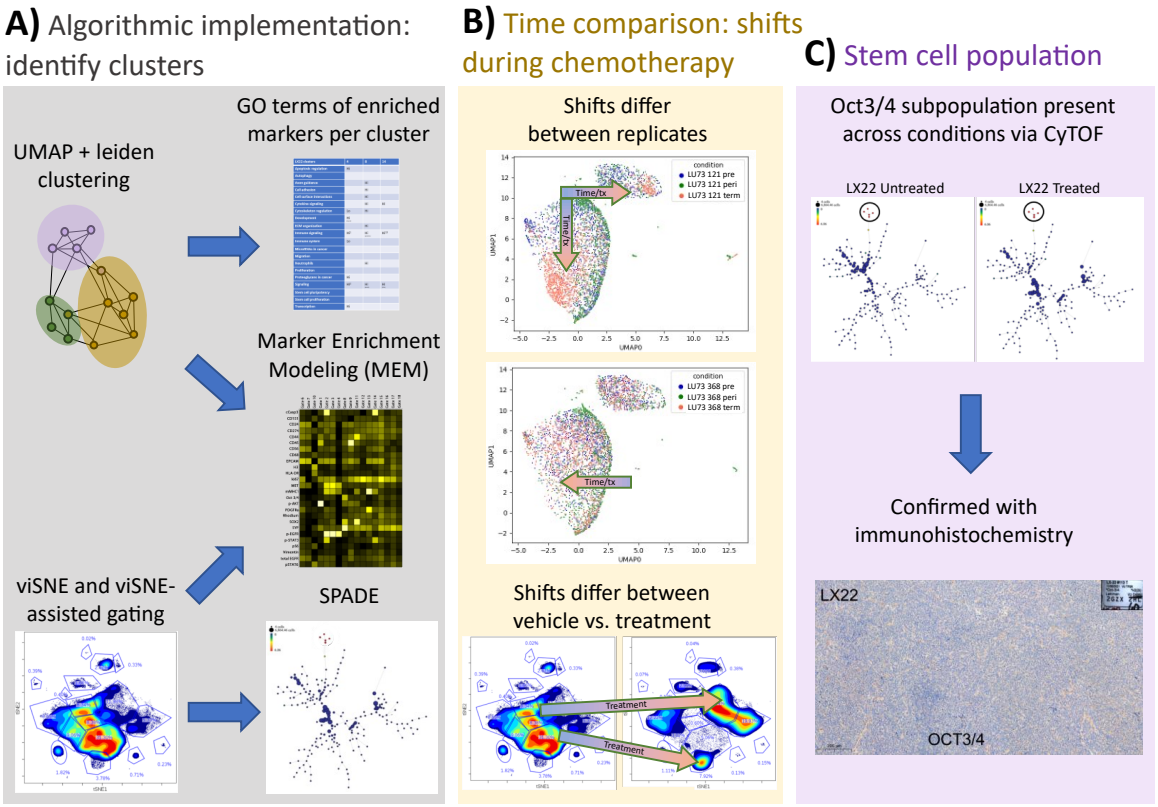


Figure 7.5: Novel algorithmic implementation, evaluation over chemotherapy treatment course, and ability to uncover rare subpopulations enable cluster identification, determination of subtype switching during treatment, and discovery of rare stem-like population in SCLC PDX tumors.

A Cluster identification is performed via multiple avenues, either UMAP dimension reduction to generate the high-dimensional graph on which Leiden clustering is performed, or viSNE dimension reduction followed by cluster assignment via viSNE-assisted manual gating. Determining enriched markers per cluster and assigning GO terms to these markers was performed on UMAP clusters; SPADE was performed on viSNE-assisted manual gated clusters. Marker enrichment modeling (MEM) was performed on both implementations of cluster assignment. All methods enable assignment of biological identity per cluster.

B CyTOF was performed on the same tumors (core sample or tumor harvest, see **Figure 7.1**) over treatment with vehicle or chemotherapy. Above, pre- and peri-treatment samples and terminal harvest sample for two replicates of LU73 denote varying responses to treatment as indicated by density shifts over time. Below, LX22 chemotherapy-treated sample compared to vehicle-treated sample over pre- and peri-treatment and terminal harvest indicates that treatment induces subtype-switching while vehicle treatment retains cell identity.

C Cluster identity assignment and identities that do not shift over treatment course reveal a stem-like subpopulation present in CyTOF data, the presence of which was subsequently confirmed via immunohistochemistry.

7.6 Discussion

Here, we have evaluated a novel time course of SCLC tumor growth and treatment, assessing multiple timepoints rather than before-vs-after treatment, and investigating treatment response using human-relevant chemotherapy doses rather than the maximal tolerated dose. We compare two replicates of a PDX as they

undergo treatment, and a set of treated-vs-untreated PDX tumors, and work toward identification of phenotypic subpopulations within the PDXs. We use complementary methods to investigate phenotypic identities and to determine that a stem or stem-like subpopulation within SCLC PDX LX22 persists during chemotherapy treatment and may represent a portion of the post-treatment tumor (**Figure 7.5**). Importantly, the use of CyTOF and its increased depth of measurement enables us to make these measurements, as the small subpopulation of stemlike cells is less likely to be detectable via scRNAseq due to its inflated zero counts from fewer cells sampled and potential transcript dropouts.

One limitation of our approach is that while we included ASCL1, NEUROD1, and YAP1 in our mass cytometry marker panel (**Table 7.1**), only YAP1 showed relatively high expression in a very small number of clusters (one to three) in each PDX. This is particular surprising given that LX22 has been denoted an SCLC-N subtype tumor, so we would expect much more NEUROD1 expression [222]; additionally, LU73 is a more classical SCLC tumor, so we would expect more ASCL1 expression. Recent transcriptional measurements of SCLC PDXs result in appreciable clusters of ASCL1-high, NEUROD1-high, or POU2F3-high, or YAP1-high cells [196, 204, 11], but our results here do not have such obvious expression. We hypothesize that this might be related to the use of scRNAseq transcript detection for these measurements, whereas here we are detecting the proteins themselves; it has been shown that there is not always a one-to-one ratio of mRNA transcript to translated protein [231]. No studies have been published using mass cytometry to evaluate SCLC [217].

Inherent in CyTOF studies is the requirement to select markers for analysis, unlike scRNAseq when sequencing enables the identification of all transcripts within the cell. The increased depth of CyTOF measurements allows for detection of rare cell populations, but requires that the investigator establish a panel of antibodies that will mark those populations. Several markers in our panel had little to no expression according to MEM (**Figure 7.3C**, **Figure D.1A**). If we replaced those markers in the panel with markers expressed in a subset of tumor cells, we may be better able to assign phenotypes to clusters within each tumor. Of course, which markers might be expressed in a subset of SCLC tumor cells beyond those with medium-to-high expression noted here (**Figure 7.3C**, **Figure D.1A**) remains to be elucidated. Future work will include CyTOF antibody panel optimization.

It is very important for such a high-morbidity and high-mortality tumor that we determine the mechanism for the recurrence and treatment resistance that occurs in almost all cases [12]. The movement between phenotypes over time shown here is in accordance with phenotypic plasticity denoted in other SCLC studies, [197, 196] and noting stem or stem-like chemostable subpopulations here is an important next step in understanding the mechanism behind SCLC treatment evasion. Boolean or mechanistic modeling, which can be performed by taking advantage of the 3-point timecourse in this data, may also help further our understanding of the mechanisms and protein networks in the untreated and treated SCLC tumor. This will lead to improved

hypotheses for treatment strategies that can then be evaluated experimentally, and eventually, clinically.

CHAPTER 8

Conclusions and future directions

8.1 Conclusions

Phenotypic plasticity is an important part of the dynamic processes underlying tumor evolution, treatment evasion and resistance ([232, 5], **Chapter 2**). The identities that a cell can access through phenotypic transitions enable a tumor to develop despite challenges posed by its microenvironment (e.g., oxygen supply, immune cells), and impact response to treatment via drug-tolerant, persistent phenotypes [63, 64, 62] or phenotypes that are unaffected by the treatment (**Chapter 2**). Evaluating potential transitions and determining phenotypes to which a tumor cell has access will be essential to develop effective treatment strategies: if one treatment results in a tumor accessing new phenotypes, "targeted landscaping" may allow us to intentionally push tumor cells toward a new phenotype but then impose a treatment known to be effective for that phenotype ([20] **Chapter 2**).

Mechanistic modeling of tumor cell populations is a critical tool for evaluating tumor heterogeneity and phenotypic transitions that enable tumor development, evolution, and resistance. For example, the model by Gupta and colleagues [13] shed new light on tumor cell behavior and altered the field's understanding of cancer stem cells. Experimental data can be used to build various types of mechanistic models, which can in turn produce predictions to be tested experimentally, enabling a systems biology approach toward understanding heterogeneity, phenotypic plasticity, and likelihood and rates of behavior within a tumor system (**Chapter 3**).

Unfortunately, challenges remain in modeling a biological system, where it may not be clear which species, actions, and interactions to include in a model. Including every potential detail may needlessly complicate the model and can lead to overfitting, while including too little in a model can lead to missing important details. Posing such questions, about multiple model options and which may best represent the system of interest, has often been evaluated with information theoretic model selection. However, I have shown here that using Bayesian inference can improve the model selection process (**Chapter 5**). Via an example building off prior work by Galipaud and colleagues, [1, 26] I show that using Bayesian inference, enabled by Multinest [29, 30, 31] that calculates the Bayesian evidence or marginal likelihood, results in a more sensible outcome for the likelihood that a variable should be included in a model. The consideration of various features of, and hypotheses represented by, multiple models that occurs after the ranking of models is much more interpretable when performed via Bayesian principles (**Chapter 5**).

Using these Bayesian principles, insight into phenotypic transitions and plasticity in the small cell lung

cancer tumor can be gained (**Chapter 6**). Performing Bayesian multimodel inference, I have predicted that SCLC tumors are highly plastic, with phenotypic transitions between any of the SCLC subtypes (**Figure 6.6**). I have also predicted that the presence of certain subtypes may affect the rates or likelihood of these phenotypic transitions (**Figure 6.5**). The ability to investigate time-course data more deeply will enable further conclusions to be drawn; in my work evaluating mass cytometry time-course data, (**Chapter 7**), phenotypic transitions are suggested in differing proportions of phenotypic populations within the same tumors over time (**Figure 7.2**). Since these tumors are sampled over time, however, it cannot be definitively claimed that phenotypic transitions within individual cells took place, as my work in **Chapter 6** predicts. I have worked toward testing a subset of SCLC phenotypic transitions, with the potential for tests of more transitions eventually (**Appendix A**, **Section 8.2**). Recently, Gopal and colleagues tested phenotypic transitions by tracking individual cells and the SCLC transcription factors they express and demonstrated all/some of the transitions I predicted using Bayesian multimodel inference in **Chapter 6** [233].

In summary, Bayesian model selection and multimodel inference has a principled and interpretable outcome, augmenting the mechanistic modeling of a tumor system such that biological hypotheses can be tested and predicted as part of the model-building process. Further, using the model features that result from the multimodel inference process, I have been able to make predictions, some of which have been validated in work at other institutions [233] and others that could still be tested. However, more steps need to be taken both to understand cancer heterogeneity and phenotypic transitions as well as to improve mechanistic modeling of tumor systems at multiple scales.

8.2 Future directions

8.2.1 Evaluating and generating further predictions from the SCLC population dynamics model in Chapter 6

There are two main predictions from the SCLC mechanistic models based on Bayesian multimodel inference (**Chapter 6**). The first is the prediction that phenotypic transitions are likely between any of the SCLC subtypes included in the model (**Figure 6.6**). This indicates that instead of a hierarchy of stem/stem-like cells differentiating into various phenotypes that can populate the SCLC tumor, that all SCLC subtypes can undergo the phenotypic transitions that can result in a multi-subtype tumor. The second is the prediction that transitions between SCLC-A and SCLC-Y subtypes maybe affected by the presence of SCLC-N and/or SCLC-A2 (**Figure 6.5**). This indicates that the subtype composition of the tumor may affect the behavior of individual subtypes within. Both of these predictions are important for understanding tumor evolution as well as treatment response depending on SCLC tumor composition.

The ideal way to test these predictions is with tumor cell tracking using fluorescent transcriptional re-

porters, to track in live cells which SCLC subtype-identifying transcription factors are active in each cell and thus in the whole population. As discussed in **Appendix A**, I have worked to fluorescently label ASCL1 in NCI-H69 cells ([234]) such that the cells will fluoresce when expressing ASCL1, and can thus be considered SCLC-A subtype cells. Losing fluorescence will indicate SCLC cells transitioning from subtype SCLC-A into another subtype; such transitions have been demonstrated previously [10, 179, 208, 11]. A cell gaining fluorescence indicates SCLC cells transitioning from a non-SCLC-A subtype into an SCLC-A subtype, which my model predicts and has only been shown in the last several months. In their publication, Gopal and colleagues [233] show via live-reporter imaging that cells of the SCLC-N phenotype (expressing NEUROD1) and cells of the SCLC-Y phenotype (expressing YAP1) can transition to SCLC-A, as my model predicts. In the ideal case, my fluorescently labeling of SCLC cells would involve two or more labels, though Gopal and colleagues note that cell viability and division rate limited them to two labels [233]. The ability to evaluate any one SCLC cell for which of any of the 5 subtypes it is (SCLC-A, -N, -A2, -P, -Y [9]) would represent a leap forward in investigating subtype behavior *in vitro*. Evaluating fluorescence changes (transitions) in the presence vs absence of SCLC-N or SCLC-A2 cells would then address my second prediction, that these subtypes may affect the rates or likelihood of these transitions.

Identifying the subtype of each SCLC cell in a population would also benefit the mechanistic SCLC model in that tracking individual cells and their behaviors, such as division, death, and transitions, would further inform parameter rates in the SCLC population dynamics mechanistic model. This represents the potential synergy between modeling and experiments, where I have used mechanistic modeling to make the two aforementioned predictions and narrow the likely parameter rates for each biological action to a certain extent, but further testing these behaviors experimentally will provide even narrower ranges of rates for these cell actions that can then be incorporated into the model. In addition, based on the outcome of experiments based on predictions, whether validating or refuting, the model can be further edited: either fully enshrining validated behaviors, or removing behaviors that were tested but not seen.

Finally, an important future endeavor will be to incorporate the potential for blocking a phenotypic transition in the model. Blocking one or more phenotypic transitions in the population dynamics model is an *in silico* experiment to investigate what will occur in the SCLC tumor if a transition-abrogating treatment were used. The predictions brought about by running such *in silico* experiments could then be further tested using fluorescent live-reporter imaging as discussed above. If we are to perform "targeted landscaping" as discussed in **Chapter 2** and above in **Section 8.1**, preventing a transition could be an important feature of this approach.

8.2.2 Extending the SCLC population dynamics model in Chapter 6

With the results from my multimodel inference analysis, the majority of possible models are excluded (**Figure 6.4**) and either a model per dataset or overall model, based on more-likely biological hypotheses that resulted from the analysis, can be used for future work.

The ideal utilization for modeling in systems biology is to model as much of the relevant system as possible ([86], **Chapter 3**), and currently the SCLC population dynamics model includes only tumor cells and no location information. Utilizing 'compartments' [116] in a population dynamics model would enable two or more "versions" of the population to be modeled, where model terms are the same between compartments, but parameter rates for each model action may be different (**Figure 8.1**). Incorporating compartments into population dynamics modeling of the SCLC tumor to incorporate the tumor *in situ* and the tumor having locally invaded (without having permeated the basement membrane, *vs.* within lung stroma, respectively, as in [11]) is underway [1].

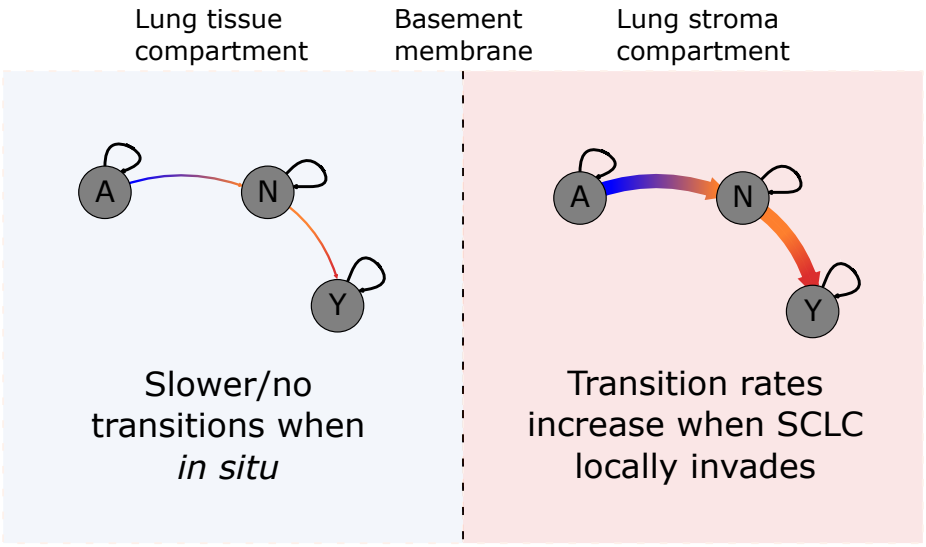


Figure 8.1: Population dynamics modeling with compartments
Compartments represent different environments where the population may exist. Parameter rates may differ from one compartment to another. Wider arrow indicates increased rate.

Modeling the SCLC microenvironment, such as intra-tumor immune cells as well as immune cells in the tumor periphery, along with stromal cells (such as cancer-associated macrophages and fibroblasts), including cells recruited to support or perform angiogenesis, would improve the relevance of the model to the system in which the tumor exists. Vasculogenic mimicry data in SCLC [203] could be incorporated into the SCLC population dynamics model and used for vascular- or angiogenesis-related parameter fitting; data from SCLC-related immune cells is being generated and could similarly be incorporated [2].

[1] Personal communication, Dr. Leonard Harris, University of Arkansas
[2] Personal communication, Dr. Julien Sage, Stanford University

The inclusion of treatment in the model would represent a large step forward for the SCLC population dynamics model's translational relevance. Incorporating the different subtypes' treatment reactions, whether to a chemotherapeutic drug, a targeted drug, or a drug that could block a transition, would enable *in silico* treatment experiments either singly or in combination. This could determine the optimal concurrent or sequential drug treatments for a tumor, including within the context of targeted landscaping as discussed earlier. Inclusion of subtype treatment responses has proven difficult, as experimental investigations have assessed treatment response in SCLC cell lines, [235, 236, 237] not accounting for which subtype(s) are present in those lines. Other investigations have assigned SCLC subtypes to treatment responses in cell lines [196] based on the predominant SCLC subtype in that cell line, but assessments of treatment responses in single SCLC subtype subpopulations within a tumor or *in vitro* population have not been performed. Moreover, even if a population is completely composed of one SCLC subtype, that subtype may not respond identically if in an overall population with multiple subtypes - this is suggested by experiments showing that neuroendocrine and non-neuroendocrine cells together have better viability during treatment than either neuroendocrine or non-neuroendocrine alone [10]. Perhaps with fluorescent reporters that can identify cells within a population as discussed in **Appendix A** and **Section 8.2.1** above, cell tracking could reveal individual subtype response to drugs either singly or more beneficially within an overall mixed population, which could then be incorporated into the population dynamics model.

Finally, the SCLC population dynamics model, and the additions proposed above, solely represent populations of cells and thus represent cell-level dynamics. Incorporating multiscale aspects into the model would also improve understanding SCLC not only as a population of cells but also from the perspective of transcriptional signaling networks. Modeling transcription factor networks within SCLC has been previously performed [180] and incorporating the cell-level outcome of transcription factor network activity into the population dynamics model will provide additional insights into tumor behavior (**Figure 8.2**). In this way, each cell in the population dynamics model would incorporate a transcriptional model within, the outcome of simulations of that transcriptional model would affect individual cell characteristics that would then be incorporated into the population activity within the model. Optimal data for the transcriptional aspects of a multiscale model would be time-course data of molecules within cells, such as single-cell RNA sequencing or mass cytometry.

8.2.3 Further investigations into mass cytometry assessment of SCLC time-course patient-derived xenografts in Chapter 7

The investigations in this dissertation into SCLC heterogeneity and phenotypic subpopulations in patient-derived xenograft (PDX) models over time using mass cytometry data represent a promising start for time-

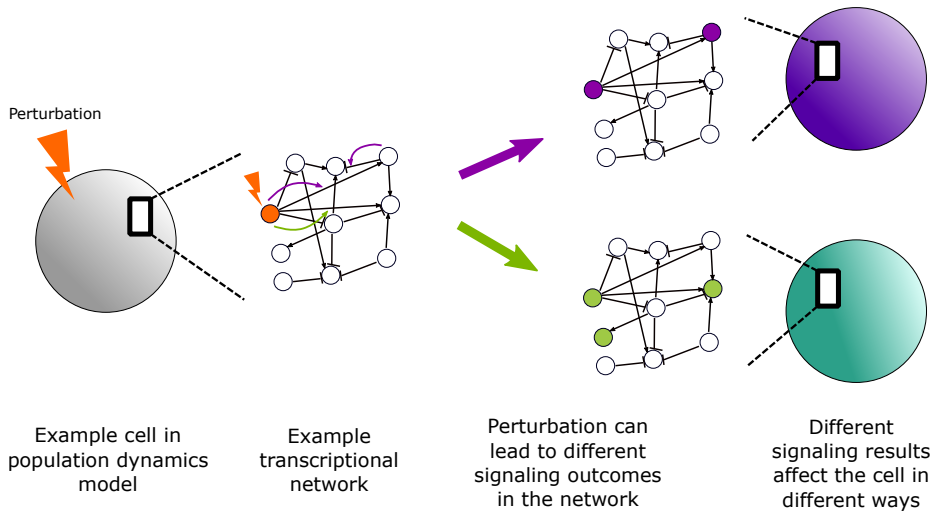


Figure 8.2: Multiscale modeling

course evaluations of SCLC tumors during evolution and treatment. Not only are the phenotypic changes assessed on the same tumors over time using serial core samples, but the PDX models used are translationally relevant because the mice are treated with subtotal chemotherapy - more relevant to the chemotherapy dosage and schedule of a human SCLC patient. However, interpretation of the data is limited because the computationally-defined phenotypic clusters do not differ ideally in marker expression. For example, as seen in **Figure 7.3**, the clusters express largely the same markers, but in differing amounts. Considered another way, some markers evaluated were found in zero clusters, while some were found in all clusters. Both of these facts lead to difficulty in optimally determining how each clusters differs from all or most others.

I hypothesize that this limitation stems from the mass cytometry panel used in these experiments. Because mass cytometry studies require the investigator to pre-select markers for analysis, antibody panels for these experiments require optimization to provide the most informative results. Mass cytometry has largely been used for immune or other hematopoietic-lineage cell investigations, [238] so more work needs to be done to develop a panel that will best represent the different cell types found in a small cell lung cancer tumor. It is also possible that a "general" SCLC panel would be difficult to develop and would need to be created based on an individual or small group of tumors' features. If the markers with low or no expression in the panels noted here were replaced with markers expressed in a subset of tumor cells - based on the proposed further optimization - phenotypes for each cluster in the tumor (**Figure 7.2**) may be more easily assignable.

More definitive phenotypes per cluster would also enable population dynamics modeling of the PDX tumor time-course data, with and without treatment. The mass cytometry data presented in **Chapter 7** does not align with SCLC canonical subtypes SCLC-A, N, A2, and Y, which are the potential subtypes that can be

used in the population dynamics modeling in **Chapter 6**. The same Bayesian multimodel inference approach taken using the CIBERSORT SCLC subtype proportion assignment in **Chapter 6** could be used, but instead with phenotypic subtypes or subpopulations found by the CyTOF data. The work in **Chapters 5** and **6** indicates that this methodology can be used for any type of population dynamics model, so it should also be applicable to the time-course mass cytometry PDX data. This mass cytometry data, indicating the increase or decrease in subpopulations over time (**Figure 7.2**), could then be used to fit a population dynamics model that includes *those* subpopulations, and to investigate either similar or additional biological hypotheses than those investigated in **Chapter 6**.

Beyond phenotypic subpopulations based on computationally defined clusters, mass cytometry also enables intracellular network evaluation because it measures proteins within or on the surface of cells. With the right antibody panel, signaling networks and changes in protein abundance in these networks could be assessed as the SCLC PDX tumor develops or responds to treatment. Modeling the measured proteins as a network or part of a network could also enable the multiscale modeling noted in **Section 8.2.2**, incorporating network dynamics and their cell-level outcomes into population-level changes over time.

8.2.4 Clinical implications of this work

The prediction that nonhierarchical phenotypic transitions likely play a role in tumor growth (and partial validation by [233]) indicates that this feature of SCLC is likely to heavily contribute to the clinical behavior seen in this tumor, with apparent high response rate followed by relapse and resistance. Treatments are in development that attempt to decrease the ability for cells to undergo epigenetic changes, some with success in some tumor contexts; due to the predictions herein, the use of treatments in SCLC may be useful.

The multimodel inference performed on the tumor proportion data in **Chapter 6** is also a promising tool for personalized cancer therapy. The ability to fit a mathematical model to a patient's own data would enable a clinician to generate predictions about how the patient's tumor will grow, and possibly how it will respond to treatment depending on the predicted subtype makeup over time. In this dissertation I add to this particular paradigm of personalized cancer modeling by adding the ability to use multimodel inference to determine the probability that a patient's tumor has the ability to perform one or more of the processes that likely contribute to SCLC recurrence and treatment resistance. Even without a specific prediction of a patient's tumor's cell population dynamics, determining whether certain subtypes are present, or certain cell-cell communications or phenotypic transitions are present could benefit a clinician aiming to prescribe personalized therapy to a patient.

