TITLE: Unified Tumor Growth Mechanisms from Multimodel Inference and Dataset Integration

AUTHORS: Samantha P. Beik[1], Leonard A. Harris[2,3,4], Michael A. Kochen[5], Julien Sage[6,7], Vito Quaranta[8,9], Carlos F. Lopez[9,10]*

ABSTRACT

Systems approaches to elucidate biological processes that impact human health leverage mathematical models encoding mechanistic hypotheses suitable for experimental validation. However, building a single model fit to one dataset may miss alternate equally valid mathematical formulations, and available data may not be sufficient to fully elucidate mechanisms underlying system behavior. Here, we overcome these limitations via a Bayesian multimodel inference (Bayesian-MMI) approach, which estimates how multiple mechanistic hypotheses explain experimental datasets, concurrently quantifying how each dataset informs each hypothesis. We apply this approach to unanswered questions about heterogeneity, lineage plasticity, and cell-cell interaction dynamics in small cell lung cancer (SCLC) tumor growth. Through available dataset integration, we find that Bayesian-MMI predictions support tumor evolution promoted by high lineage plasticity, rather than through expanding rare stem-like populations. These results highlight that given available data, any SCLC cellular subtype can contribute to tumor repopulation post-treatment, suggesting a mechanistic interpretation for tumor recalcitrance.

[1] Medical Scientist Training Program, Vanderbilt University School of Medicine, Nashville, TN, 37232, USA
[2] Department of Biomedical Engineering, University of Arkansas, Fayetteville, AR, USA
[3] Interdisciplinary Graduate Program in Cell & Molecular Biology, University of Arkansas, Fayetteville, AR, USA
[4] Cancer Biology Program, Winthrop P. Rockefeller Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, AR, USA
[5] Department of Bioengineering, University of Washington, Seattle, WA, USA
[6] Departments of Pediatrics, Stanford University, Stanford, CA, USA
[7] Departments of Genetics, Stanford University, Stanford, CA, USA
[8] Program in Chemical and Physical Biology, Vanderbilt University, Nashville, TN, USA
[9] Department of Biochemistry, Vanderbilt University, Nashville, TN, USA
[10] Altos Laboratories, Redwood City, CA, USA
* Correspondence: clopez@altoslabs.com

INTRODUCTION

A mechanistic understanding of cellular processes, explaining causal input-output relationships and predicting population behaviors (Baker et al., 2018) remains a central challenge in biology. Mathematical models provide such quantitative formulations, but two fundamental factors remain poorly addressed toward accurate descriptions of cellular processes. The first concerns model creation, where practitioners must make choices about which model species and interactions should be included or excluded. Often these choices are made in support of models with increased predictive power but narrower scope, at the cost of fundamental insights that could be gleaned from wider-ranging or multiple models (Krakauer et al., 2011). This has led to a deluge of built-to-purpose models, where different models of the