Appendix A

Using fluorescent reporters to investigate phenotypic transitions in single cells

A.1 Validating predictions from the small cell lung cancer population dynamics model

The results presented in **Chapter 6** provide predictions about phenotypic plasticity in SCLC. Phenotypic transitions are known to exist in small cell lung cancer (SCLC), and have been shown moving from the neuroendocrine subtypes to non-neuroendocrine subtypes [10, 11]. Given that the investigators that demonstrated these transitions did not find transitions in the opposite direction, (non-neuroendocrine to neuroendocrine) the only known SCLC transitions could be considered hierarchical in nature, always proceeding from neuroendocrine to non-neuroendocrine. My multimodel inference investigation of SCLC candidate mathematical models shows higher likelihood of phenotypic transitions between nearly all subtypes, regardless of differentiation hierarchy, (**Figure 6.6**) indicating phenotypic plasticity. Transitions had not been seen from non-neuroendocrine to neuroendocrine SCLC subtypes, so I aimed to develop an experimental system to test this prediction of the model; within the past few months, Gopal and colleagues demonstrated several transitions in their publication [233]. The knowledge of nonhierarchical phenotypic transitions is important both in assessing the applicability of the SCLC population dynamics model presented in **Chapter 6**, and for moving forward in understanding SCLC.

A.2 Means for experimentally demonstrating phenotypic transitions

Phenotypic transitions in other tumor and physiologic systems have been demonstrated experimentally in different ways, some requiring more inference than others. Herein I consider different methods for evaluating phenotypic transitions *in vitro*, as it is *in vitro* experiments I would perform for validating model predictions. The biological interpretation of most of the publications whose methodologies are referenced herein are provided in **Chapter 2**.

Firstly, phenotypic transitions have been demonstrated *in vitro* using methodologies that fix or destroy the cells of interest. In these situations, investigators either sample from the same population over time, or grow multiple populations of cells simultaneously and use one replicate at a particular time while the others continue to grow to be measured later in the time-course. Bulk RNA sequencing has been used, where the absence of expression of particular genes early in a time-course and then presence of these genes later indicate the appearance of a new phenotypic population in the system [55, 19]. Immunofluorescence on fixed cells has also been used, to evaluate intracellular or membrane protein expression of individual cells, where cells in the population at one timepoint express a phenotypically-representative protein, but a different

representative protein at a later timepoint [16, 60]. The advent of single-cell RNA sequencing advanced these studies, increasing throughput from evaluation of fluorescence in single cells to transcriptomes of thousands of individual cells: comparing results from different timepoints can show that representative transcripts or groups of transcripts (such as to generate a phenotype "score") are present at unique times, indicating different phenotypic presence at different times [66, 11]. A similar analysis was performed in **Chapter 7**, using mass cytometry to perform a similar, higher-throughput analysis of single-cell protein expression comparing sampled tumors over time.

There are several potential pitfalls in evaluating phenotypic transitions using methods that destroy the cells. The assumption that the cell population is well-mixed in various phenotypes might not hold, so sampling from a population may not truly be representative of the entire population. This may be less of a problem *in vitro* than *in vivo*, as *in vivo* populations are more likely to be organized rather than randomly mixed, such as in a tissue. Measuring an entire population in a replicate from a set of replicates over time (to then use another replicate's entire population at the next measurement time) can improve the issue of sampling from part of a population; however, it is still possible that there may be differences between replicate phenotypic proportions that cannot be measured in this experimental setup. Overall, evaluating phenotypic transitions in ways that destroy the cells require inferring that such transitions occur based on population changes, rather than witnessing a change in a single cell.

For witnessing the change of a cell population or an individual cell in the same cell or group of cells, *in vitro* methodologies that keep cells alive have also been used. Lineage tracing using DNA barcoding has been performed: because the DNA barcoding process labels one cell uniquely, after labeling and elapsed time the presence of cells with the same barcode but different phenotypes indicates that that a unique cell with a unique phenotype gave rise to progeny at least one of which changed its phenotype via a transition [239]. Fluorescence-activated cell sorting (FACS) is a common methodology for assessing changes in cell populations without destroying the cells, where populations are sorted to include just one phenotype, then if another or multiple phenotypes are present over time, it demonstrates that phenotypic transitions have occurred [16, 15, 51, 10, 13, 18]. FACS and subsequent analyses are most easily performed using fluorescent antibodies directed toward cell surface proteins, but can also be performed using fluorescent molecules conjugated to or expressed alongside intracellular proteins. *In situ* immunofluorescence and following single cells using video or sequential image capture has been used [45, 233].

While live-cell sorting or imaging relieves some of the issues with cell-destroying methodologies, there remain drawbacks to these methodologies, including specifically in the case of SCLC. For DNA barcoding, the presence of more than one phenotype with the same barcode could be the result of a barcode having been transfected into more than one cell. Using large numbers of barcodes relative to the number of transfected

cells makes it very unlikely that the same barcode will enter two cells, [239] but it is not impossible. Assuring that each cell received a unique barcode would make further experiments impossible as reading the barcode by DNA or RNA sequencing requires destroying the cell. In sorting cells and later evaluating if multiple phenotypes are present in a once-single-phenotype population, while FACS is highly reliable, it is not infallible, [240] and multiple phenotypes in a population may come from outgrowth of an incorrectly-sorted cell that was left in the population. For SCLC in particular, FACS or analytical flow cytometry require transfection of the cell with a fluorescent gene as described above. This is due to the fact that the SCLC phenotypes are delineated by transcription factors, which will not appear on the cell surface to be marked by antibodies while the cells are alive; SCLC analyses sorting different phenotypes were performed using a fluorescent gene expressed from a general non-neuroendocrine marker promoter [10]. FACS is still likely the optimal high-throughput methodology for assessing phenotypic transitions, and in SCLC this would require use of one or multiple fluorescence genes conjugated to or expressed with phenotypic subtype-identifying transcription factors. The methodology where the investigator can be the most certain that a cell undergoes a phenotypic transition is that of single-cell image tracking.

Recently, Gopal and colleagues demonstrated that fluorescently labeling live SCLC cells is feasible, and tracked them to demonstrate phenotypic transitions (including some nonhierarchical transitions predicted by my SCLC population dynamics model) [233]. I have worked toward developing a system for studying SCLC phenotypic transitions in a similar way, with the goal to validate a subset of predictions from the population dynamics model. While this work is incomplete, I include it here for future reference.

A.3 Ensuring planned fluorescence represents NCI-H69 ASCL1 is expressed and functional

The goal is to have the SCLC cell fluoresce when ASCL1 is expressed. I aim to perform this in NCI-H69 cells [234] via CRISPR transfection. There are several ways that, post-procedure, appearance or disappearance of fluorescence may not represent a phenotypic transition. To ensure that fluorescence changes will represent phenotypic transitions, I will utilize the following strategies.

Post-transfection and sorting for fluorescent cells, I will perform genomic sequencing to ensure that the reporter underwent homology-directed repair into ASCL1. The cells may still fluoresce even if the reporter was inserted incorrectly, in which case they will not represent ASCL1 expression. Per the Cancer Dependency Map (DepMap; https://depmap.org/portal/) [241] there are two alleles of ASCL1 in NCI-H69 cells, so I will be transfecting in two reporters (one green and one red). If I only used one reporter, and it only inserted into one allele, a lack of fluorescence in the cells may not mean a lack of ASCL1 expression as I expect it to - in such a case, it could be that the allele without reporter insertion was transcribed. Using two reporters (and confirming appropriate insertion) I will know that a lack of fluorescence truly means a lack of ASCL1

transcription.

I will also ensure that the fluorescent reporters do not interfere with ASCL1 activity. I have designed the insertion such that the reporters follow a 2A peptide (P2A), so although ASCL1 will be transcribed with the reporter, during translation the ribosome will skip and two separate peptide chains will be made, so the fluorescent reporter is not attached to the ASCL1 protein [242]. Nonetheless, I need to ensure that nothing has altered ASCL1 function during the process. I will perform ChIP-seq on ASCL1 in the parental cells (no reporter insertion) and the reporter-inserted cells to ensure that ASCL1 is located at all the genomic regions it would be found, were the reporters not inserted.

A.4 Methods, their background, and considerations

A.4.1 Inserting a reporter via homology-directed repair

I want to insert a fluorescent gene into ASCL1, at the end of the region of the gene that will be translated into protein, using homology-directed repair (HDR). Double-stranded breaks in the cell's DNA require repair to ensure proper DNA replication and transcription, and the repair pathways are generally considered to be HDR or nonhomologous end joining (NHEJ) [243]. In HDR, an exchange of DNA sequences occurs between the DNA molecule with a break and an intact DNA molecule identical to the DNA with the break [243]. The intact DNA molecule is usually the sister chromatid, which has the highest likelihood of error-free repair as an identical sequence; HDR usually occurs in the S or G2 cell cycle phases, since these are the phases where a sister chromatid is present as it has been generated during replication [243]. Creating a double-strand break and providing exogenous DNA with regions identical to those regions around the break will enable the insertion of that exogenous DNA (as in **Figure A.2B**).

A.4.2 Sequencing ASCL1

Knowing the exact sequence of the relevant portions of ASCL1 in the NCI-H69 cell lines I have will provide the highest chance of success, as I will be able to design a sequence with the highest likelihood of homology in these particular cells. While the sequence of ASCL1 has been published, (Ensembl release 108; [244]) it is possible that the patient from whom the NCI-H69 cell line was derived had a sequence variation in ASCL1 compared to the published gene, and further, since NCI-H69 is an immortalized cell line, [234] over time in culture it could have accumulated one or more mutations in the gene that would lead it to differ from the published sequence [245]).

Because I want the fluorescent protein to be transcribed and translated when ASCL1 is transcribed and translated, I aimed to insert the fluorescent marker at the end of the ASCL1 translated region. This results in the relevant portions of ASCL1, as mentioned above, being the end of translated ASCL1 (up to the stop

codon). ASCL1 has two exons and one intron, so both exons will be transcribed; however, the stop codon is located at the end of the first exon, so only the one exon will be translated (Ensembl release 108; [244]; **Figure A.1A**). Therefore, the region of interest is the middle/end of ASCL1's first exon and the beginning of its intron. I sequenced these regions using GenHunter (Nashville, TN) and was able to sequence with high quality across several replicates from base pairs 491 to 1291 and from 1073 to 1932, covering the last $\frac{2}{3}$ of the first exon, and then the end of the first exon, the intron, and about $\frac{1}{4}$ of the second exon (Ensembl release 108; [244]).

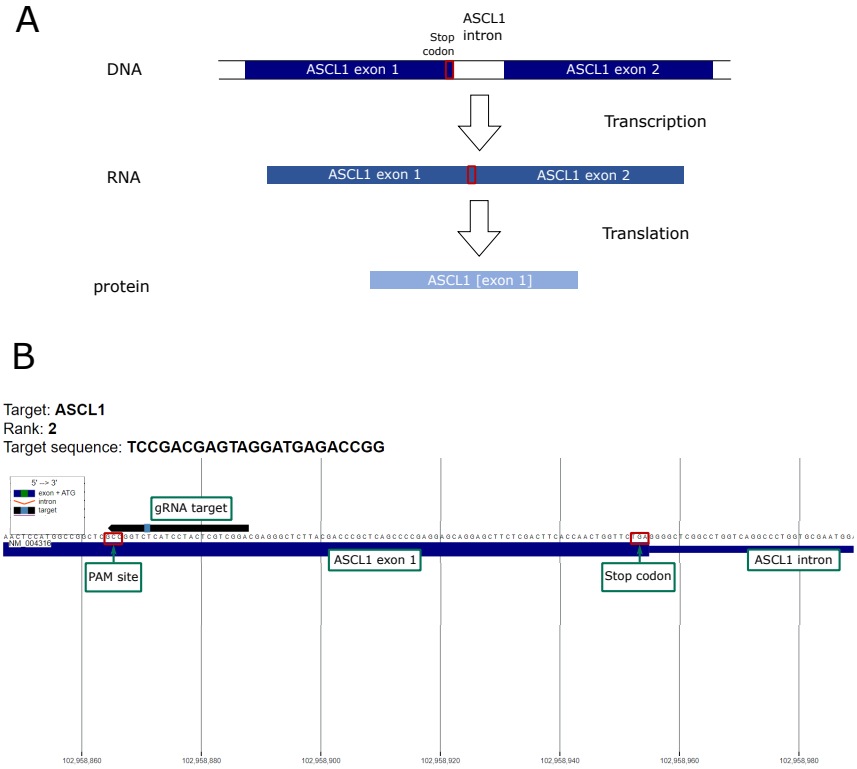


Figure A.1: ASCL1 genetic locus and annotated CHOPCHOP screenshot.

A ASCL1 gene and the process of transcription and translation. Both exons are transcribed into mRNA, but the stop codon midway through the transcript halts translation, so only ASCL1's first exon contributes to its protein structure.

B Screenshot of CHOPCHOP, a web tool for selecting target sites for CRISPR/Cas9 (https://chopchop.cbu.uib.no/), plus additional annotations. Showing the ASCL1 locus (end of exon 1 and beginning of intron) with the algorithmically-determined gRNA + PAM target sequence and its location.

A.4.3 Using CRISPR for insertion

I aim to insert the fluorescent proteins using CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats, and CRISPR-associated protein). CRISPR/Cas9 has been demonstrated to be an efficient tool for genome editing, where the Cas9 protein creates a double-strand break in the genome that initiates the cell's

DNA repair process [246]. As noted above, the double-strand break is necessary for HDR to occur, and the Cas9 protein will provide this break in the DNA strand. Where Cas9 will generate the double-strand break depends upon an RNA molecule that recognizes specific sequences in the DNA (crRNA) and a tracrRNA that pairs with crRNA and enables Cas9 to interface with the RNA. Now, a chimeric guide RNA (gRNA) containing all functions of crRNA and tracrRNA can be purchased, or hybridized from crRNA and tracrRNA, for an investigator to apply CRISPR/Cas9 in cells of interest [246]. The recognition site of a gRNA is approximately 20 base pairs that must contain the complementary sequence of the genomic target, and in addition a protospacer adjacent motif (PAM) must be adjacent to the target for the Cas protein itself to recognize [247]. However, the gRNA and Cas protein are able to complement and bind, respectively, genomic sites with base-pair mismatches, so there can be off-target binding relative to the sequence an investigator designs for the gRNA [247].

Finding the appropriate sequence for CRISPR/Cas9 to generate the double-stranded break that will allow HDR, with less off-target recognition and more on-target efficiency, (better recognizing and inducing breaks) is thus an important task. Labun and colleagues note that efficient targeting depends on "the position of specific nucleotides in the target sequence, the accessibility of the target site, and the sequence of its flanking regions" [248]. Large-scale studies measuring the efficiency of Cas9, along with machine learning-based methods, have enabled optimization of Cas protein cutting and this can be investigated via the web tool CHOPCHOP [248, 247]. I used CHOPCHOP to find gRNAs that would enable a Cas9 cut near the end of ASCL1, while prioritizing targeting and cutting efficiency. I chose the second highest-ranking gRNA returned by CHOPCHOP, which was 28 codons (84 base pairs) before the ASCL1 stop codon and on the reverse strand (**Figure A.1B**). Confirming via my ASCL1 sequencing results (**Section A.4.2**) that the gRNA sequence was identical and in the same location in my NCI-H69 cells as it is in CHOPCHOP's ASCL1 sequence, I purchased a gRNA with the recommended recognition sequence from Synthego (https://www.synthego.com/).

A.4.4 Design of the HDR construct

The Cas9 protein will make a double-strand break in the DNA at the PAM site adjacent to the complement of the gRNA sequence, and concurrently the donor DNA template, with homology to the area surrounding the cut, must be present to recombine as part of the HDR process. Because the structure of the DNA molecule allows it to loop, if the homology regions are present surrounding additional DNA, the homologous regions can recombine with the cut DNA strand and bring the additional DNA with them. Thus, the donor DNA will be comprised of two homology regions toward ASCL1, surrounding the fluorescent marker to insert into the end of ASCL1 (see **Figure A.2**).

The first consideration for the exogenous construct is the size of the homology regions. The 5' homology arm must include base pairs up to the gRNA recognition site, at which point everything 3' is considered part of the insertion (until the 3' homology arm). The design of the homology arm is a tradeoff between accuracy (longer homology arms) and efficiency (shorter arms). Some publications recommend 200 base pairs of homology [249] while others recommend 1 kilobase [250]. The sequencing I had undertaken allowed me to design homology arms of 700 base pair length on both the 5' and 3' ends of the exogenous construct. The downstream homology arm must then encompass everything after the PAM site, which includes the sequence recognized by the gRNA.

Because the gRNA site is 28 codons from ASCL1's stop codon, the first part of the donor DNA to insert must be the remaining 28 codons of ASCL1 before its stop codon. However, an important consideration with regard to the codons to insert is that the first 20 base pairs (essentially first 7 codons) of the insertion should not match the gRNA recognition sequence. If so, the Cas9 protein could target the donor DNA template (since the insertion will have the matching gRNA site) and if the DNA is successfully recombined, remaining Cas9 protein within the cell could cut the newly inserted DNA [247]. Therefore, the first 7 codons must remain the same, so that the ASCL1 protein will continue to function properly after donor insertion, but the base pairs must be altered. Due to the degeneracy of the DNA code, I altered the DNA sequence across the first 7 codons, making 8 base pair changes out of the 20 in the gRNA recognition sequence. As already noted, Cas9 and the gRNA may still bind and recognize DNA with mismatches, but the likelihood is decreased by changing 8 out of 20 base pairs.

Moving 5' to 3', the stop codon is not included after ASCL1's final 28 codons in the insert because the stop codon should not be present until after any DNA intended for translation - since an inserted fluorescent protein will need to be translated, the stop codon will be encoded after it. The sequence for P2A peptide (human optimized) is then included between the pre-stop-codon end of ASCL1 and the fluorescent marker. A P2A peptide induces ribosomal skipping, so it will separate the peptide chains before and after the P2A sequence [242]. The fluorescent protein will thus be translated whenever ASCL1 is, but it will not be conjugated to ASCL1. The benefit of this decision is that it is less likely the fluorescent protein will interfere with ASCL1's normal function, though the drawback is that the fluorescent protein will not translocate with ASCL1. Fluorescence in the cell will indicate that ASCL1 is expressed, but not the nature of its activity. I will use two fluorescence markers (see **Section A.3**) and selected human-optimized mKate2 and mAzamiGreen, inserting each sequence into its own HDR construct [251, 252, 253]. An SGS linker is placed between each noted sequence (not shown), as spacers between each domain [254], leading to the 5'-to-3' organization of the insert as such: remainder of ASCL1, SGS linker, P2A peptide, SGS linker, fluorescent protein. 3' to the fluorescent protein is the stop codon, as previously noted. See **Figures A.2A** and **A.2B** for a diagrammatic

depiction.

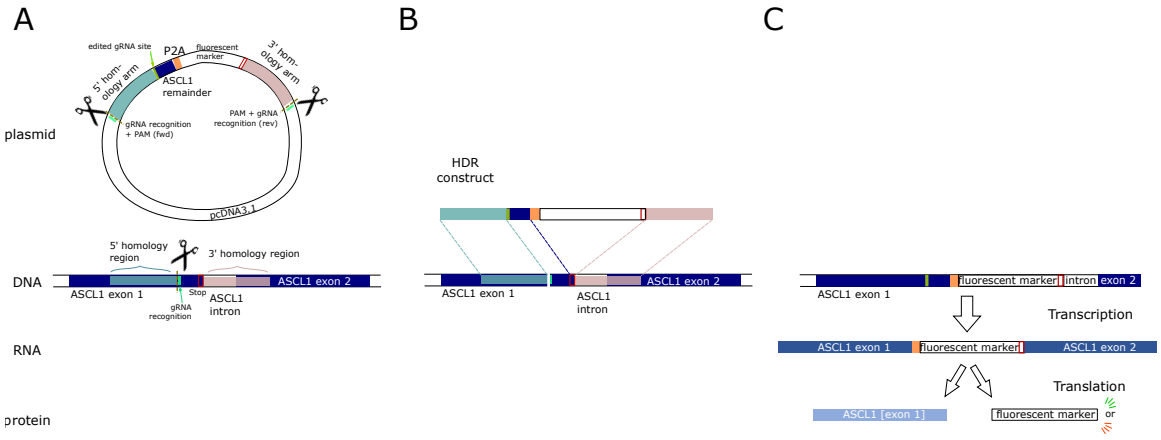


Figure A.2: Using CRISPR/Cas9 to insert a homology-directed repair construct into the cell that will result in translation of a fluorescent marker along with (but physically separate from) ASCL1.

A Above, plasmid that I designed with the HDR construct and Cas9 cut sites so that Cas9 will cut the construct out of the plasmid, (denoted by scissors) allowing it to recombine during homology-directed repair. Below, endogenous ASCL1 locus with homology regions highlighted and the Cas9 cut sites to generate the double-stranded break (scissors).

B Once Cas9 cuts the donor DNA from the plasmid, a double-stranded DNA fragment is present to undergo homology-directed repair of the double-stranded break.

C The resulting gene includes the ASCL1 exon 1 immediately followed by a P2A peptide and the fluorescent marker and stop codon, followed by the ASCL1 intron and then exon 2. Transcription leads to an mRNA molecule with exon 1, fluorescent protein, and exon 2. Translation then leads to generation of the ASCL1 protein and separately generation of the fluorescent protein.

A.4.5 Transfecting the cell

With the gRNA (**Section A.4.3**) and the HDR construct (**Section A.4.4**) designed, the next step is to plan to get the HDR construct to the location of the endogenous DNA as a fragment of double-stranded DNA. The exogeneous DNA will be delivered into the cell via transfection, and I chose to use a plasmid-based transfection for this experiment; viral DNA transfection is also possible, with the benefits of plasmid transfection being less immunogenicity and no viral integration into host genome, but drawbacks being potentially lower transfection efficiency [255]. The HDR construct within the plasmid will need to be excised to appear as the fragment of double-stranded DNA that will serve as the homologous donor DNA. The Cas9 protein already in use for targeting the endogenous DNA can be used to recognize the gRNA and PAM sequences if they flank the HDR construct: therefore, the 20-base pair gRNA sequence and PAM site appear in forward orientation before the 5' homology arm, and the PAM site and gRNA sequence in reverse orientation appear after the 3' homology arm (**Figure A.2A**, top). When Cas9+gRNA target the endogenous DNA, other Cas9+gRNA molecules will target the plasmid, generating a cut-out HDR construct (**Figure A.2B**) and a double-stranded break where the construct will be inserted (**Figure A.2A**, bottom, and **Figure A.2B**).

I purchased this HDR construct sequence within a pcDNA3.1 plasmid, generated by GenScript (https://www.genscript.com/), which I then transformed high-efficiency E. coli (New Enland Biosciences; https://www.neb.com) to take up each plasmid. I generated glycerol stocks of these transformed bacteria using Addgene's (https://www.addgene.org/) protocol and used the Qiagen (https://www.qiagen.com/us) maxiprep kit to purify plasmids from the bacteria for transfection into NCI-H69 cells.

A.4.6 Plasmid transfection

The plasmids will be transfected into the cells along with the ribonucleoprotein (RNP) complex of Cas9 (TrueCut Cas9 Protein v2, ThermoFisher Scientific; https://www.thermofisher.com) and gRNA (Synthego) [256]. I will use electroporation to transfect the cells. While several cell lines have optimized electroporation conditions published, this is not the case for NCI-H69 cells and I needed to perform the optimization myself.

A.4.6.1 Electroporation optimization

I followed the Neon Transfection System (Thermofisher Scientific) protocol for optimization, transiently transfecting NCI-H69 cells with a plasmid expressing mAzamiGreen under the EF1alpha promoter. Successful transfection with this plasmid results in green fluorescence within the cell. The Neon optimization protocol includes 24 electroporation conditions of varying voltage (from 850 V to 1700 V), pulse duration (from 10 ms to 40 ms), and pulse number (from 1 to 3), including a control with no pulse applied. I aimed to assess the viability of transfected cells and the transfection efficiency, which are generally at odds: the more voltage over time, the more plasmid will enter the cell (higher efficiency) but the more injury to the cell is likely to occur (lower viability).

After 72 hours, I visually inspected all 24 conditions using fluorescence microscopy (EVOS Cell Imaging System; ThermoFisher Scientific) and assessed the top 9 conditions. These conditions ranged from a voltage of 900 V to 1400 V, pulse duration of 10 ms to 40 ms and number from 1 to 3. I performed a second round of optimization on NCI-H69 cells using these 9 conditions, and at 72 hours assessed viability and efficiency via analytical flow cytometry, comparing to parental, untransfected NCI-H69 cells: viability was assessed based on forward and side scatter, with the gated population considered "live" cells, and efficiency was assessed by the percentage of live, single cells positive for fluorescence (GFP channel) (**Figure A.3A**). I then selected the best 4 conditions for a further round of optimization, transfecting NCI-H69 cells with these 4 conditions and at 72 hours performing the same analytical flow cytometry workflow. The best electroporation condition at 34% viable and 42% positive for fluorescence was at 1050 V, with 2 30-ms pulses (**Figure A.3B**).

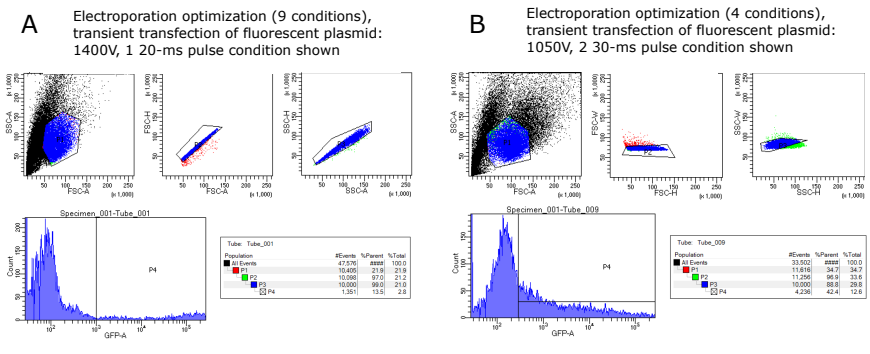


Figure A.3: Results from rounds 2 and 3 of electroporation optimization for ASCL1.

A The top 9 conditions from my running the Neon manual optimization protocol were tested. Viability was assessed via forward and side scatter (top left), with the gated population considered "live" cells. Efficiency was assessed by the percentage of live, single cells positive for fluorescence (GFP channel) (bottom left). One condition is shown.

B The top 4 conditions as assessed in A.3A were tested. Viability and efficiency were assessed similarly. The best condition, which will be used for electroporation of NCI-H69 cells going forward, is shown.

A.5 Next steps

The next step will be to perform CRISPR HDR using the Cas9, gRNA, and both fluorescent-marker HDR plasmids, with the optimized electroporation condition discussed. At least 72 hours after transfection, I will assess a sample of the cells using analytical flow cytometry to see that a proportion are double-positive, indicating that each allele of ASCL1 has been labeled. Depending on the proportion, I will grow the population: the smaller the proportion of double-positive cells, the more will need to be grown. I will then perform FACS to sort double-positive cells, so I can be sure that all my cells have ASCL1 alleles labeled. After this I will work to ensure that fluorescence is indeed indicative of ASCL1 expression and has not interfered with ASCL1 function (see **Section A.3**). Further sorting will be performed for experimental investigation of whether NCI-H69 cells undergo phenotypic transitions both away from (loss of fluorescence from a sorted, fluorescent population) or toward (gain of fluorescence from a non-fluorescent, sorted population) neuroendocrine, SCLC-A subtype identity.

Appendix B

Supplemental information for didactic example to contrast AIC vs posterior probability calculated by Bayes-MMI for model selection and multi-model inference

B.1 Comparing Galipaud et al., 2014 to our own analysis of the same simulated data

In **Chapter 5**, we use the example from [1]. We aimed to perform our own analysis on the model selection and sums of AIC weights problem presented in [1] and further discussed in [26]. The authors aimed to generate a "simulated ground truth" dataset where some predictor variables would be clearly necessary to include in a representative model for the data, while others would be unnecessary:

> We simulated a data set (sample size $n = 100$) including one response variable y and four predictor variables, x_1, x_2, x_3 and x_4. We controlled for the correlation structure both between the response variables and predictor variables and among predictor variables using a Cholesky decomposition (Genz & Bretz 2009). This method allows one predictor variable with a strong effect to be generated together with other variables with smaller tapering effects, as recommended by Burnham & Anderson (2002 p. 89, 2004). [1]

Galipaud and colleagues provide R code in the Supporting Information of their article for their simulations, and we used their `data.simulation` method to generate a dataset with `sample.size=100`, and using the same tapering predictor variable effects, $r_{y,x_1} = 0.70$, $r_{y,x_2} = 0.50$, $r_{y,x_3} = 0.05$, $r_{y,x_4} = 0.00$ [1]. This generates a dataset of which a snippet is below:

	y	x1	x2	x3	x4
1	5.927508	10.650031	9.644639	12.644369	9.585406
2	5.001812	10.255895	9.817701	9.236827	10.147333
3	5.540665	10.368334	11.221946	10.384068	10.537025
⋮	⋮	⋮	⋮	⋮	⋮
100	6.832196	11.740345	10.088608	11.213238	10.747818

Table B.1: Snippet of R simulated data array using code provided in [1]

With this simulated data, we ran nested sampling on the data to calculate Bayesian evidence (marginal likelihood) as well as AICc values. We consider our results (**Table B.2**) similar to [1] such that we can compare the two methods and the results in that publication.

β_0	x_1	x_2	x_3	x_4	k	rank, AICc	rank (AICc) in Galipaud et al. 2014	log(L)	AICc	Δ_i	w_i
-2.502	0.608	0.139			4	0	0	-48.764	103.778	0	0.313
-3.398	0.618	0.129	0.089		5	1	1	-47.728	103.877	0.099	0.298
-2.729	0.620	0.141		0.007	5	2	2	-48.776	105.973	2.195	0.105
-3.329	0.619	0.131	0.077	0.002	6	3	3	-47.739	106.115	2.338	0.097
-2.169	0.623		0.090		4	4	5	-50.264	106.778	3.001	0.070
-1.219	0.617				3	5	4	-51.362	106.848	3.070	0.068
-1.462	0.615			0.025	4	6	6	-51.299	108.847	5.069	0.025
-2.300	0.614		0.086	0.026	5	7	7	-50.241	108.903	5.125	0.024
3.357		0.164			3	8	8	-96.251	196.625	92.847	10^{-21}
2.990		0.164	0.038		4	9	9	-96.061	198.372	94.594	10^{-22}
3.355		0.164		0.001	4	10	11	-96.255	198.761	94.983	10^{-22}
3.043		0.163	0.029	0.004	5	11	13	-96.088	200.597	96.819	10^{-22}
4.990					2	12	10	-100.000	202.041	98.263	10^{-22}
4.602			0.039		3	13	12	-99.787	203.698	99.920	10^{-23}
4.901				0.009	3	14	14	-99.991	204.106	100.328	10^{-23}
4.510			0.041	0.007	4	15	15	-99.784	205.818	102.041	10^{-23}

Table B.2: Summary of nested sampling model selection results on the simulated dataset and model selection problem in Galipaud et al., 2014.

Candidate models are ranked by AICc as in Galipaud et al., 2014; rank in our analysis can be compared to the ranking in Galipaud et al., 2014 ("rank" vs. "rank in Galipaud et al., 2014" columns). Maximum log-likelihood parameter estimates and AICc are calculated from PyMultinest output. Parameter estimates are reported if present for each of the 16 candidate models. k, total number of estimable parameters; log(L), maximum log-likelihood returned by our least-squares likelihood function; AICc, AIC "corrected" for small sample size; Δ_i, $AICc - min(AICc)$ per model; w_i, Akaike weight.

B.2 SW and posterior probability on a subset of the candidate models

An important feature of SW is that it is most accurately calculated when the full candidate model set contains equal representation of every variable [24]. That is, x_1 if appears in 8 models, x_2, x_3, and x_4 must appear in 8 models to calculate an SW for each of them. That is the case in this example from [1]. However, we were interested in the case where not every variable may be represented equally in the candidate model set. Once 10 variables are present, a candidate set including equal representation of every variable would include more than 1000 models; more than 13 variables means the full candidate set includes more than 10,000 models. When many variables are present, it is possible to eliminate models from a candidate set, achieving a more tractable number of candidate models, using prior knowledge. We chose to impose synthetic prior knowledge on this 16-model candidate set, as if it is "known" that x_3 never appears along with x_4. Removing every model that includes both x_3 and x_4 left us with 12 models rather than 16, where x_1 and x_2 appeared in 6 models each, while x_3 and x_4 appeared in 4 models each.

We repeated our analysis on this subset of candidate models (**Table B.3**). Similar to our Bayesian analysis of the entire set of candidate models, x_4 does not appear until the fifth-highest ranked model; the four highest-ranked models having a cumulative probability of 0.938 (the majority of probability that the "true" model is

present); x_4 still appears only in candidate models whose posterior probability has decreased compared to the prior (**Table B.3**). The SW for x_4 in this analysis is 0.14, having changed from its SW of 0.25 using the whole candidate set (**Table B.4**). There does not appear to be an accepted threshold over which the SW for a variable indicates it should be included in a model, thus it seems practitioners choose a threshold on a per-case basis [1]. As such, it is difficult to assess what a change in SW from 0.25 to 0.14 means when using only a subset of the data – as noted, however, this is not an appropriate use for SWs anyway [24].

Model	$\log(L)$	AICc	Δ_i	w_i	rank, post. prob	rank, AICc	$\ln(Z)$	$\ln(Z)$ error (+/-)	Prior prob.	Post. prob.
$y = x_1 + x_2 + \beta$	-48.764	103.78	0	0.357	0	0	-57.280	0.048	0.083	0.459
$y = x_1 + \beta$	-51.362	106.85	3.070	0.077	1	4	-58.028	0.043	0.083	0.217
$y = x_1 + x_2 + x_3 + \beta$	-47.728	103.88	0.099	0.340	2	1	-58.250	0.054	0.083	0.174
$y = x_1 + x_3 + \beta$	-50.264	106.78	3.001	0.080	3	3	-58.936	0.049	0.083	0.088
$y = x_1 + x_2 + x_4 + \beta$	-48.776	105.97	2.195	0.119	4	2	-59.751	0.055	0.083	0.039
$y = x_1 + x_4 + \beta$	-51.299	108.85	5.069	0.028	5	5	-60.243	0.050	0.083	0.024
$y = x_2 + \beta$	-96.251	196.63	92.85	10^{-21}	6	6	-102.989	0.044	0.083	10^{-21}
$y = \beta$	-100.00	202.04	98.26	10^{-22}	7	9	-104.787	0.038	0.083	10^{-21}
$y = x_2 + x_3 + \beta$	-96.061	198.37	94.59	10^{-21}	8	7	-104.864	0.050	0.083	10^{-21}
$y = x_2 + x_4 + \beta$	-96.255	198.76	94.98	10^{-22}	9	8	-105.643	0.051	0.083	10^{-21}
$y = x_3 + \beta$	-99.787	203.70	99.92	10^{-23}	10	10	-106.782	0.045	0.083	10^{-21}
$y = x_4 + \beta$	-99.991	204.11	100.3	10^{-23}	11	11	-107.266	0.046	0.083	10^{-21}

Table B.3: Summary of AICc and nested sampling model selection results using a partial candidate set. Candidate models are ranked by posterior probability ("post. prob."). Ranking in this Bayesian analysis can be compared to ranking via AICc ("rank, post. prob.", vs. "rank, AICc"). Both AICc-related calculations (second through fifth column) and Bayesian calculations (eight through final column) are shown for the partial candidate set results. $\log(L)$, maximum log-likelihood; AICc, AIC "corrected" for small sample size; Δ_i, $AICc - min(AICc)$ per model; w_i, Akaike weight; $\ln(Z)$, the natural log of the Bayesian evidence/marginal likelihood (Z); $\ln(Z)$ *error*, the error returned by PyMultinest; prior prob., the prior probability that a model is the "correct" model; post. prob., posterior probability that the model is "correct", calculated as $\frac{Z_i * \Pr(i)}{\sum_j Z_j * \Pr(j)}$.

For our Bayesian analysis, the posterior probability that the inclusion of x_4 is supported by the data is 0.09, (**Table B.4**) and is the same for the full candidate set and for the "prior knowledge excluded" candidate set. Here, x_4 is just as unlikely to be included in the "true" model whether assessing a subset or the full set of candidate models. Interestingly, the posterior probabilities of all predictor variables $x_1 - x_3$ have identical values when analyzing a subset of the data (**Table B.4**). While this is a promising result, as it indicates that using only a subset of the data does not affect the results of Bayesian analysis as it does the SW analysis, more tests need to be done to determine whether this is generally true or true in only some cases. While there was no change, had there been one, Bayesian principles dictate that we can assess if this is a significant change (unlike the SW approach, where it is unclear what a change means). With equal prior probability per model variable, a posterior probability of 0.75 or more, or probability of 0.25 or less, would be considered substantial evidence for inclusion or exclusion of that variable, respectively ([178]; see **Section 5.6**). Thus,

a change in posterior probability per variable could be measured against that standard, denoting whether the data informs whether to include a variable in the model or not.

Variable	SW: Galipaud et al. 2014	SW: this manuscript	SW: candidate subset	Prior probability	Posterior probability	Posterior probability: candidate subset
x_1	1	1	1	0.5	1	1
x_2	0.94	0.81	0.82	0.5	0.67	0.67
x_3	0.37	0.49	0.42	0.5	0.28	0.42
x_4	0.37	0.25	0.15	0.5	0.09	0.12

Table B.4: SW and posterior probability calculations for each model variable in both full candidate set and partial candidate set examples.

Prior probability for a variable is set at 0.5, meaning a variable's prior probability for can be calculated per candidate model by dividing 0.5 by the number of models in which the variable appears. Prior probability values only impact posterior probability scores and not SW calculations.

Appendix C

Supplemental information: Unified Tumor Growth Mechanisms from Multimodel Inference and Dataset Integration

C.1 Parameter priors for SCLC mechanistic model of intratumoral heterogeneity.

NCI-H69 cell line (SCLC-A) doubles at 51.1 +/- 3.1 hours, equivalent to 0.469 doublings per day [257].
NCI-H82 (SCLC-N) doubles at 25.5 +/- 4.2 hours, equivalent to 0.783 doublings per day [257].
NCI-H841 (SCLC-Y) doubles at 31.2 +/- 3.6 hours, equivalent to 0.769 doublings per day [257].
DMS53 (SCLC-A2) doubles at $\sim$ 127 hours, equivalent to 0.1898 doublings per day [258].
Average of apoptotic indices for SCLC cell lines is 0.081, used as 0.081 deaths per day [259].
Little is known about timing of phenotypic transitions so a very permissive range was used in the model, with values considered uniformly likely from 0.01 transitions per day (1 transition every $\sim$ 3 months) to 3 transitions per day. We find this to be a reasonable permissive range based on mechanistic modeling of the epithelial-to-mesenchymal transition in breast cancer cells and stem cell differentiation [260, 261], as transition rates for SCLC subtypes have not been reported in the literature. In these reports, the EMT transition was fit to between 10 [260] and 20 [261] days.
Changes in growth rates specifically due to inter-subtype effects have not been recorded. NE viability (luciferase) and division (EdU incorporation) are increased by NonNE when plated together [10]; NonNE growth decreases in the presence of NE [179]. We used a 5% increase or decrease of the baseline value (depending on subtype and interaction) as the parameter prior for each affected rate.

Table C.1: Existing data pertaining to SCLC intratumoral heterogeneity and communication used for rate parameter priors.

C.2 Simulations using best-fitted parameters, as opposed to randomly-selected parameters from the prior distributions, replicate subtype proportions at steady state

Each different dataset (different model selection run) differentiated between more likely and less likely models for that particular dataset. Given that model selection determines Bayesian evidence scores by balancing model complexity with goodness of fit, we expected that the subtype proportions at steady state in the highest scoring models would replicate cell subtype proportions in the data.

Simulating all fitted models, using parameters selected only from the prior marginal distributions, the subtype proportions at steady state for each subtype tended to fall at 0 (0% of the simulated tumor) or 1 (100% of the simulated tumor), indicating that parameters chosen at random from the prior distributions do not fit the data (**Figure C.3**). However, in selecting parameters from the posterior, fitted, distributions – those representing the highest-scoring parameter sets – the simulations matched the data much more closely. There still remained simulations where the subtypes fell either at 0% or 100% of the simulated tumor, but the subtype steady state fell within the probabilistic representation of the data more of the time (**Figure C.3**).

For the TKO dataset, not taking into account subtype N, for which all simulations both using prior param-

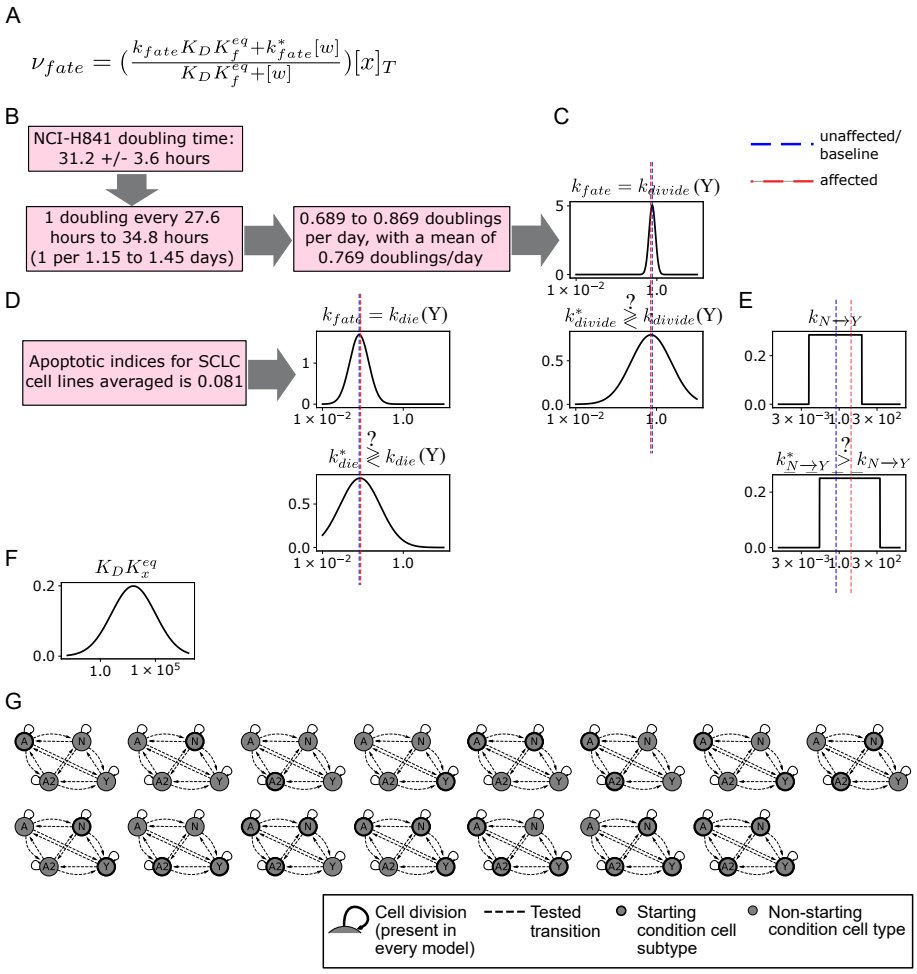


Figure C.1: Prior probabilities values and schematics.

A Rate of a cell fate (division, death, or phenotypic transition) for x (ν_{fate}) can be calculated as a function of the population size of the effector cell w (see **Figure 6.3**) [211].

B Example calculation of division rate parameter prior for H841, representation of subtype Y, (see **Table C.1**) converting doubling times to "per day" units.

C Division prior for subtype Y, (blue dashed line centered at the mean) as well as inter-subtype effect on division, whose mean is centered 5% lower (red dashed line; see **Table C.1**) with wider variance to account for more uncertainty in inter-subtype effects.

D Example calculation and visualization of death rate parameter prior for Y (blue dashed line at mean) and inter-subtype effect on death (red dashed line at mean, 5% higher).

E Example uniform transition prior, (see **Table C.1**) here showing N to Y transition; blue dashed line at baseline transition rate center, red dashed line at inter-subtype effect transition rate center.

F Equilibrium assumption prior, representing $K_D K_x^{eq}$ in the equation in (C.1A). Each affected interaction has a unique $K_D K_x^{eq}$ prior, but all such priors have identical values (centered at 1000) before fitting.

G Different model initiation hypotheses, where a model can be initiated by one or more subtypes (thick black outline) depending on the subtypes present (four-subtype topology shown). With 15 options and equal prior probabilities, the prior probability for each initiation hypothesis is 6.67%.

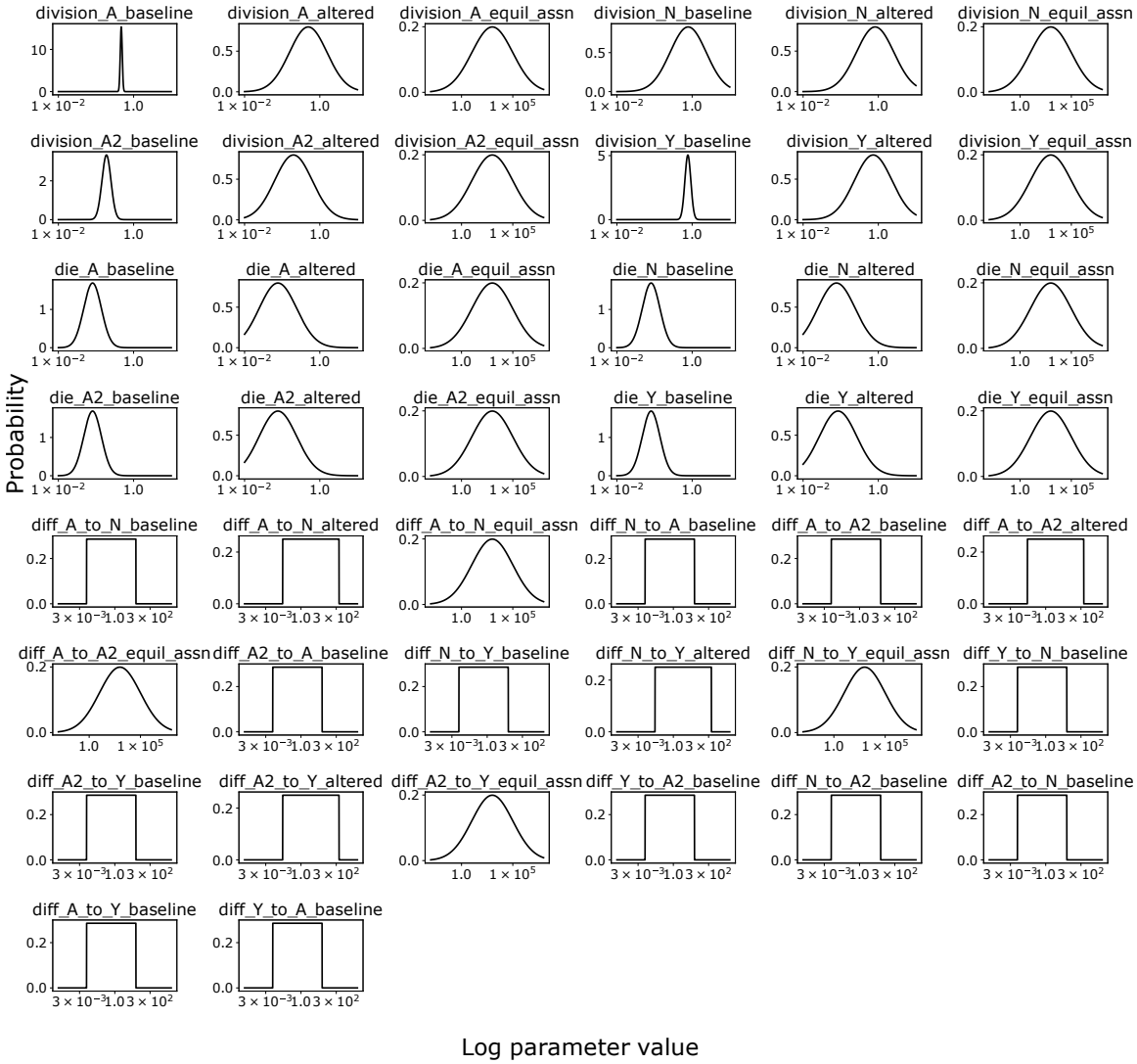


Figure C.2: Parameter prior distributions for all possible reactions in a candidate population dynamics model. If a candidate model does not contain a reaction, for example a model with the topology A, N, and Y does not include A2 and thus will not include A2 division, death, or transitions to/from A2, then the rate parameter priors for A2-related reactions will not be included as a parameter prior for model fitting.

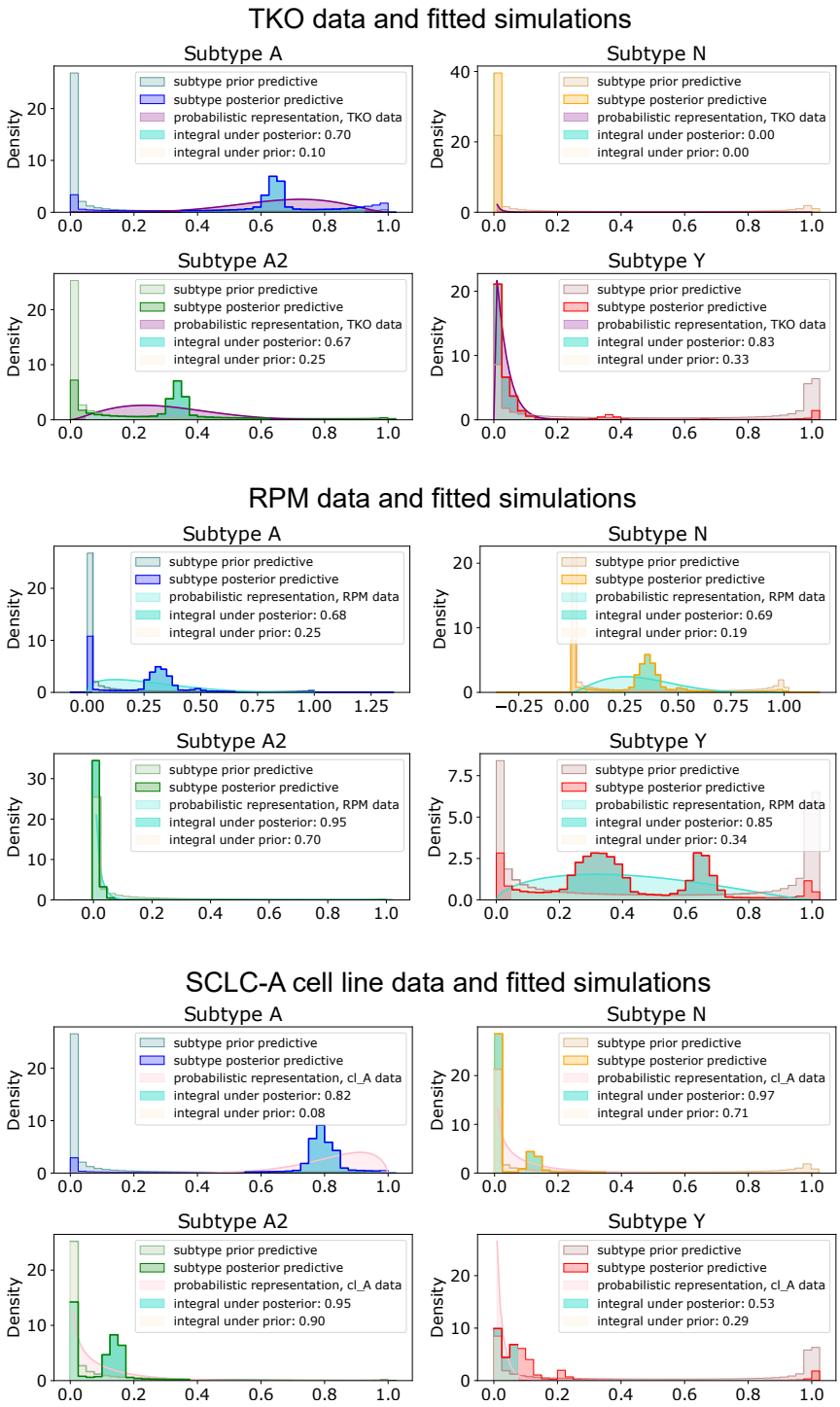


Figure C.3: Nested sampling's fitting results in better-fitting simulations than simulations using randomly selected parameter values.

Data distribution prior predictive distribution, and posterior predictive distribution for each dataset and all candidate models. Data is fit to a Beta distribution, bounded by zero and one, and used in the likelihood function input for Multinest (see **Methods Section 4.2.4**). Prior predictive distribution represents model simulations using parameters randomly drawn from the prior. Posterior predictive is generated by model simulations using best-fitting parameters returned by Multinest. See **Section C.2** for more detail.

eters and fitted parameters resulted in a steady state of 0%, on average 23% of the prior predictive density fell within the 95% confidence interval of the probabilistic representation of the data (**Figure C.3**, purple). After parameter fitting and evidence calculation, on average 73% of the posterior predictive density fell within the 95% confidence interval of the data. (Since the data, the prior predictive, and the posterior predictive are all probabilistic densities, each curve integrates to 1 or 100%.) Investigating these numbers it is clear that the model selection / parameter fitting process was unable to bring simulated subtype steady states completely within the 95% confidence interval of the TKO data subtype proportions, though it has brought it closer than the prior predictive. We would interpret this as the fact that the TKO data used for fitting was not able to completely outweigh the subtype proportions resulting from simulations using the prior parameters, which is a known possibility with regard to using Bayes' Theorem to fit a model's parameter sets to data [191].

For the RPM dataset, not taking into account subtype A2, for which all simulations both using prior parameters and fitted parameters resulted in a steady state of 0%, on average 26% of the prior predictive density fell within the 95% confidence interval of the probabilistic representation of the data. After parameter fitting and evidence calculation, on average 74% of the posterior predictive density fell within the 95% confidence interval of the data. Interestingly, the posterior predictive density for the Y subtype was bimodal (**Figure C.3**, middle), which we expect is related to the very wide range of the probabilistic representation of Y proportions in this dataset (**Figure 6.2B**).

For the SCLC-A clustered cell lines dataset, on average 50% of the prior predictive density fell within the 95% confidence interval of the probabilistic representation of the data. After parameter fitting and evidence calculation, on average 82% of the posterior predictive density fell within the 95% confidence interval of the data. In all, simulations using fitted parameters regardless of dataset resulted in a 30-50% better correspondence to subtype proportions in the data, indicating that the process of model selection and model averaging resulted in models and parameter sets that were able to represent the data at hand to a satisfactory extent.

C.3 All datasets support alteration of phenotypic transition rates in the presence of N or A2 subtypes

The highest likelihood model topologies (**Figure 6.5A**, blue) for the TKO GEMM data, along with the four-subtype topology, are compared in **Figure 6.5B** (left). Three model variables have significantly different parameter rates across model topologies: (i) the A-to-Y transition, (ii) the Y-to-A transition; and (iii) the A-to-A2 transition. The A-to-Y transition has a slower rate if A2 is present in the population and the Y-to-A transition has a faster rate if A2 is present. However, the presence of N along with A2 does not change the rate of either transition. Similarly, only Y affects the A-to-A2 transition, increasing its rate. N does decrease the rate of the A2-to-A transition despite no effect from Y. The Y-to-A2 transition rate is an increased in the presence of A2. These observations have mechanistic implications: A2 may represent an intermediate

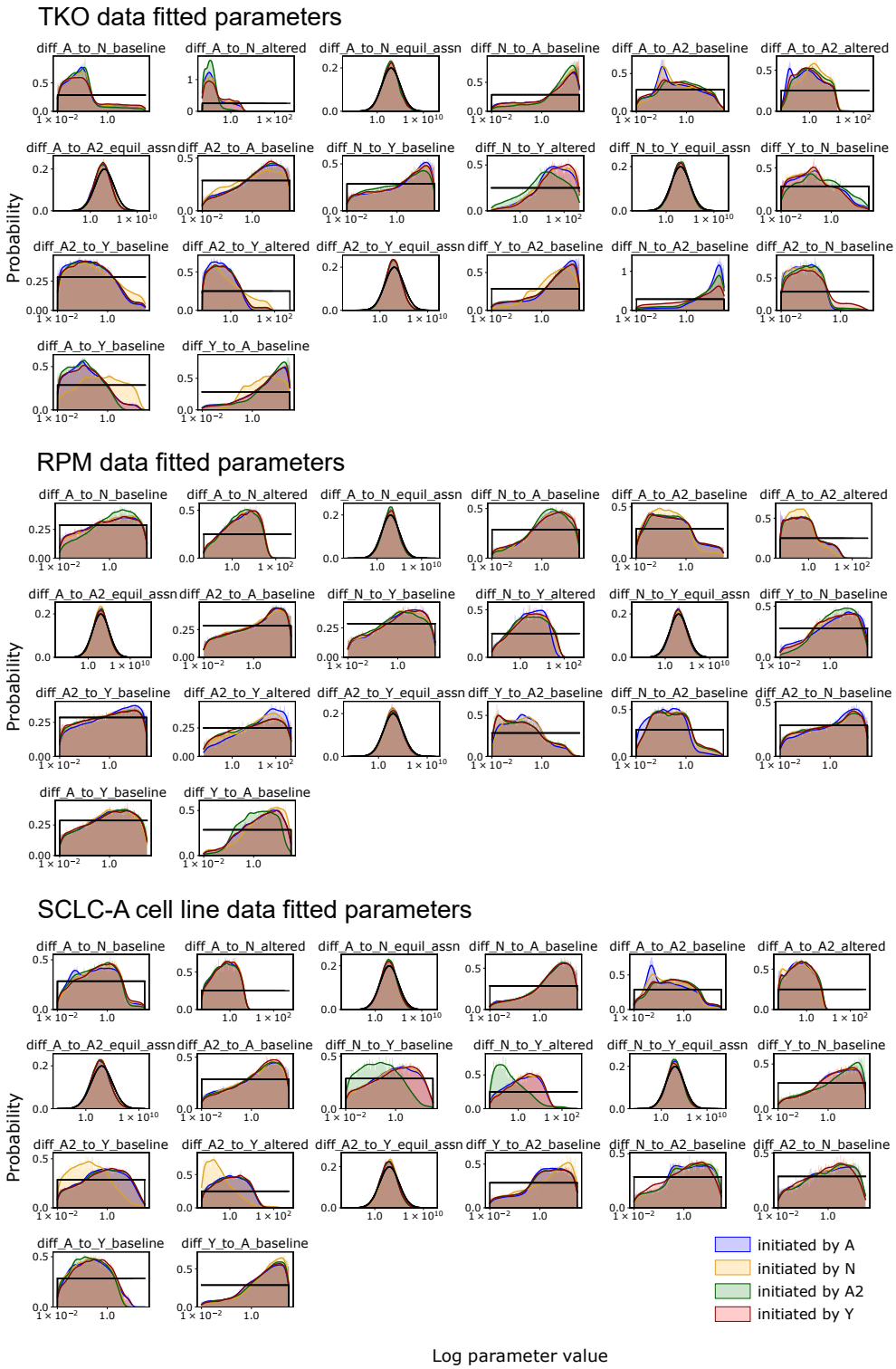


Figure C.4: Rate parameter posterior marginal distributions, after applying modeling averaging based on candidate model posterior probability.

Rate parameter distributions are plotted based on candidate model initiating subtype, where the subtype that initiates the tumor in the simulation results in nonidentical posterior distributions for some model parameters.

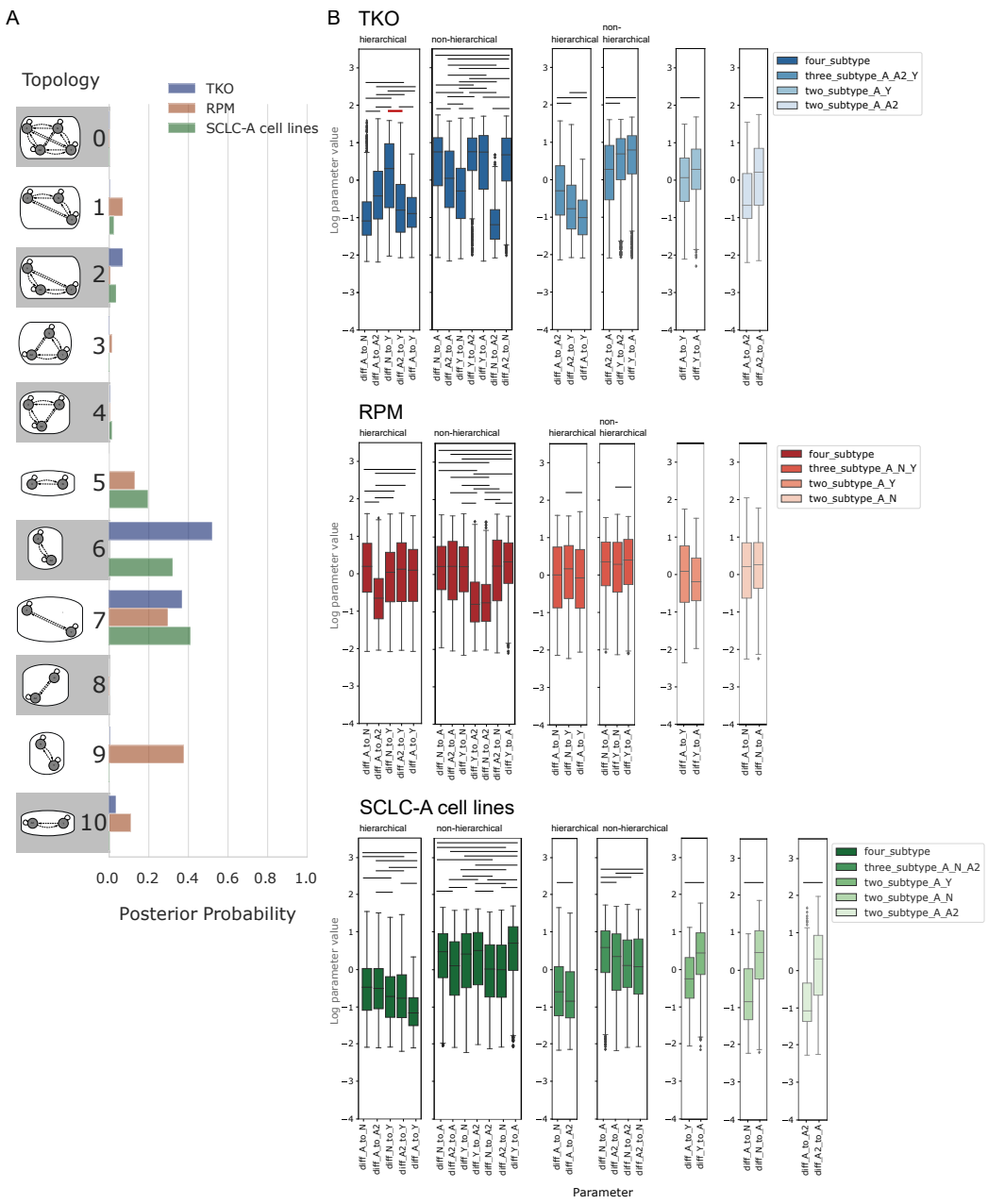


Figure C.5: Transition parameter rates vary in similar ways across datasets.

A Hypothesis assessment of model topologies per dataset, posterior probabilities based on all candidate models, with no filtering based on initiating subtype (see **Figure C.3**). Model topologies represented by images and corresponding numbers along the y-axis.

B Comparison of phenotypic transition parameter posterior marginal distributions, BMA-weighted, per dataset, separated by topology. In 3- and 4-subtype topologies, distributions are further separated by hierarchical or non-hierarchical transition status. Bars indicate significance between samples from BMA parameter distributions at family-wise error rate (FWER) of 0.01, using one-way ANOVA plus Tukey HSD.

Red bar: comparison noted in the main text.

subpopulation in the tumor that is longer-lived, and will only slowly transition to Y. In the topology with A, A2, and Y (**Figure 6.5A**, structure 2), the A-to-A2 transition takes up more of the flux in the network. Additionally, the N-to-Y transition is faster relative to the A2-to-Y transition (**Figure C.5B**), suggesting that N is a shorter-lived intermediate in the A-to-N-to-Y transition. This result aligns with previous experiments [11] where N was identified as a short-lived state in the A-to-N-to-Y transition. We therefore predict that A2 and N are involved in regulating the relative abundance of, and flux between, A and Y in the tumor.

We also compared the highest likelihood model topologies (**Figure 6.5A**, red) for RPM-fitted models, as well as the four-subtype topology (**Figure 6.5B**, middle). Five parameter rates are significantly different across model topologies. These are (i) the A-to-Y transition, (ii) the Y-to-A transition; (iii) the A-to-N transition; (iv) the N-to-Y transition; and (v) the N division rate. The same transitions affected in the TKO-fitted models (A-to-Y and Y-to-N transitions) are affected in the RPM models in the same ways, despite the experimental data being different. We thus predict that N and A2 are modulating the transition between, and relative abundance of, A and Y. Unlike in the TKO data, when A2 is present the flux through the system spends more time in the N subtype, with more frequent transitions to N and less frequent transitions to Y; additionally, the rate of the A2-to-N transition is faster than the N-to-A2 transition (**Figure C.5B**). We predict that while N may be a shorter-lived intermediate than A2, A2 regulates the flux from A-to-N-to-Y.

Next, we compared the highest likelihood model topologies (**Figure 6.5A**, green) for the SCLC-A cell line data and the four-subtype topology (**Figure 6.5B**, middle). Six parameter rates are significantly different across model topologies, five of which recapitulate rate alterations based on the presence or absence of different subtypes in TKO or RPM datasets. The rate alteration unique to the SCLC-A dataset is the A2-to-A transition, which is less frequent in the four-subtype topology, indicating that the presence of Y decreases this rate. However, in the TKO dataset, the A2-to-A transition decreases with addition of N.

C.4 Posterior probabilities for SCLC mechanistic model terms.

Model term	Calculated candidate model prior per hypothesis / summed prior	Hypothesis posterior probability (model-averaged)	Odds ratio
A to N transition	N/A	N/A	N/A
A to A2 transition	$\Pr(M \mid H_{A\to A2}) = 0.0008$, $\Pr(M \mid H_{no}) = 0.0006$ / sum 0.5 vs 0.5	$\Pr(H_{A\to A2} \mid D) = 0.62$	1.63
N to Y transition	N/A	N/A	N/A
A2 to Y transition	$\Pr(M \mid H_{A2\to Y}) = 0.0008$, $\Pr(M \mid H_{no}) = 0.004$ / sum 0.5 vs 0.5	$\Pr(H_{A2\to Y} \mid D) = 0.51$	1.04
A to Y transition	$\Pr(M \mid H_{A\to Y}) = 0.001$, $\Pr(M \mid H_{no}) = 0.003$ / sum 0.5 vs 0.5	$\Pr(H_{A\to Y} \mid D) = 0.64$	1.78
N to A2 transition	N/A	N/A	N/A
A2 to N transition	N/A	N/A	N/A
N to A transition	N/A	N/A	N/A
A2 to A transition	$\Pr(M \mid H_{A2\to A}) = 0.002$, $\Pr(M \mid H_{no}) = 0.001$ / sum 0.5 vs 0.5	$\Pr(H_{A2\to A} \mid D) = 0.71$	2.45
Y to N transition	N/A	N/A	N/A
Y to A2 transition	$\Pr(M \mid H_{Y\to A2}) = 0.0015$, $\Pr(M \mid H_{no}) = 0.0013$ / sum 0.5 vs 0.5	$\Pr(H_{Y\to A2} \mid D) = 0.68$	2.13
Y to A transition	$\Pr(M \mid H_{Y\to A}) = 0.0018$, $\Pr(M \mid H_{no}) = 0.0011$ / sum 0.5 vs 0.5	$\Pr(H_{Y\to A} \mid D) = 0.81$	4.26
Non-NE affects division, death	$\Pr(M \mid H_{div_eff}) = 0.0012$, $\Pr(M \mid H_{no}) = 0.0016$ / sum 0.5 vs 0.5	$\Pr(H_{div_eff} \mid D) = 0.44$	0.78
Y affects division & death vs A2&Y affect division& death	$\Pr(M \mid H_{div_eff_Y}) = 0.0016$, $\Pr(M \mid H_{div_eff_A2_Y}) = 0.0016$, $\Pr(M \mid H_{no}) = 0.0011$ / sum 0.33 vs 0.33 vs 0.33	$\Pr(H_{div_eff_Y} \mid D) = 0.28$, $\Pr(H_{div_eff_A2_Y} \mid D) = 0.33$	0.39, 0.50
Non-NE affects early transitions (A to N, A to A2)	$\Pr(M \mid H_{early_eff}) = 0.0011$, $\Pr(M \mid H_{no}) = 0.0018$ / sum 0.5 vs 0.5	$\Pr(H_{early_eff} \mid D) = 0.35$	0.54
Y affects early transitions (A to N, A to A2) vs A2 & Y affect these	$\Pr(M \mid H_{early_eff_Y}) = 0.0015$, $\Pr(M \mid H_{early_eff_A2_Y}) = 0.0015$, $\Pr(M \mid H_{no}) = 0.0012$ / sum 0.33 vs 0.33 vs 0.33	$\Pr(H_{early_eff_Y} \mid D) = 0.24$, $\Pr(H_{early_eff_A2_Y} \mid D) = 0.28$	0.32, 0.39
Non-NE affects late transitions (N to Y, A2 to Y)	$\Pr(M \mid H_{late_eff}) = 0.0021$, $\Pr(M \mid H_{no}) = 0.0010$ / sum 0.5 vs 0.5	$\Pr(H_{late_eff} \mid D) = 0.29$	0.41
Y affects late transitions (N to Y, A2 to Y) vs A2 & Y affect these	$\Pr(M \mid H_{late_eff_Y}) = 0.0028$, $\Pr(M \mid H_{late_eff_A2_Y}) = 0.0027$, $\Pr(M \mid H_{no}) = 0.0007$ / sum 0.33 vs 0.33 vs 0.33	$\Pr(H_{late_eff_Y} \mid D) = 0.21$, $\Pr(H_{late_eff_A2_Y} \mid D) = 0.24$	0.27, 0.32
If Non-NE effect is real, is it more likely coming from Y or from A2&Y?	$\Pr(M \mid H_{eff_from_Y}) = 0.0015$, $\Pr(M \mid H_{eff_from_A2_Y}) = 0.0015$ / sum 0.5 vs 0.5	$\Pr(H_{eff_from_Y} \mid D) = 0.47$, $\Pr(H_{eff_from_A2_Y} \mid D) = 0.53$	0.89, 1.13

Table C.2: Model variable posterior probabilities after hypothesis exploration using multimodel inference, TKO data in high-probability 3-subtype topology.

Model term	Calculated candidate model prior per hypothesis / summed prior	Hypothesis posterior probability (model-averaged)	Odds ratio
A to N transition	$\Pr(M \mid H_{A\to N}) = 0.0014$, $\Pr(M \mid H_{no}) = 0.0086$ / sum 0.5 vs 0.5	$\Pr(H_{A\to N} \mid D) = 0.71$	2.45
A to A2 transition	N/A	N/A	N/A
N to Y transition	$\Pr(M \mid H_{N\to Y}) = 0.0015$, $\Pr(M \mid H_{no}) = 0.0068$ / sum 0.5 vs 0.5	$\Pr(H_{N\to Y} \mid D) = 0.6$	1.50
A2 to Y transition	N/A	N/A	N/A
A to Y transition	$\Pr(M \mid H_{A\to Y}) = 0.0017$, $\Pr(M \mid H_{no}) = 0.0042$ / sum 0.5 vs 0.5	$\Pr(H_{A\to Y} \mid D) = 0.71$	2.45
N to A2 transition	N/A	N/A	N/A
A2 to N transition	N/A	N/A	N/A
N to A transition	$\Pr(M \mid H_{N\to A}) = 0.0027$, $\Pr(M \mid H_{no}) = 0.0022$ / sum 0.5 vs 0.5	$\Pr(H_{N\to A} \mid D) = 0.78$	3.55
A2 to A transition	N/A	N/A	N/A
Y to N transition	$\Pr(M \mid H_{Y\to N}) = 0.0028$, $\Pr(M \mid H_{no}) = 0.0022$ / sum 0.5 vs 0.5	$\Pr(H_{Y\to N} \mid D) = 0.68$	2.13
Y to A2 transition	N/A	N/A	N/A
Y to A transition	$\Pr(M \mid H_{Y\to A}) = 0.0033$, $\Pr(M \mid H_{no}) = 0.0020$ / sum 0.5 vs 0.5	$\Pr(H_{Y\to A} \mid D) = 0.81$	4.26
Non-NE affects division, death	$\Pr\big(M \mid H_{div_eff}\big) = 0.0025$, $\Pr(M \mid H_{no}) = 0.0024$ / sum 0.5 vs 0.5	$\Pr\big(H_{div_eff} \mid D\big) = 0.46$	0.85
Y affects division & death vs A2&Y affect division& death	N/A (since A2 as part effect not possible, calculations are same as Non-NE affects division, death, above	N/A	N/A
Non-NE affects early transitions (A to N, A to A2)	$\Pr\big(M \mid H_{early_eff}\big) = 0.0023$, $\Pr(M \mid H_{no}) = 0.0027$ / sum 0.5 vs 0.5	$\Pr\big(H_{early_eff} \mid D\big) = 0.44$	0.79
Y affects early transitions (A to N, A to A2) vs A2 & Y affect these	N/A (since A2 as part effect not possible, calculations are same as Non-NE affects early transitions, above	N/A	N/A
Non-NE affects late transitions (N to Y, A2 to Y)	$\Pr\big(M \mid H_{late_eff}\big) = 0.0042$, $\Pr(M \mid H_{no}) = 0.0017$ / sum 0.5 vs 0.5	$\Pr\big(H_{late_eff} \mid D\big) = 0.37$	0.59
Y affects late transitions (N to Y, A2 to Y) vs A2 & Y affect these	N/A (since A2 as part effect not possible, calculations are same as Non-NE affects late transitions above	N/A	N/A
If Non-NE effect is real, is it more likely coming from Y or from A2&Y?	N/A (A2 as part of the Non-NE effect not possible)	N/A	N/A

Table C.3: Model variable posterior probabilities after hypothesis exploration using multimodel inference, RPM data in high-probability 3-subtype topology.

Model term	Calculated candidate model prior per hypothesis / summed prior	Hypothesis posterior probability (model-averaged)	Odds ratio
A to N transition	$\Pr(M \mid H_{A\to N}) = 0.0031$, $\Pr(M \mid H_{no}) = 0.011$ / sum 0.5 vs 0.5	$\Pr(H_{A\to N} \mid D) = 0.61$	1.56
A to A2 transition	$\Pr(M \mid H_{A\to A2}) = 0.0031$, $\Pr(M \mid H_{no}) = 0.011$ / sum 0.5 vs 0.5	$\Pr(H_{A\to A2} \mid D) = 0.52$	1.10
N to Y transition	N/A	N/A	N/A
A2 to Y transition	N/A	N/A	N/A
A to Y transition	N/A	N/A	N/A
N to A2 transition	$\Pr(M \mid H_{N\to A2}) = 0.0038$, $\Pr(M \mid H_{no}) = 0.0064$ / sum 0.5 vs 0.5	$\Pr(H_{N\to A2} \mid D) = 0.66$	1.94
A2 to N transition	$\Pr(M \mid H_{A2\to N}) = 0.0038$, $\Pr(M \mid H_{no}) = 0.0064$ / sum 0.5 vs 0.5	$\Pr(H_{A2\to N} \mid D) = 0.66$	1.94
N to A transition	$\Pr(M \mid H_{N\to A}) = 0.0058$, $\Pr(M \mid H_{no}) = 0.0041$ / sum 0.5 vs 0.5	$\Pr(H_{N\to A} \mid D) = 0.74$	2.85
A2 to A transition	$\Pr(M \mid H_{A2\to A}) = 0.0058$, $\Pr(M \mid H_{no}) = 0.0041$ / sum 0.5 vs 0.5	$\Pr(H_{A2\to A} \mid D) = 0.61$	1.56
Y to N transition	N/A	N/A	N/A
Y to A2 transition	N/A	N/A	N/A
Y to A transition	N/A	N/A	N/A
Non-NE affects division, death	$\Pr\big(M \mid H_{div_eff}\big) = 0.005$, $\Pr(M \mid H_{no}) = 0.0046$ / sum 0.5 vs 0.5	$\Pr\big(H_{div_eff} \mid D\big) = 0.16$	0.19
Y affects division & death vs A2&Y affect division& death	N/A (since Y effect not possible, calculations are the same as Non-NE affects division & death, above)	N/A	N/A
Non-NE affects early transitions (A to N, A to A2)	$\Pr\big(M \mid H_{early_eff}\big) = 0.0051$, $\Pr(M \mid H_{no}) = 0.0045$ / sum 0.5 vs 0.5	$\Pr\big(H_{early_eff} \mid D\big) = 0.17$	0.21
Y affects early transitions (A to N, A to A2) vs A2 & Y affect these	N/A (since Y effect not possible, calculations are the same as Non-NE affects early transitions, above)	N/A	N/A
Non-NE affects late transitions (N to Y, A2 to Y)	N/A (Y not in this model so no transitions toward it)	N/A	N/A
Y affects late transitions (N to Y, A2 to Y) vs A2 & Y affect these	N/A (Y not in this model so no transitions toward it)	N/A	N/A
If Non-NE effect is real, is it more likely coming from Y or from A2&Y?	N/A (Y effect not possible)	N/A	N/A

Table C.4: Model variable posterior probabilities after hypothesis exploration using multimodel inference, SCLC-A cell line data in high-probability 3-subtype topology.

Appendix D

Supplemental Figures and Tables: Evaluating small cell lung carcinoma time-course patient-derived xenografts reveals phenotype switching with treatment and chemotherapy-stable subpopulations.

D.1 Supplemental Figures

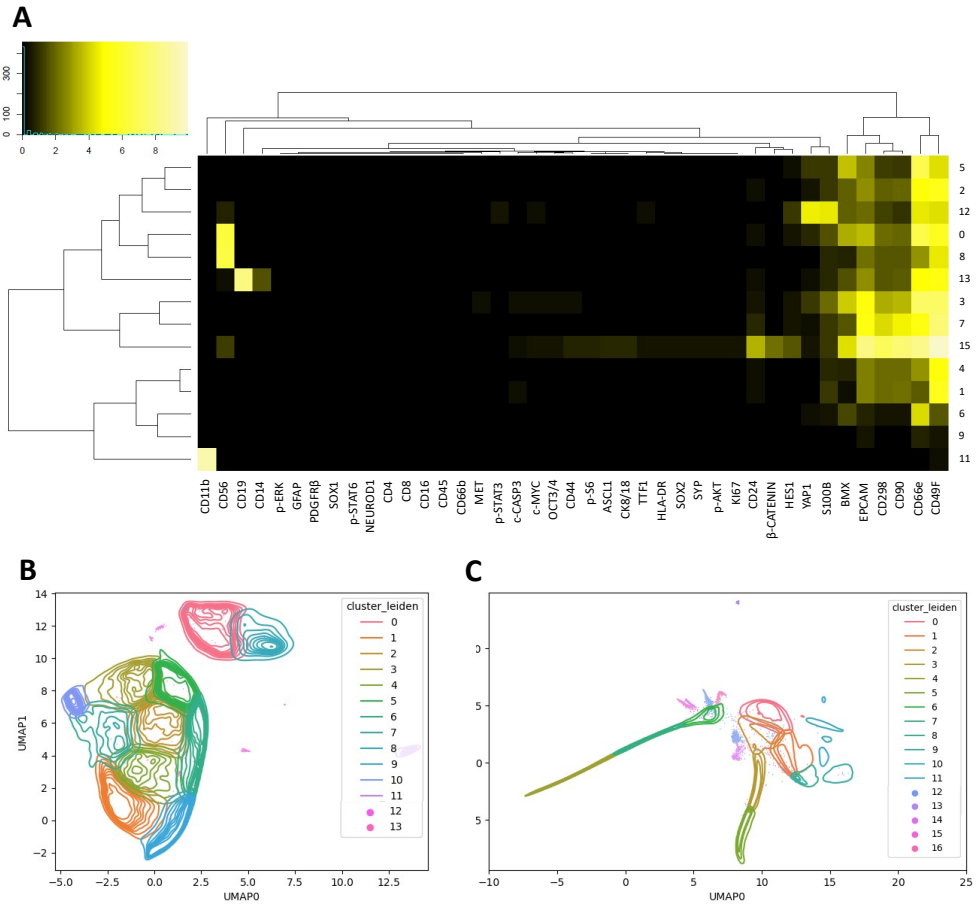


Figure D.1: LU73 MEM results and LX22 and LU73 cluster contour plots.

A A. Marker enrichment modeling (MEM) plot for all LU73 conditions and clusters, with black to yellow hue based on expression level per marker per cluster. Markers have been clustered based on expression level correlation between clusters (top dendrogram), and Leiden clusters of the combined dataset have been clustered based on similarity across markers (left dendrogram).

(D.1B,D.1C) Leiden clustering performed on the high-dimensional graph generated by the UMAP algorithm (see **Figure 7.2A**, middle and right). PDX data (all samples and conditions per PDX) shown as a contour plot of the Leiden clusters (lines in legend), with scatter plotting of the smallest clusters (dots in legend). Cluster number is based on the size of the cluster (largest cluster is number 0, smallest cluster is the highest number). Clusters shown here were used to determine proportion of cells in sample/condition occupying each cluster over time (**Figures 7.2B** and **7.2C**, bottom) and all analyses in **Figure 7.3**.

B LU73

C LX22

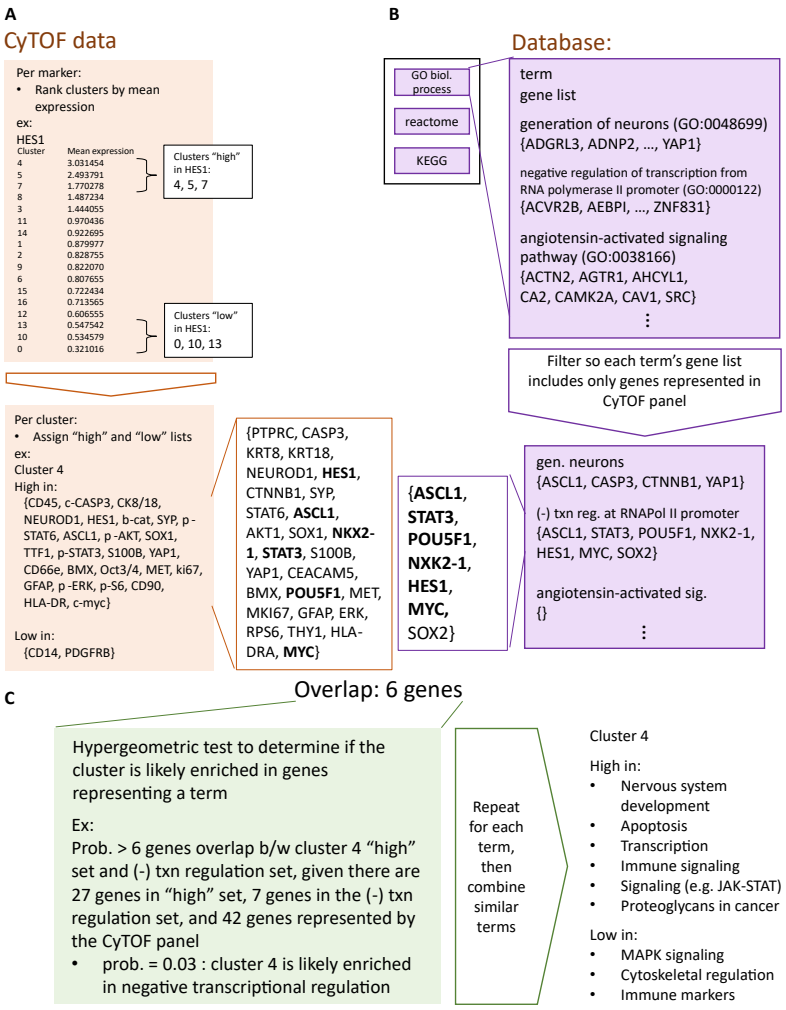


Figure D.2: Identifying phenotypes per cluster employs marker assignment and comparison to Gene Ontology (GO), Reactome pathways, and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases

A Assigning markers to clusters. Per marker, clusters were ranked by mean expression of that marker. The 3 highest-ranked clusters for a marker were denoted as "high" in that marker: for example, as pictured, the three clusters high in HES1 are clusters 4, 5, and 7. The 3 lowest-ranked clusters for a marker were denoted "low" in that marker: as pictured, low in HES1 are clusters 0, 10, and 13. This results in each cluster with a list of markers in which it is high and low, to be compared to database terms to investigate biological categories to be high or low in.

B Preparing databases. 3 databases were searched to match each cluster's lists of markers, with each cluster's "high" marker list searched separately from "low". Pictured is a subset of the GO biological process library (2018; library hosted by Enrichr [262, 263, 264]) arranged by GO term (underlined) and gene list. Each GO term list includes many genes, the set of potential genes to be found being all possible genes in the genome. (Lists from Reactome and KEGG terms may also include any gene in the genome.) Yet, the CyTOF panel includes only select proteins: only 42 potential genes to possibly find. Thus, each database library's gene lists were filtered to include only those 42 genes as the remainder could never be found in the CyTOF data.

C Comparing overlap between markers in each cluster list and database gene list to match biological process terms to clusters. The number of genes between each cluster list and each term in each database is used for a hypergeometric test, along with the number of all genes in the cluster list, number of all (filtered) genes in each term list, and all genes that could be found (42 genes). This calculates whether a cluster's list (whether high or low) is likely to be enriched in genes representing a biological process term, or if the overlap found is due to random chance.

D.2 Supplemental Tables

LU73 clusters	0	1	2	3	4	5	6	7	8	9	10	11	12	13
Apoptosis regulation	Hi						Lo			Lo	Hi	Lo	Hi	
Autophagy	Hi									Lo				
Axon guidance	Hi								Hi			Lo		
Cell adhesion					Lo			Hi	Hi	Lo		Lo		
Cell surface interactions														
Cytokine signaling							Lo							
Cytoskeleton regulation			Lo[2]							Lo				
Development	Hi					Hi		Hi			Hi[6]	Lo[8]		
ECM organization									Lo					
Immune signaling		Lo (INFγ,IL2)								Lo[3]				Hi
Immune system		Lo (adaptive)			Lo (adaptive)				Hi			Hi		
Metabolism										Lo[4]				
MicroRNAs in cancer										Lo				
Migration			Lo					Hi		Lo	Hi	Lo		
Neutrophils		Lo							Lo			Hi		
Proliferation						Hi							Hi	
Proteoglycans in cancer		Lo								Lo				
Signaling		Lo[1]		Hi		Hi (ΔNP63)	Lo		Hi	Lo[5]	Hi[7]	Lo (MAPK)	Hi[9]	
Stem cell pluripotency										Lo (Notch)			Hi	
Stem cell proliferation														
Transcription						Hi	Lo				Hi		Hi	

Table D.1: LU73 cluster phenotypic identity via GO terms

Clusters were distinguished as 'high' or 'low' in a marker if that cluster was in the top three or lowest three clusters for mean expression of that marker. Some clusters are high or low in several markers, and some in fewer. Here, Gene Ontology has been used for the combinations of markers assigned to each cluster. Any additional details from GO terms per cluster are either included in the table or here as a footnote if the details include a list.

1: FGF, TGF-β, Rac1/MMP2, VEGF, Rap1, BDNF

2: endoderm, endocrine, CNS, "branching morphogenesis of epithelial tube"

3: IL-2, 4, 5; TGF-beta; T cell signaling

4: "central carbon metabolism in cancer" (& metabolism general)

5: PKB, MAPK, PDGF, Jak-STAT, PI3k-Akt, Rap1, Hippo-Merlin, TSLP, VEGF, Wnt, Notch

6: Neuron differentiation, gland development, "branching morphogenesis of epithelial tube", cardiac progenitor, endoderm

7: BDNF, Wnt, thyroid hormone

8: gland development, cardiac progenitor, endoderm

9: TGF-β, Jak-STAT, Rac1/MMP2, Wnt

LX22 clusters	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Apoptosis regulation			Lo		Hi	Lo	Lo						Hi			Hi	
Autophagy													Hi				
Axon guidance							Lo		Hi					Lo			
Cell adhesion						Lo		Lo	Hi	Hi				Lo			
Cell surface interactions							Lo	Lo	Hi			Hi		Lo			
Cytokine signaling						Lo			Hi				Lo	Lo	Hi		
Cytoskeleton regulation					Lo		Lo	Lo	Hi								
Development			Lo		Hi (neural)		Lo	Hi (neural)						Lo^{10}			
ECM organization						Lo	Lo	Lo	Hi					Lo			
Immune signaling					Hi^{1}	Lo^{3}	Lo^{6}		Hi (cytokine)		Lo (IL-2)			Lo^{11}	Hi^{13}		Lo^{14}
Immune system			Lo (adaptive)		Lo			Lo		Hi	Lo			Lo^{12}			
Metabolism																	
MicroRNAs in cancer						Lo							Hi				
Migration							Lo										
Neutrophils									Hi								
Proliferation						Lo											
Proteoglycans in cancer					Hi						Lo		Hi				
Signaling	Lo (ERbb4)				Hi^{2}	Lo^{4}	Lo^{7}	Hi^{8}	Hi (MAPK)				Hi^{9}		Hi (PI3K)	Lo (JAK-STAT)	
Stem cell pluripotency						Lo						Lo					
Stem cell proliferation														Lo			
Transcription			Lo		Hi	Lo		Hi	Lo		Lo						

Table D.2: LX22 cluster phenotypic identity via GO terms

Clusters were distinguished as 'high' or 'low' in a marker if that cluster was in the top three or lowest three clusters for mean expression of that marker. Some clusters are high or low in several markers, and some in fewer. Here, Gene Ontology has been used for the combinations of markers assigned to each cluster. Any additional details from GO terms per cluster are included here as a footnote.

1: IL-2, 4; TGFβ
2: Jak-STAT, Rac1/MMP2 pathway, prolactin signaling, TSLP, VEGF
3: BCR; IL-2,5
4: NGF, BDNF, Hippo-Merlin, PI3K-Akt, ERBB4, PDGF, Jak-STAT, CXCR4, MAPK cascade
5: hematopoietic, cardiac progenitor
6: "interleukins", Fc epsilon receptor signaling
7: PKB, MAPK cascade, Rap1, DAP12, EGFR, FGFR, NGF
8: TSLP, VEGF
9: ERK1, ERK2 cascade; MAPK cascade; Hippo-Merlin
10: skin, gland, CNS, "branching morphogenesis of an epithelial tube"
11: Antigen processing and presentation, T cells
12: IL-2, 4; PD-1; TCR signaling
13: BCR, IL-5
14: T-cell receptor signaling, antigen receptor-mediated signaling pathway

References

[1] Matthias Galipaud, Mark A. F. Gillingham, Morgan David, and François-Xavier Dechaume-Moncharmont. Ecologists overestimate the importance of predictor variables in model averaging: a plea for cautious interpretations. *Methods in Ecology and Evolution*, 5:983–991, 10 2014.

[2] Rebecca L. Siegel, Kimberly D. Miller, Nikita Sandeep Wagle, and Ahmedin Jemal. Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73:17–48, 1 2023.

[3] P C Nowell. The clonal evolution of tumor cell populations. *Science (New York, N.Y.)*, 194:23–28, 10 1976.

[4] Mel Greaves and Carlo C. Maley. Clonal evolution in cancer. *Nature*, 481:306–313, 1 2012.

[5] Andriy Marusyk, Vanessa Almendro, and Kornelia Polyak. Intra-tumour heterogeneity: A looking glass for cancer? *Nature Reviews Cancer*, 12:323–334, 5 2012.

[6] Jean Pascal Capp. Cancer stem cells: From historical roots to a new perspective. *Journal of Oncology*, 2019, 6 2019.

[7] Charles M. Rudin, Elisabeth Brambilla, Corinne Faivre-Finn, and Julien Sage. Small-cell lung cancer. *Nature Reviews Disease Primers*, 7, 12 2021.

[8] Julie George, Jing Shan Lim, Se Jin Jang, Yupeng Cun, Luka Ozretić, Gu Kong, Frauke Leenders, Xin Lu, Lynnette Fernández-Cuesta, Graziella Bosco, Christian Müller, Ilona Dahmen, Nadine S. Jahchan, Kwon-Sik Park, Dian Yang, Anthony N. Karnezis, Dedeepya Vaka, Angela Torres, Maia Segura Wang, Jan O. Korbel, Roopika Menon, Sung-Min Chun, Deokhoon Kim, Matt Wilkerson, Neil Hayes, David Engelmann, Brigitte Pützer, Marc Bos, Sebastian Michels, Ignacija Vlasic, Danila Seidel, Berit Pinther, Philipp Schaub, Christian Becker, Janine Altmüller, Jun Yokota, Takashi Kohno, Reika Iwakawa, Koji Tsuta, Masayuki Noguchi, Thomas Muley, Hans Hoffmann, Philipp A. Schnabel, Iver Petersen, Yuan Chen, Alex Soltermann, Verena Tischler, Chang min Choi, Yong-Hee Kim, Pierre P. Massion, Yong Zou, Dragana Jovanovic, Milica Kontic, Gavin M. Wright, Prudence A. Russell, Benjamin Solomon, Ina Koch, Michael Lindner, Lucia A. Muscarella, Annamaria la Torre, John K. Field, Marko Jakopovic, Jelena Knezevic, Esmeralda Castaños-Vélez, Luca Roz, Ugo Pastorino, Odd-Terje Brustugun, Marius Lund-Iversen, Erik Thunnissen, Jens Köhler, Martin Schuler, Johan Botling, Martin Sandelin, Montserrat Sanchez-Cespedes, Helga B. Salvesen, Viktor Achter, Ulrich Lang, Magdalena Bogus, Peter M. Schneider, Thomas Zander, Sascha Ansén, Michael Hallek, Jürgen Wolf, Martin Vingron, Yasushi Yatabe, William D. Travis, Peter Nürnberg, Christian Reinhardt, Sven Perner, Lukas Heukamp, Reinhard Büttner, Stefan A. Haas, Elisabeth Brambilla, Martin Peifer, Julien Sage, and Roman K. Thomas. Comprehensive genomic profiles of small cell lung cancer. *Nature*, 524:47–53, 8 2015.

[9] Charles M. Rudin, John T. Poirier, Lauren Averett Byers, Caroline Dive, Afshin Dowlati, Julie George, John V. Heymach, Jane E. Johnson, Jonathan M. Lehman, David MacPherson, Pierre P. Massion, John D. Minna, Trudy G. Oliver, Vito Quaranta, Julien Sage, Roman K. Thomas, Christopher R. Vakoc, and Adi F. Gazdar. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nature Reviews Cancer*, 19:289–297, 3 2019.

[10] Jing Shan Lim, Alvaro Ibaseta, Marcus M. Fischer, Belinda Cancilla, Gilbert O'Young, Sandra Cristea, Vincent C. Luca, DIan Yang, Nadine S. Jahchan, Cécile Hamard, Martine Antoine, Marie Wislez, Christina Kong, Jennifer Cain, Yu Wang Liu, Ann M. Kapoun, K. Christopher Garcia, Timothy Hoey, Christopher L. Murriel, and Julien Sage. Intratumoural heterogeneity generated by notch signalling promotes small-cell lung cancer. *Nature*, 545:360–364, 2017.

[11] Abbie S. Ireland, Alexi M. Micinski, David W. Kastner, Bingqian Guo, Sarah J. Wait, Kyle B. Spainhower, Christopher C. Conley, Opal S. Chen, Matthew R. Guthrie, Danny Soltero, Yi Qiao,

Xiaomeng Huang, Szabolcs Tarapcsák, Siddhartha Devarakonda, Milind D. Chalishazar, Jason Gertz, Justin C. Moser, Gabor Marth, Sonam Puri, Benjamin L. Witt, Benjamin T. Spike, and Trudy G. Oliver. Myc drives temporal evolution of small cell lung cancer subtypes by reprogramming neuroendocrine fate. *Cancer Cell*, 38:60–78.e12, 2020.

[12] Jonathan M. Lehman, Mary E. Gwin, and Pierre P. Massion. Immunotherapy and targeted therapy for small cell lung cancer: There is hope. *Current Oncology Reports*, 19:1–9, 7 2017.

[13] Piyush B. Gupta, Christine M. Fillmore, Guozhi Jiang, Sagi D. Shapira, Kai Tao, Charlotte Kuperwasser, and Eric S. Lander. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146:633–644, 8 2011.

[14] Xiufang Chen, Yue Wang, Tianquan Feng, Ming Yi, Xingan Zhang, and Da Zhou. The overshoot and phenotypic equilibrium in characterizing cancer dynamics of reversible phenotypic plasticity. *Journal of Theoretical Biology*, 390:40–49, 2 2016.

[15] Angela Oliveira Pisco, Amy Brock, Joseph Zhou, Andreas Moor, Mitra Mojtahedi, Dean Jackson, and Sui Huang. Non-darwinian dynamics in therapy-induced cancer drug resistance. *Nature Communications*, 4:1–11, 9 2013.

[16] Alessandro L. Sellerio, Emilio Ciusani, Noa Bossel Ben-Moshe, Stefania Coco, Andrea Piccinini, Christopher R. Myers, James P. Sethna, Costanza Giampietro, Stefano Zapperi, and Caterina A.M. La Porta. Overshoot during phenotypic switching of cancer cell populations. *Scientific Reports 2015 5:1*, 5:1–14, 10 2015.

[17] Margaret P. Chapman, Tyler Risom, Anil J. Aswani, Ellen M. Langer, Rosalie C. Sears, and Claire J. Tomlin. Modeling differentiation-state transitions linked to therapeutic escape in triple-negative breast cancer. *PLOS Computational Biology*, 2019.

[18] Aaron Goldman, Biswanath Majumder, Andrew Dhawan, Sudharshan Ravi, David Goldman, Mohammad Kohandel, Pradip K. Majumder, and Shiladitya Sengupta. Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. *Nature Communications*, 6:1–13, 2 2015.

[19] Yapeng Su, Wei Wei, Lidia Robert, Min Xue, Jennifer Tsoi, Angel Garcia-Diaz, Blanca Homet Moreno, Jungwoo Kim, Rachel H. Ng, Jihoon W. Lee, Richard C. Koya, Begonya Comin-Anduix, Thomas G. Graeber, Antoni Ribas, and James R. Heath. Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. *Proceedings of the National Academy of Sciences of the United States of America*, 114:13679–13684, 12 2017.

[20] B Bishal Paudel, Leonard A Harris, Keisha N Hardeman, Arwa A Abugable, Corey E Hayford, Darren R Tyson, and Vito Quaranta. A nonquiescent "idling" population state in drug-treated, braf-mutated melanoma. *Biophysical journal*, 114:1499–1511, 3 2018.

[21] Mikahl Banwarth-Kuhn and Suzanne Sindi. How and why to build a mathematical model: A case study using prion aggregation. *Journal of Biological Chemistry*, 295:5022–5035, 4 2020.

[22] Michael P. H. Stumpf. Multi-model and network inference based on ensemble estimates: avoiding the madness of crowds. *Journal of the Royal Society Interface*, 17, 10 2020.

[23] Clare E. Giacomantonio and Geoffrey J. Goodhill. A boolean model of the gene regulatory network underlying mammalian cortical area development. *PLOS Computational Biology*, 6:e1000936, 2010.

[24] Kenneth P. & Burnham and David R. Anderson. *Model Selection and Multimodel Inference: a Practical Information-theoretic Approach*. 2nd ed. Springer, New York, 2002.

[25] Stéphanie Portet. A primer on model selection using the akaike information criterion. *Infectious Disease Modelling*, 5:111, 1 2020.

[26] Matthias Galipaud, Mark A.F. Gillingham, and François Xavier Dechaume-Moncharmont. A farewell to the sum of akaike weights: The benefits of alternative metrics for variable importance estimations in model selection. *Methods in Ecology and Evolution*, 8:1668–1678, 12 2017.

[27] Tiago M. Fragoso, Wesley Bertoli, and Francisco Louzada. Bayesian model averaging: A systematic review and conceptual classification. *International Statistical Review*, 86:1–28, 4 2018.

[28] John Skilling. Nested sampling. *AIP Conference Proceedings*, 735:395, 11 2004.

[29] F. Feroz and M. P. Hobson. Multimodal nested sampling: an efficient and robust alternative to markov chain monte carlo methods for astronomical data analyses. *Monthly Notices of the Royal Astronomical Society*, 384:449–463, 2 2008.

[30] F. Feroz, M. P. Hobson, and M. Bridges. Multinest: an efficient and robust bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, 398:1601–1614, 10 2009.

[31] Farhan Feroz, Michael P. Hobson, Ewan Cameron, and Anthony N. Pettitt. Importance nested sampling and the multinest algorithm. *The Open Journal of Astrophysics*, 2:11120, 11 2019.

[32] A O Pisco and S Huang. Non-genetic cancer cell plasticity and therapy-induced stemness in tumour relapse: 'what does not kill me strengthens me'. *British Journal of Cancer*, 112:1725–1732, 2015.

[33] C. H. Waddington. Canalization of development and the inheritance of acquired characters. *Nature*, 150:563–565, 11 1942.

[34] Sui Huang. Tumor progression: Chance and necessity in darwinian and lamarckian somatic (mutationless) evolution. *Progress in Biophysics and Molecular Biology*, 110:69–86, 9 2012.

[35] Amy Brock, Hannah Chang, and Sui Huang. Non-genetic heterogeneity–a mutation-independent driving force for the somatic evolution of tumours. *Nature Reviews Genetics*, 10:336–342, 5 2009.

[36] Kateri A Moore and Ihor R Lemischka. Stem cells and their niches. *Science*, 311:1880–1885, 2006.

[37] Michael F. Clarke, John E. Dick, Peter B. Dirks, Connie J. Eaves, Catriona H.M. Jamieson, D. Leanne Jones, Jane Visvader, Irving L. Weissman, and Geoffrey M. Wahl. Cancer stem cells - perspectives on current status and future directions: Aacr workshop on cancer stem cells. *Cancer Research*, 66:9339–9344, 10 2006.

[38] William H. Matsui. Cancer stem cell signaling pathways. *Medicine (United States)*, 95:S8–S19, 2016.

[39] Eduard Batlle and Hans Clevers. Cancer stem cells revisited. *Nature Medicine 2017 23:10*, 23:1124–1134, 10 2017.

[40] Keith S. Hoek, Ossia M. Eichhoff, Natalie C. Schlegel, Udo Döbbeling, Nikita Kobert, Leo Schaerer, Silvio Hemmi, and Reinhard Dummer. In vivo switching of human melanoma cells between proliferative and invasive states. *Cancer Research*, 68:650–656, 2 2008.

[41] Alexander Roesch, Mizuho Fukunaga-Kalabis, Elizabeth C. Schmidt, Susan E. Zabierowski, Patricia A. Brafford, Adina Vultur, Devraj Basu, Phyllis Gimotty, Thomas Vogt, and Meenhard Herlyn. A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell*, 141:583–594, 5 2010.

[42] Ryoichiro Kageyama, Toshiyuki Ohtsuka, Hiromi Shimojo, and Itaru Imayoshi. Dynamic notch signaling in neural progenitor cells and a revised view of lateral inhibition. *Nature Neuroscience 2008 11:11*, 11:1247–1251, 10 2008.

[43] Elsa Quintana, Mark Shackleton, Hannah R. Foster, Douglas R. Fullen, Michael S. Sabel, Timothy M. Johnson, and Sean J. Morrison. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell*, 18:510–523, 11 2010.

[44] Christine L. Chaffer, Ines Brueckmann, Christina Scheel, Alicia J. Kaestli, Paul A. Wiggins, Leonardo O. Rodrigues, Mary Brooks, Ferenc Reinhardt, Ying Suc, Kornelia Polyak, Lisa M. Arendt, Charlotte Kuperwasser, Brian Bierie, and Robert A. Weinberg. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Sciences of the United States of America*, 108:7950–7955, 5 2011.

[45] G. Yang, Y. Quan, W. Wang, Q. Fu, J. Wu, T. Mei, J. Li, Y. Tang, C. Luo, Q. Ouyang, S. Chen, L. Wu, T. K. Hei, and Y. Wang. Dynamic equilibrium between cancer stem cells and non-stem cancer cells in human sw620 and mcf-7 cancer cell populations. *British Journal of Cancer 2012 106:9*, 106:1512–1519, 4 2012.

[46] Sarah Schwitalla, Alexander A. Fingerle, Patrizia Cammareri, Tim Nebelsiek, Serkan I. Göktuna, Paul K. Ziegler, Ozge Canli, Jarom Heijmans, David J. Huels, Guenievre Moreaux, Rudolf A. Rupec, Markus Gerhard, Roland Schmid, Nick Barker, Hans Clevers, Roland Lang, Jens Neumann, Thomas Kirchner, Makoto M. Taketo, Gijs R. Van Den Brink, Owen J. Sansom, Melek C. Arkan, and Florian R. Greten. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell*, 152:25–38, 1 2013.

[47] Y. B. Hu, C. Yan, L. Mu, Y. L. Mi, H. Zhao, H. Hu, X. L. Li, D. D. Tao, Y. Q. Wu, J. P. Gong, and J. C. Qin. Exosomal wnt-induced dedifferentiation of colorectal cancer cells contributes to chemotherapy resistance. *Oncogene 2018 38:11*, 38:1951–1965, 11 2018.

[48] Ana C. DeCarvalho, Kevin Nelson, Nancy Lemke, Norman L. Lehman, Ali S. Arbab, Steven Kalkanis, and Tom Mikkelsen. Gliosarcoma stem cells undergo glial and mesenchymal differentiation in vivo. *Stem Cells*, 28:181–190, 2 2010.

[49] Artem D. Berezovsky, Laila M. Poisson, David Cherba, Craig P. Webb, Andrea D. Transou, Nancy W. Lemke, Xin Hong, Laura A. Hasselbach, Susan M. Irtenkauf, Tom Mikkelsen, and Ana C. de Carvalho. Sox2 promotes malignancy in glioblastoma by regulating plasticity and astrocytic differentiation. *Neoplasia*, 16:193–206, 3 2014.

[50] Arthur D. Lander, Kimberly K. Gokoffski, Frederic Y.M. Wan, Qing Nie, and Anne L. Calof. Cell lineages and the logic of proliferative control. *PLOS Biology*, 7:e1000015, 1 2009.

[51] Rong Wang, Kalyani Chadalavada, Jennifer Wilshire, Urszula Kowalik, Koos E. Hovinga, Adam Geber, Boris Fligelman, Margaret Leversha, Cameron Brennan, and Viviane Tabar. Glioblastoma stem-like cells give rise to tumour endothelium. *Nature*, 468:829–833, 11 2010.

[52] Lucia Ricci-Vitiani, Roberto Pallini, Mauro Biffoni, Matilde Todaro, Gloria Invernici, Tonia Cenci, Giulio Maira, Eugenio Agostino Parati, Giorgio Stassi, Luigi Maria Larocca, and Ruggero De Maria. Tumour vascularization via endothelial differentiation of glioblastoma stem-like cells. *Nature*, 468:824–828, 11 2010.

[53] Nadine S. Jahchan, Jing Shan Lim, Becky Bola, Karen Morris, Garrett Seitz, Kim Q. Tran, Lei Xu, Francesca Trapani, Christopher J. Morrow, Sandra Cristea, Garry L. Coles, Dian Yang, Dedeepya Vaka, Michael S. Kareta, Julie George, Pawel K. Mazur, Thuyen Nguyen, Wade C. Anderson, Scott J. Dylla, Fiona Blackhall, Martin Peifer, Caroline Dive, and Julien Sage. Identification and targeting of long-term tumor-propagating cells in small cell lung cancer. *Cell Reports*, 16:644–656, 2016.

[54] Ga Young Lee, Jae Seung Shim, Bin Cho, Joo Young Jung, Dong Soon Lee, and Il Hoan Oh. Stochastic acquisition of a stem cell-like state and drug tolerance in leukemia cells stressed by radiation. *International Journal of Hematology*, 93:27–35, 1 2011.

[55] Laura Ghisolfi, Andrew C. Keates, Xingwang Hu, Dong ki Lee, and Chiang J. Li. Ionizing radiation induces stemness in cancer cells. *PLOS ONE*, 7:e43628, 8 2012.

[56] Anne Schroeder, Andreas Herrmann, Gregory Cherryholmes, Claudia Kowolik, Ralf Buettner, Sumanta Pal, Hua Yu, Gerhard Müller-Newen, and Richard Jove. Loss of androgen receptor expression promotes a stem-like cell phenotype in prostate cancer through stat3 signaling. *Cancer Research*, 74:1227–1237, 2 2014.

[57] Brian B. Liau, Cem Sievers, Laura K. Donohue, Shawn M. Gillespie, William A. Flavahan, Tyler E. Miller, Andrew S. Venteicher, Christine H. Hebert, Christopher D. Carey, Scott J. Rodig, Sarah J. Shareef, Fadi J. Najm, Peter van Galen, Hiroaki Wakimoto, Daniel P. Cahill, Jeremy N. Rich, Jon C. Aster, Mario L. Suvà, Anoop P. Patel, and Bradley E. Bernstein. Adaptive chromatin remodeling drives glioblastoma stem cell plasticity and drug tolerance. *Cell Stem Cell*, 20:233–246, 2 2017.

[58] Barbara Banelli, Elisa Carra, Federica Barbieri, Roberto Würth, Federica Parodi, Alessandra Pattarozzi, Roberta Carosio, Alessandra Forlani, Giorgio Allemanni, Daniela Marubbi, Tullio Florio, Antonio Daga, and Massimo Romani. The histone demethylase kdm5a is a key factor for the resistance to temozolomide in glioblastoma. *Cell Cycle*, 14:3418–3429, 2015.

[59] Pan Yu Chen, Mandar Deepak Muzumdar, Kimberly Judith Dorans, Rebecca Robbins, Arjun Bhutkar, Amanda Del Rosario, Philipp Mertins, Jana Qiao, Anette Claudia Schafer, Frank Gertler, Steven Carr, and Tyler Jacks. Adaptive and reversible resistance to kras inhibition in pancreatic cancer cells. *Cancer Research*, 78:985–1002, 2 2018.

[60] Tyler Risom, Ellen M. Langer, Margaret P. Chapman, Juha Rantala, Andrew J. Fields, Christopher Boniface, Mariano J. Alvarez, Nicholas D. Kendsersky, Carl R. Pelz, Katherine Johnson-Camacho, Lacey E. Dobrolecki, Koei Chin, Anil J. Aswani, Nicholas J. Wang, Andrea Califano, Michael T. Lewis, Claire J. Tomlin, Paul T. Spellman, Andrew Adey, Joe W. Gray, and Rosalie C. Sears. Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. *Nature Communications*, 9:1–17, 9 2018.

[61] Elizabeth A. Kuczynski, Daniel J. Sargent, Axel Grothey, and Robert S. Kerbel. Drug rechallenge and treatment beyond progression—implications for drug resistance. *Nature Reviews Clinical Oncology*, 10:571–587, 9 2013.

[62] Sreenath V. Sharma, Diana Y. Lee, Bihua Li, Margaret P. Quinlan, Fumiyuki Takahashi, Shyamala Maheswaran, Ultan McDermott, Nancy Azizian, Lee Zou, Michael A. Fischbach, Kwok Kin Wong, Kathleyn Brandstetter, Ben Wittner, Sridhar Ramaswamy, Marie Classon, and Jeff Settleman. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141:69–80, 4 2010.

[63] Michael Ramirez, Satwik Rajaram, Robert J. Steininger, Daria Osipchuk, Maike A. Roth, Leanna S. Morinishi, Louise Evans, Weiyue Ji, Chien Hsiang Hsu, Kevin Thurley, Shuguang Wei, Anwu Zhou, Prasad R. Koduru, Bruce A. Posner, Lani F. Wu, and Steven J. Altschuler. Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. *Nature Communications*, 7:1–8, 2 2016.

[64] Aaron N Hata, Matthew J Niederst, Hannah L Archibald, Maria Gomez-Caraballo, Faria M Siddiqui, Hillary E Mulvey, Yosef E Maruvka, Fei Ji, Hyo-Eun C Bhang, Viveksagar Krishnamurthy Radhakrishna, Giulia Siravegna, Haichuan Hu, Sana Raoof, Elizabeth Lockerman, Anuj Kalsy, Dana Lee, Celina L Keating, David A Ruddy, Leah J Damon, Adam S Crystal, Carlotta Costa, Zofia Piotrowska, Alberto Bardelli, Anthony J Iafrate, Ruslan I Sadreyev, Frank Stegmeier, Gad Getz, Lecia V Sequist, Anthony C Faber, and Jeffrey A Engelman. Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nature Medicine*, 22:262–269, 2 2016.

[65] Sydney M Shaffer, Margaret C Dunagin, Stefan R Torborg, Eduardo A Torre, Benjamin Emert, Clemens Krepler, Marilda Beqiri, Katrin Sproesser, Patricia A Brafford, Min Xiao, Elliott Eggan, Ioannis N Anastopoulos, Cesar A Vargas-Garcia, Abhyudai Singh, Katherine L Nathanson, Meenhard Herlyn, and Arjun Raj. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature*, 546:431–435, 2017.

[66] Ankur Sharma, Elaine Yiqun Cao, Vibhor Kumar, Xiaoqian Zhang, Hui Sun Leong, Angeline Mei Lin Wong, Neeraja Ramakrishnan, Muhammad Hakimullah, Hui Min Vivian Teo, Fui Teen Chong, Shumei Chia, Matan Thangavelu Thangavelu, Xue Lin Kwang, Ruta Gupta, Jonathan R. Clark, Giridharan

Periyasamy, N. Gopalakrishna Iyer, and Ramanuj DasGupta. Longitudinal single-cell rna sequencing of patient-derived primary cells reveals drug-induced infidelity in stem cell hierarchy. *Nature Communications*, 2018.

[67] Min Zou, Roxanne Toivanen, Antonina Mitrofanova, Nicolas Floch, Sheida Hayati, Yanping Sun, Clémentine Le Magnen, Daniel Chester, Elahe A. Mostaghel, Andrea Califano, Mark A. Rubin, Michael M. Shen, and Cory Abate-Shen. Transdifferentiation as a mechanism of treatment resistance in a mouse model of castration-resistant prostate cancer. *Cancer Discovery*, 2017.

[68] Ping Mu, Zeda Zhang, Matteo Benelli, Wouter R. Karthaus, Elizabeth Hoover, Chi-Chao Chen, John Wongvipat, Sheng-Yu Ku, Dong Gao, Zhen Cao, Neel Shah, Elizabeth J. Adams, Wassim Abida, Philip A. Watson, Davide Prandi, Chun-Hao Huang, Elisa de Stanchina, Scott W. Lowe, Leigh Ellis, Himisha Beltran, Mark A. Rubin, David W. Goodrich, Francesca Demichelis, and Charles L. Sawyers. ¡i¿sox2¡/i¿ promotes lineage plasticity and antiandrogen resistance in ¡i¿tp53¡/i¿ - and ¡i¿rb1¡/i¿ - deficient prostate cancer. *Science*, 2017.

[69] Sheng Yu Ku, Spencer Rosario, Yanqing Wang, Ping Mu, Mukund Seshadri, Zachary W. Goodrich, Maxwell M. Goodrich, David P. Labbé, Eduardo Cortes Gomez, Jianmin Wang, Henry W. Long, Bo Xu, Myles Brown, Massimo Loda, Charles L. Sawyers, Leigh Ellis, and David W. Goodrich. Rb1 and trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science*, 355:78–83, 1 2017.

[70] Mannan Nouri, Josselin Caradec, Amy Anne Lubik, Na Li, Brett G. Hollier, Mandeep Takhar, Manuel Altimirano-Dimas, Mengqian Chen, Mani Roshan-Moniri, Miriam Butler, Melanie Lehman, Jennifer Bishop, Sarah Truong, Shih-Chieh Huang, Dawn Cochrane, Michael Cox, Colin Collins, Martin Gleave, Nicholas Erho, Mohamed Alshalafa, Elai Davicioni, Colleen Nelson, Sheryl Gregory-Evans, R. Jeffrey Karnes, Robert B. Jenkins, Eric A. Klein, Ralph Buttyan, Mannan Nouri, Josselin Caradec, Amy Anne Lubik, Na Li, Brett G. Hollier, Mandeep Takhar, Manuel Altimirano-Dimas, Mengqian Chen, Mani Roshan-Moniri, Miriam Butler, Melanie Lehman, Jennifer Bishop, Sarah Truong, Shih-Chieh Huang, Dawn Cochrane, Michael Cox, Colin Collins, Martin Gleave, Nicholas Erho, Mohamed Alshalafa, Elai Davicioni, Colleen Nelson, Sheryl Gregory-Evans, R. Jeffrey Karnes, Robert B. Jenkins, Eric A. Klein, and Ralph Buttyan. Therapy-induced developmental reprogramming of prostate cancer cells and acquired therapy resistance. *Oncotarget*, 8:18949–18967, 1 2017.

[71] Daniel A. Charlebois and Gábor Balázsi. Modeling cell population dynamics. *In Silico Biology*, 13:21–39, 12 2018.

[72] István Albert, Juilee Thakar, Song Li, Ranran Zhang, and Réka Albert. Boolean network simulations for life scientists. *Source Code for Biology and Medicine*, 3:1–8, 11 2008.

[73] John Metzcar, Yafei Wang, Randy Heiland, and Paul Macklin. A review of cell-based computational modeling in cancer biology. *JCO Clin. Cancer Inform.*, 3:1–13, 2 2019.

[74] Thomas S. Deisboeck, Zhihui Wang, Paul Macklin, and Vittorio Cristini. Multiscale cancer modeling. *Annual Review of Biomedical Engineering*, 13:127–155, 8 2011.

[75] Weikang Wang, Yi Quan, Qibin Fu, Yu Liu, Ying Liang, Jingwen Wu, Gen Yang, Chunxiong Luo, Qi Ouyang, and Yugang Wang. Dynamics between cancer cell subpopulations reveals a model coordinating with both hierarchical and stochastic concepts. *PloS one*, 9:e84654, 1 2014.

[76] Da Zhou, Shanjun Mao, Jing Cheng, Kaiyi Chen, Xiaofang Cao, and Jie Hu. A bayesian statistical analysis of stochastic phenotypic plasticity model of cancer cells. *Journal of Theoretical Biology*, 2018.

[77] Andriy Marusyk and Kornelia Polyak. Tumor heterogeneity: Causes and consequences. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1805:105–117, 1 2010.

[78] Sui Huang. The molecular and mathematical basis of waddington's epigenetic landscape: a framework for post-darwinian biology? *BioEssays*, 34:149–157, 2 2012.

[79] Sui Huang. Genetic and non-genetic instability in tumor progression: link between the fitness landscape and the epigenetic landscape of cancer cells. *Cancer and Metastasis Reviews*, 32:423–448, 12 2013.

[80] Arjun Raj and Alexander van Oudenaarden. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell*, 135:216–226, 10 2008.

[81] Lazaros Batsilas, Alexander M. Berezhkovskii, and Stanislav Y. Shvartsman. Stochastic model of autocrine and paracrine signals in cell culture assays. *Biophysical Journal*, 85:3659–3665, 12 2003.

[82] Sabrina L Spencer, Suzanne Gaudet, John G Albeck, John M Burke, and Peter K Sorger. Non-genetic origins of cell-to-cell variability in trail-induced apoptosis. *Nature*, 459:428–432, 5 2009.

[83] Dimitrios Anastasiou. Tumour microenvironment factors shaping the cancer metabolism landscape. *British Journal of Cancer*, 116:277–286, 1 2017.

[84] Andrea Sottoriva, Haeyoun Kang, Zhicheng Ma, Trevor A Graham, Matthew P Salomon, Junsong Zhao, Paul Marjoram, Kimberly Siegmund, Michael F Press, Darryl Shibata, and Christina Curtis. A big bang model of human colorectal tumor growth. *Nature Genetics*, 47:209–216, 2 2015.

[85] Corbin E Meacham and Sean J Morrison. Tumour heterogeneity and cancer cell plasticity. *Nature*, 501:328–337, 9 2013.

[86] Henrica M. J. Werner, Gordon B. Mills, and Prahlad T. Ram. Cancer systems biology: a peek into the future of patient care? *Nature Reviews Clinical Oncology*, 11:167–176, 3 2014.

[87] Alexander R. A. Anderson and Vito Quaranta. Integrative mathematical oncology. *Nature Reviews Cancer*, 8:227–234, 3 2008.

[88] Gary An, Qi Mi, Joyeeta Dutta-Moscato, and Yoram Vodovotz. Agent-based models in translational systems biology. *WIRES Syst. Biol. Med.*, 1:159–171, 9 2009.

[89] Jan Poleszczuk, Paul Macklin, and Heiko Enderling. Agent-based modeling of cancer stem cell driven solid tumor growth. *Methods in Molecular Biology*, 1516:335–346, 2016.

[90] Heiko Enderling and Philip Hahnfeldt. Cancer stem cells in solid tumors: Is 'evading apoptosis' a hallmark of cancer? *Progress in Biophysics and Molecular Biology*, 106:391–399, 2011.

[91] Kerri-Ann Norton, Travis Wallace, Niranjan B. Pandey, and Aleksander S. Popel. An agent-based model of triple-negative breast cancer: the interplay between chemokine receptor ccr5 expression, cancer stem cells, and hypoxia. *BMC Systems Biology*, 11:68, 12 2017.

[92] Jan Poleszczuk and Heiko Enderling. Cancer stem cell plasticity as tumor growth promoter and catalyst of population collapse. *Stem Cells International*, 2016:1–12, 12 2016.

[93] Jakob Nikolas Kather, Jan Poleszczuk, Meggy Suarez-Carmona, Johannes Krisam, Pornpimol Charoentong, Nektarios A. Valous, Cleo Aron Weis, Luca Tavernar, Florian Leiss, Esther Herpel, Fee Klupp, Alexis Ulrich, Martin Schneider, Alexander Marx, Dirk Jäger, and Niels Halama. In silico modeling of immunotherapy and stroma-targeting therapies in human colorectal cancer. *Cancer Research*, 77:6442–6452, 2017.

[94] S.H. Sabzpoushan and Fateme Pourhasanzade. A new method for shrinking tumor based on microenvironmental factors: Introducing a stochastic agent-based model of avascular tumor growth. *Physica A*, 508:771–787, 10 2018.

[95] Damian Stichel, Alistair M. Middleton, Benedikt F. Müller, Sofia Depner, Ursula Klingmüller, Kai Breuhahn, and Franziska Matthäus. An individual-based model for collective cancer cell migration explains speed dynamics and phenotype variability in response to growth factors. *npj Systems Biology and Applications*, 2017.

[96] Stefan Hoehme, Francois Bertaux, William Weens, Bettina Grasl-Kraupp, Jan G. Hengstler, and Dirk Drasdo. Model prediction and validation of an order mechanism controlling the spatiotemporal phenotype of early hepatocellular carcinoma. *Bulletin of Mathematical Biology*, 80:1134–1171, 2018.

[97] Paul Van Liedekerke, Johannes Neitsch, Tim Johann, Kevin Alessandri, Pierre Nassoy, and Dirk Drasdo. Quantitative agent-based modeling reveals mechanical stress response of growing tumor spheroids is predictable over various growth conditions and cell lines. *PLOS Computational Biology*, 2019.

[98] Ignacio A Rodriguez-Brenes and Dominik Wodarz. Preventing clonal evolutionary processes in cancer: Insights from mathematical models. *Proceedings of the National Academy of Sciences of the United States of America*, 112:8843–8850, 7 2015.

[99] Margriet M. Palm, Marjet Elemans, and Joost B. Beltman. Heritable tumor cell division rate heterogeneity induces clonal dominance. *PLoS Computational Biology*, 2018.

[100] Marco S Nobile, Thalia Vlachou, Simone Spolaor, Daniela Bossi, Paolo Cazzaniga, Luisa Lanfrancone, Giancarlo Mauri, Pier Giuseppe Pelicci, and Daniela Besozzi. Modeling cell proliferation in human acute myeloid leukemia xenografts. *Bioinformatics*, 2 2019.

[101] N. Minar, R. Burkhart, C. Langton, and M. Askenazi. The swarm simulation system: a toolkit for building multi-agent simulations. Working Paper 96-06-042, Santa Fe Institute, 1996.

[102] Seth Tisue and Uri Wilensky. NetLogo: A simple environment for modeling complexity. In Ali Minai and Yaneer Bar-Yam, editors, *Proceedings of the Fifth International Conference on Complex Systems ICCS 2004*, pages 16–21, 2004.

[103] Alexey Solovyev, Maxim Mikheev, Leming Zhou, Joyeeta Dutta-Moscato, Cordelia Ziraldo, Gary An, Yoram Vodovotz, and Qi Mi. Spark. *International Journal of Agent Technologies and Systems*, 2:18–30, 7 2010.

[104] Nicholson Collier and Michael North. Parallel agent-based simulation with repast for high performance computing. *SIMULATION*, 89:1215–1235, 10 2013.

[105] F. Kargi. Re-interpretation of the logistic equation for batch microbial growth in relation to monod kinetics. *Letters in Applied Microbiology*, 48:398–401, 4 2009.

[106] Franck. Courchamp, Ludek. Berec, and Joanna. Gascoigne. *Allee effects in ecology and conservation*. Oxford University Press, 2008.

[107] María Fuentes-Garí, Ruth Misener, David García-Munzer, Eirini Velliou, Michael C. Georgiadis, Margaritis Kostoglou, Efstratios N. Pistikopoulos, Nicki Panoskaltsis, and Athanasios Mantalaris. A mathematical model of subpopulation kinetics for the deconvolution of leukaemia heterogeneity. *Journal of The Royal Society Interface*, 12, 7 2015.

[108] James M. Greene, Doron Levy, King Leung Fung, Paloma S. Souza, Michael M. Gottesman, and Orit Lavi. Modeling intrinsic heterogeneity and growth of cancer cells. *Journal of Theoretical Biology*, 367:262–277, 2 2015.

[109] Stephen Smith and Ramon Grima. Spatial stochastic intracellular kinetics: a review of modelling approaches. *Bulletin of Mathematical Biology*, 81:2960–3009, 5 2018.

[110] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58:35–55, 2007.

[111] S. Hoops, S. Sahle, R. Gauges, C. Lee, J. Pahle, N. Simus, M. Singhal, L. Xu, P. Mendes, and U. Kummer. Copasi–a complex pathway simulator. *Bioinformatics*, 22:3067–3074, 12 2006.

[112] Diana C. Resasco, Fei Gao, Frank Morgan, Igor L. Novak, James C. Schaff, and Boris M. Slepchenko. Virtual cell: computational tools for modeling in cell biology. *WIRES Syst. Biol. Med.*, 4:129–140, 3 2012.

[113] Kevin R Sanft, Sheng Wu, Min Roh, Jin Fu, Rone Kwei Lim, and Linda R Petzold. Stochkit2: software for discrete stochastic simulation of biochemical systems with events. *Bioinformatics*, 27:2457–2458, 2011.

[114] D. Fange, A. Mahmutovic, and J. Elf. Mesord 1.0: Stochastic reaction-diffusion simulations in the microscopic limit. *Bioinformatics*, 28:3155–3157, 12 2012.

[115] Brian Drawert, Stefan Engblom, and Andreas Hellander. Urdme: a modular framework for stochastic simulation of reaction-transport processes in complex geometries. *BMC systems biology*, 6:76, 6 2012.

[116] Carlos F Lopez, Jeremy L Muhlich, John A Bachman, and Peter K Sorger. Programming biological models in python using pysb. *Molecular Systems Biology*, 9:646, 2 2013.

[117] L.A. Harris, J.S. Hogg, J.-J. Tapia, J.A.P. Sekar, S. Gupta, I. Korsunsky, A. Arora, D. Barua, R.P. Sheehan, and J.R. Faeder. Bionetgen 2.2: Advances in rule-based modeling. *Bioinformatics*, 32:3366–3368, 2016.

[118] Doris P. Tabassum and Kornelia Polyak. Tumorigenesis: It takes a village. *Nature Reviews Cancer*, 15:473–483, 8 2015.

[119] Joseph Xu Zhou, Angela Oliveira Pisco, Hong Qian, and Sui Huang. Nonequilibrium population dynamics of phenotype conversion of cancer cells. *PLOS ONE*, 9:e110714, 12 2014.

[120] Elena Piretto, Marcello Delitala, and Mario Ferraro. Combination therapies and intra-tumoral competition: Insights from mathematical modeling. *Journal of Theoretical Biology*, 446:149–159, 6 2018.

[121] Diego Chowell, James Napier, Rohan Gupta, Karen S. Anderson, Carlo C. Maley, and Melissa A. Wilson Sayres. Modeling the subclonal evolution of cancer cell populations. *Cancer Research*, 78:830–839, 2018.

[122] Kamran Kaveh, Mohammad Kohandel, and Siv Sivaloganathan. Replicator dynamics of cancer stem cell: Selection in the presence of differentiation and plasticity. *Mathematical Biosciences*, 272:64–75, 2016.

[123] Hualiang Zhong, Stephen Brown, Suneetha Devpura, X. Allen Li, and Indrin J. Chetty. Kinetic modeling of tumor regression incorporating the concept of cancer stem-like cells for patients with locally advanced lung cancer. *Theoretical Biology and Medical Modelling*, 15:23, 12 2018.

[124] Kirill S. Korolev, Joao B. Xavier, and Jeff Gore. Turning ecology and evolution against cancer. *Nature Reviews Cancer*, 14:371–380, 2014.

[125] Anna Konstorum, Thomas Hillen, and John Lowengrub. Feedback regulation in a cancer stem cell model can cause an allee effect. *Bulletin of Mathematical Biology*, 78:754–785, 2016.

[126] Joel S. Brown, Jessica J. Cunningham, and Robert A. Gatenby. Aggregation effects and population-based dynamics as a source of therapy resistance in cancer. *IEEE Transactions on Biomedical Engineering*, 64:512–518, 2017.

[127] Zoltan Neufeld, William von Witt, Dora Lakatos, Jiaming Wang, Balazs Hegedus, and Andras Czirok. The role of allee effect in modelling post resection recurrence of glioblastoma. *PLoS Computational Biology*, 13:e1005818, 2017.

[128] Daniel C Kirouac, Caryn Ito, Elizabeth Csaszar, Aline Roch, Mei Yu, Edward A Sykes, Gary D Bader, and Peter W Zandstra. Dynamic interaction networks in a hierarchically organized tissue. *Molecular Systems Biology*, 6:417, 10 2010.

[129] Mingyang Lu, Bin Huang, Samir M Hanash, José N Onuchic, and Eshel Ben-Jacob. Modeling putative therapeutic implications of exosome exchange between tumor and immune cells. *Proceedings of the National Academy of Sciences of the United States of America*, 111:E4165–E4174, 10 2014.

[130] Mary E. Sehl, Miki Shimada, Alfonso Landeros, Kenneth Lange, and Max S. Wicha. Modeling of cancer stem cell state transitions predicts therapeutic response. *PLOS ONE*, 10:e0135797, 9 2015.

[131] Xiaoqiang Sun, Jiguang Bao, and Yongzhao Shao. Mathematical modeling of therapy-induced cancer drug resistance: connecting cancer mechanisms to population survival rates. *Scientific Reports*, 6:22498, 2016.

[132] Martina Baar, Loren Coquille, Hannah Mayer, Michael Hölzel, Meri Rogava, Thomas Tüting, and Anton Bovier. A stochastic model for immunotherapy of cancer. *Scientific Reports*, 6:24169, 7 2016.

[133] Martin Meier-Schellersheim, Iain D. C. Fraser, and Frederick Klauschen. Multiscale modeling for biologists. *WIRES Syst. Biol. Med.*, 1:4–14, 7 2010.

[134] Joseph O. Dada and Pedro Mendes. Multi-scale modelling and simulation in systems biology. *Integrative Biology*, 3:86–96, 2 2011.

[135] Joseph Walpole, Jason A. Papin, and Shayn M. Peirce. Multiscale computational models of complex biological systems. *Annual Review of Biomedical Engineering*, 15:137–154, 7 2013.

[136] Jessica S Yu and Neda Bagheri. Multi-class and multi-scale models of complex biological phenomena. *Current Opinion in Biotechnology*, 39:167–173, 2016.

[137] Alexander R.A. Anderson, Alissa M. Weaver, Peter T. Cummings, and Vito Quaranta. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell*, 127:905–915, 2006.

[138] Alexander R.A. Anderson, Mohamed Hassanein, Kevin M. Branch, Jenny Lu, Nichole A. Lobdell, Julie Maier, David Basanta, Brandy Weidow, Archana Narasanna, Carlos L. Arteaga, Albertb Reynolds, Vito Quaranta, Lourdes Estrada, and Alissa M. Weaver. Microenvironmental independence associated with tumor progression. *Cancer Research*, 69:8797–8806, 2009.

[139] Minna Qiao, Dan Wu, Michelle Carey, Xiaobo Zhou, and Le Zhang. Multi-scale agent-based multiple myeloma cancer modeling and the related study of the balance between osteoclasts and osteoblasts. *PLoS ONE*, 10:e0143206, 2015.

[140] Arig Ibrahim-Hashim, Mark Robertson-Tessi, Pedro M. Enriquez-Navas, Mehdi Damaghi, Yoganand Balagurunathan, Jonathan W. Wojtkowiak, Shonagh Russell, Kam Yoonseok, Mark C. Lloyd, Marilyn M. Bui, Joel S. Brown, Alexander R.A. Anderson, Robert J. Gillies, and Robert A. Gatenby. Defining cancer subpopulations by adaptive strategies rather than molecular properties provides novel insights into intratumoral evolution. *Cancer Research*, 77:2242–2254, 2017.

[141] Mehrdad Ghadiri, Mahshid Heidari, Sayed Amir Marashi, and Seyed Hasan Mousavi. A multiscale agent-based framework integrated with a constraint-based metabolic network model of cancer for simulating avascular tumor growth. *Molecular BioSystems*, 13:1888–1897, 2017.

[142] Nick Jagiella, Benedikt Müller, Margareta Müller, Irene E. Vignon-Clementel, and Dirk Drasdo. Inferring growth control mechanisms in growing multi-cellular spheroids of nsclc cells from spatial-temporal image data. *PLoS Computational Biology*, 12:e1004412, 2016.

[143] Joseph D. Butner, Vittorio Cristini, and Zhihui Wang. Development of a three dimensional, multiscale agent-based model of ductal carcinoma in situ. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 86–89, 2017.

[144] Eunjung Kim, Jae Young Kim, Matthew A. Smith, Eric B. Haura, and Alexander R.A. Anderson. Cell signaling heterogeneity is modulated by both cell-intrinsic and -extrinsic mechanisms: An integrated approach to understanding targeted therapy. *PLoS Biology*, 16:e2002930, 2018.

[145] A. Hoekstra, B. Chopard, and P. Coveney. Multiscale modelling and simulation: a position paper. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372:20130377, 6 2014.

[146] Gaelle Letort, Arnau Montagud, Gautier Stoll, Randy Heiland, Emmanuel Barillot, Paul MacKlin, Andrei Zinovyev, and Laurence Calzone. Physiboss: A multi-scale agent-based modelling framework integrating physical dimension and cell signalling. *Bioinformatics*, 35:1188–1196, 2019.

[147] Maciej H. Swat, Gilberto L. Thomas, Julio M. Belmonte, Abbas Shirinifard, Dimitrij Hmeljak, and James A. Glazier. Multi-scale modeling of tissues using compucell3d. *Methods in Cell Biology*, 110:325–366, 1 2012.

[148] Ahmadreza Ghaffarizadeh, Randy Heiland, Samuel H. Friedman, Shannon M. Mumenthaler, and Paul Macklin. Physicell: an open source physics-based cell simulator for 3-d multicellular systems. *PLOS Computational Biology*, 14:e1005991, 2 2018.

[149] Gary R. Mirams, Christopher J. Arthurs, Miguel O. Bernabeu, Rafel Bordas, Jonathan Cooper, Alberto Corrias, Yohan Davit, Sara-Jane Dunn, Alexander G. Fletcher, Daniel G. Harvey, Megan E. Marsh, James M. Osborne, Pras Pathmanathan, Joe Pitt-Francis, James Southern, Nejib Zemzemi, and David J. Gavaghan. Chaste: an open source c++ library for computational physiology and biology. *PLoS Computational Biology*, 9:e1002970, 3 2013.

[150] Rafael Bravo, Etienne Baratchart, Jeffrey West, Ryan O. Schenck, Anna K. Miller, Jill Gallaher, Chandler D. Gatenbee, David Basanta, Mark Robertson-Tessi, and Alexander R. A. Anderson. Hybrid automata library: a modular platform for efficient hybrid modeling with real-time visualization. *bioRxiv*, page 411538, 12 2018.

[151] Jörn Starruß, Walter de Back, Lutz Brusch, and Andreas Deutsch. Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology. *Bioinformatics*, 30:1331–1332, 5 2014.

[152] Huaming Yan, Monica Romero-Lopez, H.B. Frieboes, C.C.W. Hughes, and J.S. Lowengrub. Multiscale modeling of glioblastoma suggests that the partial disruption of vessel/cancer stem cell crosstalk can promote tumor regression without increasing invasiveness. *IEEE Transactions on Biomedical Engineering*, 64:538–548, 2016.

[153] Eva Maria Galan-Moya, Armelle Le Guelte, Evelyne Lima Fernandes, Cécile Thirant, Julie Dwyer, Nicolas Bidere, Pierre-Olivier Couraud, Mark G H Scott, Marie-Pierre Junier, Hervé Chneiweiss, and Julie Gavard. Secreted factors from brain endothelial cells maintain glioblastoma stem-like cell expansion through the mtor pathway. *EMBO reports*, 12:470–476, 5 2011.

[154] Guang-Ning Yan, Lang Yang, Yang-Fan Lv, Yu Shi, Li-Li Shen, Xiao-Hong Yao, Qiao-Nan Guo, Peng Zhang, You-Hong Cui, Xia Zhang, Xiu-Wu Bian, and De-Yu Guo. Endothelial cells promote stem-like phenotype of glioma cells through activating the hedgehog pathway. *The Journal of Pathology*, 234:11–22, 9 2014.

[155] Adi F. Gazdar, Paul A. Bunn, and John D. Minna. Small-cell lung cancer: What we know, what we need to know and the path forward. *Nature Reviews Cancer*, 17:725–737, 2017.

[156] Roy S. Herbst, Giuseppe Giaccone, Filippo de Marinis, Niels Reinmuth, Alain Vergnenegre, Carlos H. Barrios, Masahiro Morise, Enriqueta Felip, Zoran Andric, Sarayut Geater, Mustafa Özgüroğlu, Wei Zou, Alan Sandler, Ida Enquist, Kimberly Komatsubara, Yu Deng, Hiroshi Kuriki, Xiaohui Wen, Mark McCleland, Simonetta Mocci, Jacek Jassem, and David R. Spigel. Atezolizumab for first-line treatment of pd-l1–selected patients with nsclc. *New England Journal of Medicine*, 383(14):1328–1339, 2020. PMID: 32997907.

[157] Yeela Shamai, Dalia Cohn Alperovich, Zohar Yakhini, Karl Skorecki, and Maty Tzukerman. Reciprocal reprogramming of cancer cells and associated mesenchymal stem cells in gastric cancer. *STEM CELLS*, 37:176–189, 2 2019.

[158] Yan Ting Shue, Jing Shan Lim, and Julien Sage. Tumor heterogeneity in small cell lung cancer defined and investigated in pre-clinical mouse models. *Translational Lung Cancer Research*, 7(1), 2018.

[159] Joaquim Calbo, Erwin van Montfort, Natalie Proost, Ellen van Drunen, H. Berna Beverloo, Ralph Meuwissen, and Anton Berns. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. *Cancer Cell*, 19:244–256, 2 2011.

[160] Ravi Salgia, Isa Mambetsariev, Blake Hewelt, Srisairam Achuthan, Haiqing Li, Valeriy Poroyko, Yingyu Wang, Martin Sattler, Ravi Salgia, Isa Mambetsariev, Blake Hewelt, Srisairam Achuthan, Haiqing Li, Valeriy Poroyko, Yingyu Wang, Martin Sattler, Ravi Salgia, Isa Mambetsariev, Blake Hewelt, Srisairam Achuthan, Haiqing Li, Valeriy Poroyko, Yingyu Wang, and Martin Sattler. Modeling small cell lung cancer (sclc) biology through deterministic and stochastic mathematical models. *Oncotarget*, 9:26226–26242, 5 2018.

[161] Eshan D. Mitra, Raquel Dias, Richard G. Posner, and William S. Hlavacek. Using both qualitative and quantitative data in parameter identification for systems biology models. *Nature Communications*, 9:3901, 12 2018.

[162] Alejandro F. Villaverde, Antonio Barreiro, and Antonis Papachristodoulou. Structural identifiability of dynamic systems biology models. *PLOS Computational Biology*, 12:e1005153, 10 2016.

[163] Paul Kirk, Thomas Thorne, and Michael PH Stumpf. Model selection in systems and synthetic biology. *Current Opinion in Biotechnology*, 24:767–774, 4 2013.

[164] John D Blischak, Emily R Davenport, and Greg Wilson. A quick introduction to version control with git and github. *PLoS Computational Biology*, 12:e1004668, 1 2016.

[165] Vijayalakshmi Chelliah, Camille Laibe, and Nicolas Le Novère. Biomodels database: a repository of mathematical models of biological processes. *Methods in Molecular Biology*, 1021:189–199, 2013.

[166] Wen Xu, Adam M Smith, James R Faeder, and G Elisabeta Marai. Rulebender: a visual interface for rule-based modeling. *Bioinformatics*, 27:1721–1722, 4 2011.

[167] John E Wenskovitch, Leonard A Harris, Jose-Juan Tapia, James R Faeder, and G Elisabeta Marai. Mosbie: a tool for comparison and analysis of rule-based biochemical models. *BMC Bioinformatics*, 15:316, 1 2014.

[168] Lily A Chylek, Leonard A Harris, Chang-Shung Tung, James R Faeder, Carlos F Lopez, and William S Hlavacek. Rule-based modeling: a computational approach for studying biomolecular site dynamics in cell signaling systems. *Wiley Interdisciplinary Reviews Systems Biology and Medicine*, 6:13–36, 9 2014.

[169] Lily A Chylek, Leonard A Harris, James R Faeder, and William S Hlavacek. Modeling for (physical) biologists: an introduction to the rule-based approach. *Physical Biology*, 12:045007, 7 2015.

[170] Wes McKinney. Data Structures for Statistical Computing in Python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 56 – 61, 2010.

[171] L.A. Harris, M.S. Nobile, J.C. Pino, A.L.R. Lubbock, D. Besozzi, G. Mauri, P. Cazzaniga, and C.F. Lopez. Gpu-powered model analysis with pysb/cupsoda. *Bioinformatics*, 33:3492–3494, 2017.

[172] Erin M Shockley, Jasper A Vrugt, and Carlos F Lopez. Pydream: high-dimensional parameter inference for biological models in python. *Bioinformatics*, 34:695–697, 2 2018.

[173] Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS computational biology*, 3:1871–1878, 10 2007.

[174] Z. Zi. Sensitivity analysis approaches applied to systems biology models. *IET Systems Biology*, 5:336–346, 11 2011.

[175] Maksat Ashyraliyev, Yves Fomekong-Nanfack, Jaap A. Kaandorp, and Joke G. Blom. Systems biology: parameter estimation for biochemical models. *FEBS Journal*, 276:886–902, 2 2009.

[176] Hoda Eydgahi, William W Chen, Jeremy L Muhlich, Dennis Vitkup, John N Tsitsiklis, and Peter K Sorger. Properties of cell death models calibrated and compared using bayesian approaches. *Molecular systems biology*, 9:644, 1 2013.

[177] Christian Tönsing, Jens Timmer, and Clemens Kreutz. Cause and cure of sloppiness in ordinary differential equation models. *Physical Review E*, 90:023303, 8 2014.

[178] Robert E Kass and Adrian E Raftery. Bayes factors. *Journal of the American Statistical Association*, 90:773–795, 1995.

[179] Gurkan Mollaoglu, Matthew R. Guthrie, Stefanie Böhm, Johannes Brägelmann, Ismail Can, Paul M. Ballieu, Annika Marx, Julie George, Christine Heinen, Milind D. Chalishazar, Haixia Cheng, Abbie S. Ireland, Kendall E. Denning, Anandaroop Mukhopadhyay, Jeffery M. Vahrenkamp, Kristofer C. Berrett, Timothy L. Mosbruger, Jun Wang, Jessica L. Kohan, Mohamed E. Salama, Benjamin L. Witt, Martin Peifer, Roman K. Thomas, Jason Gertz, Jane E. Johnson, Adi F. Gazdar, Robert J. Wechsler-Reya, Martin L. Sos, and Trudy G. Oliver. Myc drives progression of small cell lung cancer to a variant neuroendocrine subtype with vulnerability to aurora kinase inhibition. *Cancer Cell*, 31:270–285, 2 2017.

[180] David J. Wooten, Sarah M. Groves, Darren R. Tyson, Qi Liu, Jing S. Lim, Réka Albert, Carlos F. Lopez, Julien Sage, and Vito Quaranta. Systems-level network modeling of small cell lung cancer subtypes identifies master regulators and destabilizers. *PLOS Computational Biology*, 15:e1007343, 10 2019.

[181] Eshan D Mitra and William S Hlavacek. Parameter estimation and uncertainty quantification for systems biology models. *Current Opinion in Systems Biology*, 18:9–18, 2019.

[182] Don van Ravenzwaaij, Pete Cassey, and Scott D. Brown. A simple introduction to markov chain monte–carlo sampling. *Psychonomic Bulletin and Review*, 25:143–154, 2 2018.

[183] John. Kruschke. *Doing Bayesian data analysis : a tutorial introduction with R.* Elsevier Science & Technology, 2 edition, 2014.

[184] Paul H Garthwaite and Emmanuel Mubwandarikwa. Selection of weights for weighted model averaging. *Aust. N. Z. J. Stat*, 52:363–382, 2010.

[185] Nikesh Kotecha, Peter O. Krutzik, and Jonathan M. Irish. Web-based analysis and publication of flow cytometry experiments. *Current Protocols in Cytometry*, 53:10.17.1–10.17.24, 7 2010.

[186] Nalin Leelatian, Kirsten E. Diggins, and Jonathan M. Irish. Characterizing phenotypes and signaling networks of single human cells by mass cytometry. *Methods in Molecular Biology*, 1346:99–113, 2015.

[187] Kirsten E. Diggins, Allison R. Greenplate, Nalin Leelatian, Cara E. Wogsland, and Jonathan M. Irish. Characterizing cell subsets using marker enrichment modeling. *Nature Methods 2017 14:3*, 14:275–278, 1 2017.

[188] Max Hinne, Quentin F. Gronau, Don van den Bergh, and Eric-Jan Wagenmakers. A conceptual introduction to bayesian model averaging. *Advances in Methods and Practices in Psychological Science*, 3:200–215, 2020.

[189] Manfred Wildner. In memory of william of occam. *Lancet*, 354:2172, 12 1999.

[190] Zihang Lu and Wendy Lou. Bayesian approaches to variable selection: A comparative study from practical perspectives. *International Journal of Biostatistics*, 3 2021.

[191] Rens van de Schoot, Sarah Depaoli, Ruth King, Bianca Kramer, Kaspar Märtens, Mahlet G. Tadesse, Marina Vannucci, Andrew Gelman, Duco Veen, Joukje Willemsen, and Christopher Yau. Bayesian statistics and modelling. *Nature Reviews Methods Primers 2021 1:1*, 1:1–26, 1 2021.

[192] Ruth E. Baker, Jose Maria Peña, Jayaratnam Jayamohan, and Antoine Jérusalem. Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters*, 14, 2018.

[193] Stefan M. Kallenberger, Joël Beaudouin, Juliane Claus, Carmen Fischer, Peter K. Sorger, Stefan Legewie, and Roland Eils. Intra- and interdimeric caspase-8 self-cleavage controls strength and timing of cd95-induced apoptosis. *Science Signaling*, 7, 3 2014.

[194] Michael A. Kochen and Carlos F. Lopez. A probabilistic approach to explore signal execution mechanisms with limited experimental data. *Frontiers in Genetics*, 11:686, 7 2020.

[195] Sabrina L. Spencer and Peter K. Sorger. Measuring and modeling apoptosis in single cells. *Cell*, 144:926–939, 3 2011.

[196] Carl M. Gay, C. Allison Stewart, Elizabeth M. Park, Lixia Diao, Sarah M. Groves, Simon Heeke, Barzin Y. Nabet, Junya Fujimoto, Luisa M. Solis, Wei Lu, Yuanxin Xi, Robert J. Cardnell, Qi Wang, Giulia Fabbri, Kasey R. Cargill, Natalie I. Vokes, Kavya Ramkumar, Bingnan Zhang, Carminia M. Della Corte, Paul Robson, Stephen G. Swisher, Jack A. Roth, Bonnie S. Glisson, David S. Shames, Ignacio I. Wistuba, Jing Wang, Vito Quaranta, John Minna, John V. Heymach, and Lauren Averett Byers. Patterns of transcription factor programs and immune pathway activation define four major subtypes of sclc with distinct therapeutic vulnerabilities. *Cancer Cell*, 39:346–360.e7, 3 2021.

[197] Sarah M Groves, Geena V Ildefonso, Caitlin O Mcatee, Trudy G Oliver, Alissa M Weaver, and Vito Quaranta. Archetype tasks link intratumoral heterogeneity to plasticity and cancer hallmarks in small cell lung cancer. *Cell Systems*, 13:690–710, 2022.

[198] Andrew T. Tredennick, Giles Hooker, Stephen P. Ellner, and Peter B. Adler. A practical guide to selecting models for exploration, inference, and prediction in ecology. *Ecology*, 102:e03336, 6 2021.

[199] Wendy S. Parker. Ensemble modeling, uncertainty and robust predictions. *Wiley Interdisciplinary Reviews: Climate Change*, 4:213–223, 5 2013.

[200] László Zsolt Garamszegi and Roger Mundry. Multimodel-inference in comparative analyses. In László Zsolt Garamszegi, editor, *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice*, chapter 12, pages 305–331. Springer Berlin, 2014.

[201] Paola Lecca. Machine learning for causal inference in biological networks: Perspectives of this challenge. *Frontiers in Bioinformatics*, 1:45, 9 2021.

[202] Karen G. Zeman, Christina E. Brzezniak, and Corey A. Carter. Recalcitrant small cell lung cancer: the argument for optimism. *Journal of Thoracic Disease*, 9:E295–E296, 3 2017.

[203] Stuart C. Williamson, Robert L. Metcalf, Francesca Trapani, Sumitra Mohan, Jenny Antonello, Benjamin Abbott, Hui Sun Leong, Christopher P.E. Chester, Nicole Simms, Radoslaw Polanski, Daisuke Nonaka, Lynsey Priest, Alberto Fusi, Fredrika Carlsson, Anders Carlsson, Mary J.C. Hendrix, Richard E.B. Seftor, Elisabeth A. Seftor, Dominic G. Rothwell, Andrew Hughes, James Hicks, Crispin Miller, Peter Kuhn, Ged Brady, Kathryn L. Simpson, Fiona H. Blackhall, and Caroline Dive. Vasculogenic mimicry in small cell lung cancer. *Nature Communications 2016 7:1*, 7:1–14, 11 2016.

[204] Joseph M. Chan, Álvaro Quintanal-Villalonga, Vianne Ran Gao, Yubin Xie, Viola Allaj, Ojasvi Chaudhary, Ignas Masilionis, Jacklynn Egger, Andrew Chow, Thomas Walle, Marissa Mattar, Dig V.K. Yarlagadda, James L. Wang, Fathema Uddin, Michael Offin, Metamia Ciampricotti, Besnik Qeriqi, Amber Bahr, Elisa de Stanchina, Umesh K. Bhanot, W. Victoria Lai, Matthew J. Bott, David R. Jones, Arvin Ruiz, Marina K. Baine, Yanyun Li, Natasha Rekhtman, John T. Poirier, Tal Nawy, Triparna Sen, Linas Mazutis, Travis J. Hollmann, Dana Pe'er, and Charles M. Rudin. Signatures of plasticity, metastasis,

and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell*, 39:1479–1496.e18, 11 2021.

[205] Kathryn L. Simpson, Ruth Stoney, Kristopher K. Frese, Nicole Simms, William Rowe, Simon P. Pearce, Sam Humphrey, Laura Booth, Derrick Morgan, Marek Dynowski, Francesca Trapani, Alessia Catozzi, Mitchell Revill, Thomas Helps, Melanie Galvin, Luc Girard, Daisuke Nonaka, Louise Carter, Matthew G. Krebs, Natalie Cook, Mathew Carter, Lynsey Priest, Alastair Kerr, Adi F. Gazdar, Fiona Blackhall, and Caroline Dive. A biobank of small cell lung cancer cdx models elucidates inter- and intratumoral phenotypic heterogeneity. *Nature Cancer 2020 1:4*, 1:437–451, 4 2020.

[206] Yu Han Huang, Olaf Klingbeil, Xue Yan He, Xiaoli S. Wu, Gayatri Arun, Bin Lu, Tim D.D. Somerville, Joseph P. Milazzo, John E. Wilkinson, Osama E. Demerdash, David L. Spector, Mikala Egeblad, Junwei Shi, and Christopher R. Vakoc. Pou2f3 is a master regulator of a tuft cell-like variant of small cell lung cancer. *Genes & Development*, 32:915–928, 7 2018.

[207] Youcef Ouadah, Enrique R. Rojas, Daniel P. Riordan, Sarah Capostagno, Christin S. Kuo, and Mark A. Krasnow. Rare pulmonary neuroendocrine cells are stem cells regulated by rb, p53, and notch. *Cell*, 179:403–416.e23, 10 2019.

[208] Yan Ting Shue, Alexandros P Drainas, Nancy Yanzhe Li, Sarah M Pearsall, Derrick Morgan, Nasa Sinnott-Armstrong, Susan Q Hipkins, Garry L Coles, Jing Shan Lim, Anthony E Oro, Kathryn L Simpson, Caroline Dive, and Julien Sage. A conserved yap/notch/rest network controls the neuroendocrine cell fate in the lungs. *Nature Communications*, 13, 2022.

[209] Rachelle R Olsen, Abbie S Ireland, David W Kastner, Sarah M Groves, Kyle B Spainhower, Karine Pozo, Demetra P Kelenis, Christopher P Whitney, Matthew R Guthrie, Sarah J Wait, Danny Soltero, Benjamin L Witt, Vito Quaranta, Jane E Johnson, and Trudy G Oliver. Ascl1 represses a sox9 + neural crest stem-like state in small cell lung cancer. *Genes & Development*, 35:847–869, 2021.

[210] Jordi Barretina, Giordano Caponigro, Nicolas Stransky, Kavitha Venkatesan, Adam A. Margolin, Sungjoon Kim, Christopher J. Wilson, Joseph Lehár, Gregory V. Kryukov, Dmitriy Sonkin, Anupama Reddy, Manway Liu, Lauren Murray, Michael F. Berger, John E. Monahan, Paula Morais, Jodi Meltzer, Adam Korejwa, Judit Jané-Valbuena, Felipa A. Mapa, Joseph Thibault, Eva Bric-Furlong, Pichai Raman, Aaron Shipway, Ingo H. Engels, Jill Cheng, Guoying K. Yu, Jianjun Yu, Peter Aspesi, Melanie De Silva, Kalpana Jagtap, Michael D. Jones, Li Wang, Charles Hatton, Emanuele Palescandolo, Supriya Gupta, Scott Mahan, Carrie Sougnez, Robert C. Onofrio, Ted Liefeld, Laura MacConaill, Wendy Winckler, Michael Reich, Nanxin Li, Jill P. Mesirov, Stacey B. Gabriel, Gad Getz, Kristin Ardlie, Vivien Chan, Vic E. Myer, Barbara L. Weber, Jeff Porter, Markus Warmuth, Peter Finan, Jennifer L. Harris, Matthew Meyerson, Todd R. Golub, Michael P. Morrissey, William R. Sellers, Robert Schlegel, and Levi A. Garraway. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature 2012 483:7391*, 483:603–607, 3 2012.

[211] Leonard A. Harris, Samantha Beik, Patricia M.M. Ozawa, Lizandra Jimenez, and Alissa M. Weaver. Modeling heterogeneous tumor growth dynamics and cell–cell interactions at single-cell and cell-population resolution. *Current Opinion in Systems Biology*, 17:24–34, 10 2019.

[212] C. Allison Stewart, Carl M. Gay, Yuanxin Xi, Santhosh Sivajothi, V. Sivakamasundari, Junya Fujimoto, Mohan Bolisetty, Patrice M. Hartsfield, Veerakumar Balasubramaniyan, Milind D. Chalishazar, Cesar Moran, Neda Kalhor, John Stewart, Hai Tran, Stephen G. Swisher, Jack A. Roth, Jianjun Zhang, John de Groot, Bonnie Glisson, Trudy G. Oliver, John V. Heymach, Ignacio Wistuba, Paul Robson, Jing Wang, and Lauren Averett Byers. Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. *Nature Cancer*, 1:423–436, 4 2020.

[213] Corey E. Hayford, Darren R. Tyson, C. Jack Robbins, Peter L. Frick, Vito Quaranta, and Leonard A. Harris. *An in vitro model of tumor heterogeneity resolves genetic, epigenetic, and stochastic sources of cell state variability*, volume 19. Public Library of Science, 2021.

[214] Sydney M Shaffer, Margaret C Dunagin, Stefan R Torborg, Eduardo A Torre, Benjamin Emert, Clemens Krepler, Marilda Beqiri, Katrin Sproesser, Patricia A Brafford, Min Xiao, Elliott Eggan, Ioannis N Anastopoulos, Cesar A Vargas-Garcia, Abhyudai Singh, Katherine L Nathanson, Meenhard Herlyn, and Arjun Raj. Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature Publishing Group*, 2018.

[215] Spencer S Watson, Mark Dane, Koei Chin, Oliver Jonas, Joe W Gray, and James E Korkola. Microenvironment-mediated mechanisms of resistance to her2 inhibitors differ between her2+ breast cancer subtypes. *Cell Systems*, 6:329–342, 2018.

[216] Min-Chul Kwon, Natalie Proost, Ji-Ying Song, Kate D Sutherland, John Zevenhoven, and Anton Berns. Paracrine signaling between tumor subclones of mouse sclc: a critical role of ets transcription factor pea3 in facilitating metastasis. *Genes & Development*, 2015.

[217] John T. Poirier, Julie George, Taofeek K. Owonikoko, Anton Berns, Elisabeth Brambilla, Lauren A. Byers, David Carbone, Huanhuan J. Chen, Camilla L. Christensen, Caroline Dive, Anna F. Farago, Ramaswamy Govindan, Christine Hann, Matthew D. Hellmann, Leora Horn, Jane E. Johnson, Young S. Ju, Sumin Kang, Mark Krasnow, James Lee, Se Hoon Lee, Jonathan Lehman, Benjamin Lok, Christine Lovly, David MacPherson, David McFadden, John Minna, Matthew Oser, Keunchil Park, Kwon Sik Park, Yves Pommier, Vito Quaranta, Neal Ready, Julien Sage, Giorgio Scagliotti, Martin L. Sos, Kate D. Sutherland, William D. Travis, Christopher R. Vakoc, Sarah J. Wait, Ignacio Wistuba, Kwok Kin Wong, Hua Zhang, Jillian Daigneault, Jacinta Wiens, Charles M. Rudin, and Trudy G. Oliver. New approaches to sclc therapy: From the laboratory to the clinic. *Journal of Thoracic Oncology*, 15:520–540, 4 2020.

[218] Matthew H. Spitzer and Garry P. Nolan. Mass cytometry: Single cells, many features. *Cell*, 165:780–791, 5 2016.

[219] Byungjin Hwang, Ji Hyun Lee, and Duhee Bang. Single-cell rna sequencing technologies and bioinformatics pipelines. *Experimental & Molecular Medicine 2018 50:8*, 50:1–14, 8 2018.

[220] Yukie Kashima, Yosuke Togashi, Shota Fukuoka, Takahiro Kamada, Takuma Irie, Ayako Suzuki, Yoshiaki Nakamura, Kohei Shitara, Tatsunori Minamide, Taku Yoshida, Naofumi Taoka, Tatsuya Kawase, Teiji Wada, Koichiro Inaki, Masataka Chihara, Yukihiko Ebisuno, Sakiyo Tsukamoto, Ryo Fujii, Akihiro Ohashi, Yutaka Suzuki, Katsuya Tsuchihara, Hiroyoshi Nishikawa, and Toshihiko Doi. Potentiality of multiple modalities for single-cell analyses to evaluate the tumor microenvironment in clinical specimens. *Scientific Reports*, 11:341, 2018.

[221] Christine L. Hann and Charles M. Rudin. Management of small-cell lung cancer: Incremental changes but hope for the future. *Oncology (Williston Park, N.Y.)*, 22:1486, 11 2008.

[222] Haobin Chen, Lisa Gesumaria, Young-Kwon Park, Trudy G. Oliver, Dinah S. Singer, Kai Ge, and David S. Schrump. Bet inhibitors target the sclc-n subtype of small-cell lung cancer by blocking neurod1 transactivation. *Molecular Cancer Research*, pages OF1–OF11, 12 2022.

[223] Marcello Stanzione, Jun Zhong, Edmond Wong, Thomas J. LaSalle, Jillian F. Wise, Antoine Simoneau, David T. Myers, Sarah Phat, Moshe Sade-Feldman, Michael S. Lawrence, M. Kyle Hadden, Lee Zou, Anna F. Farago, Nicholas J. Dyson, and Benjamin J. Drapkin. Translesion dna synthesis mediates acquired resistance to olaparib plus temozolomide in small cell lung cancer. *Science Advances*, 8:1229, 5 2022.

[224] Rachel Finck, Erin F. Simonds, Astraea Jager, Smita Krishnaswamy, Karen Sachs, Wendy Fantl, Dana Pe'er, Garry P. Nolan, and Sean C. Bendall. Normalization of mass cytometry data with bead standards. *Cytometry Part A*, 83A:483–494, 5 2013.

[225] V. A. Traag, L. Waltman, and N. J. van Eck. From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports 2019 9:1*, 9:1–12, 3 2019.

[226] Stephen Lyle and Nathan Moore. Quiescent, slow-cycling stem cell populations in cancer: A review of the evidence and discussion of significance. *Journal of Oncology*, 2011.

[227] Wanyin Chen, Jihu Dong, Jacques Haiech, Marie Claude Kilhoffer, and Maria Zeniou. Cancer stem cell quiescence and plasticity as major challenges in cancer therapy. *Stem Cells International*, 2016, 2016.

[228] Dinoop Ravindran Menon, Heinz Hammerlindl, Joachim Torrano, Helmut Schaider, and Mayumi Fujita. Epigenetics and metabolism at the crossroads of stress-induced plasticity, stemness and therapeutic resistance in cancer. *Theranostics*, 10:6261, 2020.

[229] El Ad David Amir, Kara L. Davis, Michelle D. Tadmor, Erin F. Simonds, Jacob H. Levine, Sean C. Bendall, Daniel K. Shenfeld, Smita Krishnaswamy, Garry P. Nolan, and Dana Pe'Er. visne enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature Biotechnology 2013 31:6*, 31:545–552, 5 2013.

[230] Peng Qiu, Erin F. Simonds, Sean C. Bendall, Kenneth D. Gibbs, Robert V. Bruggner, Michael D. Linderman, Karen Sachs, Garry P. Nolan, and Sylvia K. Plevritis. Extracting a cellular hierarchy from high-dimensional cytometry data with spade. *Nature Biotechnology 2011 29:10*, 29:886–891, 10 2011.

[231] Shaoheng Liang, Vakul Mohanty, Jinzhuang Dou, Qi Miao, Yuefan Huang, Muharrem Müftüoğlu, Li Ding, Weiyi Peng, and Ken Chen. Single-cell manifold-preserving feature selection for detecting rare cell populations. *Nature Computational Science 2021 1:5*, 1:374–384, 5 2021.

[232] Piyush B. Gupta, Ievgenia Pastushenko, Adam Skibinski, Cedric Blanpain, and Charlotte Kuperwasser. Phenotypic plasticity: Driver of cancer initiation, progression, and therapy resistance. *Cell Stem Cell*, 24(1):65–78, 2019.

[233] Priyanka Gopal, Aaron Petty, Kevin Rogacki, Titas Bera, Rohan Bareja, Craig D Peacock, and Mohamed E Abazeed. Multivalent state transitions shape the intratumoral composition of small cell lung carcinoma. *Science Advances*, 8:eabp8674, 2022.

[234] Adi F. Gazdar, Desmond N. Carney, Edward K. Russell, Harvie L. Sims, Stephen B. Baylin, Paul A. Bunn Jr., John G. Guccion, and John D. Minna. Establishment of continuous, clonable cultures of small-cell carcinoma of lung which have amine precursor uptake and decarboxylation cell properties. *Cancer Research*, 40:3502–3507, 1980.

[235] Carl M. Gay, Pan Tong, Robert J. Cardnell, Triparna Sen, Xiao Su, Jun Ma, Rasha O. Bara, Faye M. Johnson, Chris Wakefield, John V. Heymach, Jing Wang, and Lauren A. Byers. Differential sensitivity analysis for resistant malignancies (disarm) identifies common candidate therapies across platinum-resistant cancers. *Clinical Cancer Research*, 25:346–357, 1 2019.

[236] C. Allison Stewart, Pan Tong, Robert J. Cardnell, Triparna Sen, Lerong Li, Carl M. Gay, Fatemah Masrorpour, Youhong Fan, Rasha O. Bara, Ying Feng, Yuanbin Ru, Junya Fujimoto, Samrat T. Kundu, Leonard E. Post, Guoying K. Yu, Yuqiao Shen, Bonnie S. Glisson, Ignacio Wistuba, John V. Heymach, Don L. Gibbons, Jing Wang, Lauren A. Byers, C. Allison Stewart, Pan Tong, Robert J. Cardnell, Triparna Sen, Lerong Li, Carl M. Gay, Fatemah Masrorpour, Youhong Fan, Rasha O. Bara, Ying Feng, Yuanbin Ru, Junya Fujimoto, Samrat T. Kundu, Leonard E. Post, Guoying K. Yu, Yuqiao Shen, Bonnie S. Glisson, Ignacio Wistuba, John V. Heymach, Don L. Gibbons, Jing Wang, and Lauren A. Byers. Dynamic variations in epithelial-to-mesenchymal transition (emt), atm, and slfn11 govern response to parp inhibitors and cisplatin in small cell lung cancer. *Oncotarget*, 8:28575–28587, 2 2017.

[237] Eric Polley, Mark Kunkel, David Evans, Thomas Silvers, Rene Delosh, Julie Laudeman, Chad Ogle, Russell Reinhart, Michael Selby, John Connelly, Erik Harris, Nicole Fer, Dmitriy Sonkin, Gurmeet Kaur, Anne Monks, Shakun Malik, Joel Morris, and Beverly A. Teicher. Small cell lung cancer screen of oncology drugs, investigational agents, and gene and microrna expression. *JNCI: Journal of the National Cancer Institute*, 108, 10 2016.

[238] Bita Sahaf, Adeeb Rahman, Holden T. Maecker, and Sean C. Bendall. High-parameter immune profiling with cytof. In Magdalena Thurin, Alessandra Cesano, and Francesco M. Marincola, editors, *Biomarkers for Immunotherapy of Cancer: Methods and Protocols*, pages 351–368. Springer New York, New York, NY, 2020.

[239] Justus M. Kebschull and Anthony M. Zador. Cellular barcoding: lineage tracing, screening and beyond. *Nature Methods*, 15(11):871–879, Nov 2018.

[240] Livius Penter, Kerstin Dietze, Lars Bullinger, Jörg Westermann, Hans-Peter Rahn, and Leo Hansmann. Facs single cell index sorting is highly reliable and determines immune phenotypes of clonally expanded t cells. *European Journal of Immunology*, 48(7):1248–1250, 2018.

[241] Mahmoud Ghandi, Franklin W. Huang, Judit Jané-Valbuena, Gregory V. Kryukov, Christopher C. Lo, E. Robert McDonald, Jordi Barretina, Ellen T. Gelfand, Craig M. Bielski, Haoxin Li, Kevin Hu, Alexander Y. Andreev-Drakhlin, Jaegil Kim, Julian M. Hess, Brian J. Haas, François Aguet, Barbara A. Weir, Michael V. Rothberg, Brenton R. Paolella, Michael S. Lawrence, Rehan Akbani, Yiling Lu, Hong L. Tiv, Prafulla C. Gokhale, Antoine de Weck, Ali Amin Mansour, Coyin Oh, Juliann Shih, Kevin Hadi, Yanay Rosen, Jonathan Bistline, Kavitha Venkatesan, Anupama Reddy, Dmitriy Sonkin, Manway Liu, Joseph Lehar, Joshua M. Korn, Dale A. Porter, Michael D. Jones, Javad Golji, Giordano Caponigro, Jordan E. Taylor, Caitlin M. Dunning, Amanda L. Creech, Allison C. Warren, James M. McFarland, Mahdi Zamanighomi, Audrey Kauffmann, Nicolas Stransky, Marcin Imielinski, Yosef E. Maruvka, Andrew D. Cherniack, Aviad Tsherniak, Francisca Vazquez, Jacob D. Jaffe, Andrew A. Lane, David M. Weinstock, Cory M. Johannessen, Michael P. Morrissey, Frank Stegmeier, Robert Schlegel, William C. Hahn, Gad Getz, Gordon B. Mills, Jesse S. Boehm, Todd R. Golub, Levi A. Garraway, and William R. Sellers. Next-generation characterization of the cancer cell line encyclopedia. *Nature 2019 569:7757*, 569:503–508, 5 2019.

[242] Mingjie Liu, Saad Rehman, Xidian Tang, Kui Gu, Qinlei Fan, Dekun Chen, and Wentao Ma. Methodologies for improving hdr efficiency. *Frontiers in Genetics*, 10:691, 1 2019.

[243] B. Pardo, B. Gómez-González, and A. Aguilera. Dna double-strand break repair: How to fix a broken relationship. *Cellular and Molecular Life Sciences*, 66:1039–1056, 3 2009.

[244] Fiona Cunningham, James E. Allen, Jamie Allen, Jorge Alvarez-Jarreta, M. Ridwan Amode, Irina M. Armean, Olanrewaju Austine-Orimoloye, Andrey G. Azov, If Barnes, Ruth Bennett, Andrew Berry, Jyothish Bhai, Alexandra Bignell, Konstantinos Billis, Sanjay Boddu, Lucy Brooks, Mehrnaz Charkhchi, Carla Cummins, Luca Da Rin Fioretto, Claire Davidson, Kamalkumar Dodiya, Sarah Donaldson, Bilal El Houdaigui, Tamara El Naboulsi, Reham Fatima, Carlos Garcia Giron, Thiago Genez, Jose Gonzalez Martinez, Cristina Guijarro-Clarke, Arthur Gymer, Matthew Hardy, Zoe Hollis, Thibaut Hourlier, Toby Hunt, Thomas Juettemann, Vinay Kaikala, Mike Kay, Ilias Lavidas, Tuan Le, Diana Lemos, Jose Carlos Marugán, Shamika Mohanan, Aleena Mushtaq, Marc Naven, Denye N. Ogeh, Anne Parker, Andrew Parton, Malcolm Perry, Ivana Pilizota, Irina Prosovetskaia, Manoj Pandian Sakthivel, Ahamed Imran Abdul Salam, Bianca M. Schmitt, Helen Schuilenburg, Dan Sheppard, Jose G. Perez-Silva, William Stark, Emily Steed, Kyösti Sutinen, Ranjit Sukumaran, Dulika Sumathipala, Marie Marthe Suner, Michal Szpak, Anja Thormann, Francesca Floriana Tricomi, David Urbina-Gómez, Andres Veidenberg, Thomas A. Walsh, Brandon Walts, Natalie Willhoft, Andrea Winterbottom, Elizabeth Wass, Marc Chakiachvili, Bethany Flint, Adam Frankish, Stefano Giorgetti, Leanne Haggerty, Sarah E. Hunt, Garth R. Iisley, Jane E. Loveland, Fergal J. Martin, Benjamin Moore, Jonathan M. Mudge, Matthieu Muffato, Emily Perry, Magali Ruffier, John Tate, David Thybert, Stephen J. Trevanion, Sarah Dyer, Peter W. Harrison, Kevin L. Howe, Andrew D. Yates, Daniel R. Zerbino, and Paul Flicek. Ensembl 2022. *Nucleic Acids Research*, 50:D988–D995, 1 2022.

[245] Uri Ben-David, Benjamin Siranosian, Gavin Ha, Helen Tang, Yaara Oren, Kunihiko Hinohara, Craig A. Strathdee, Joshua Dempster, Nicholas J. Lyons, Robert Burns, Anwesha Nag, Guillaume Kugener, Beth Cimini, Peter Tsvetkov, Yosef E. Maruvka, Ryan O'Rourke, Anthony Garrity, Andrew A. Tubelli, Pratiti Bandopadhayay, Aviad Tsherniak, Francisca Vazquez, Bang Wong, Chet Birger, Mahmoud Ghandi, Aaron R. Thorner, Joshua A. Bittker, Matthew Meyerson, Gad Getz,

Rameen Beroukhim, and Todd R. Golub. Genetic and transcriptional evolution alters cancer cell line drug response. *Nature*, 560(7718):325–330, Aug 2018.

[246] Feng Zhang, Yan Wen, and Xiong Guo. CRISPR/Cas9 for genome editing: progress, implications and challenges. *Human Molecular Genetics*, 23(R1):R40–R46, 03 2014.

[247] Kornel Labun, Maximilian Krause, Yamila Torres Cleuren, and Eivind Valen. Crispr genome editing made easy through the chopchop website. *Current Protocols*, 1(4):e46, 2021.

[248] Kornel Labun, Tessa G Montague, Maximilian Krause, Yamila N Torres Cleuren, Håkon Tjeldnes, and Eivind Valen. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, 47(W1):W171–W174, 05 2019.

[249] Kai Li, Gang Wang, Troels Andersen, Pingzhu Zhou, and William T. Pu. Optimization of genome engineering approaches with the crispr/cas9 system. *PLOS ONE*, 9:e105779, 8 2014.

[250] Fei Song and Knut Stieger. Optimizing the dna donor template for homology-directed repair of double-strand breaks. *Molecular Therapy - Nucleic Acids*, 7:53–60, 6 2017.

[251] Talley J. Lambert. Fpbase: a community-editable fluorescent protein database. *Nature Methods*, 16(4):277–278, Apr 2019.

[252] Dmitry Shcherbo, Christopher S. Murphy, Galina V. Ermakova, Elena A. Solovieva, Tatiana V. Chepurnykh, Aleksandr S. Shcheglov, Vladislav V. Verkhusha, Vladimir Z. Pletnev, Kristin L. Hazelwood, Patrick M. Roche, Sergey Lukyanov, Andrey G. Zaraisky, Michael W. Davidson, and Dmitriy M. Chudakov. Far-red fluorescent tags for protein imaging in living tissues. *Biochemical Journal*, 418(3):567–574, 02 2009.

[253] Satoshi Karasawa, Toshio Araki, Miki Yamamoto-Hino, and Atsushi Miyawaki. A green-emitting fluorescent protein from galaxeidae coral and its monomeric version for use in fluorescent labeling. *Journal of Biological Chemistry*, 278:34167–34171, 9 2003.

[254] Vishnu Priyanka Reddy Chichili, Veerendra Kumar, and J. Sivaraman. Linkers in the structural biology of protein–protein interactions. *Protein Science*, 22(2):153–167, 2013.

[255] Zhi Xiong Chong, Swee Keong Yeap, and Wan Yong Ho. Transfection types, methods and strategies: A technical review. *PeerJ*, 9:e11165, 4 2021.

[256] Sojung Kim, Daesik Kim, Seung Woo Cho, Jungeun Kim, and Jin Soo Kim. Highly efficient rna-guided genome editing in human cells via delivery of purified cas9 ribonucleoproteins. *Genome Research*, 24:1012–1019, 6 2014.

[257] Gurmeet Kaur, Russell A. Reinhart, Anne Monks, David Evans, Joel Morris, Eric Polley, and Beverly A. Teicher. Bromodomain and hedgehog pathway targets in small cell lung cancer. *Cancer Letters*, 371:225–239, 2 2016.

[258] Olive S. Pettengill, George D. Sorenson, Doris H. Wurster-Hill, Thomas J. Curphey, Walter W. Noll, Charles C. Gate, and L. Herbert Maurer. Isolation and growth characteristics of continuous cell lines from small-cell carcinoma of the lung. *Cancer*, 45, 1980.

[259] Florin Sirzén, Boris Zhivotovsky, Anders Nilsson, Jonas Bergh, and Rolf Lewensohn. Higher spontaneous apoptotic index in small cell compared with non-small cell lung carcinoma cell lines; lack of correlation with bcl-2/bax. *Lung Cancer*, 22:1–13, 10 1998.

[260] C. Turner and M. Kohandel. Quantitative approaches to cancer stem cells and epithelial – mesenchymal transition. *Seminars in Cancer Biology*, 22:374–378, 10 2012.

[261] Mingyang Lu, Mohit Kumar Jolly, Herbert Levine, José N. Onuchic, and Eshel Ben-Jacob. Micrornа-based regulation of epithelial–hybrid–mesenchymal fate determination. *Proceedings of the National Academy of Sciences*, 110:18144–18149, 11 2013.

[262] Zhuorui Xie, Allison Bailey, Maxim V. Kuleshov, Daniel J.B. Clarke, John E. Evangelista, Sherry L. Jenkins, Alexander Lachmann, Megan L. Wojciechowicz, Eryk Kropiwnicki, Kathleen M. Jagodnik, Minji Jeon, and Avi Ma'ayan. Gene set knowledge discovery with enrichr. *Current Protocols*, 1:e90, 3 2021.

[263] Maxim V. Kuleshov, Matthew R. Jones, Andrew D. Rouillard, Nicolas F. Fernandez, Qiaonan Duan, Zichen Wang, Simon Koplev, Sherry L. Jenkins, Kathleen M. Jagodnik, Alexander Lachmann, Michael G. McDermott, Caroline D. Monteiro, Gregory W. Gundersen, and Avi Maayan. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, 44:W90–W97, 7 2016.

[264] Edward Y. Chen, Christopher M. Tan, Yan Kou, Qiaonan Duan, Zichen Wang, Gabriela V. Meirelles, Neil R. Clark, and Avi Ma'ayan. Enrichr: Interactive and collaborative html5 gene list enrichment analysis tool. *BMC Bioinformatics*, 14:1–14, 4 2013.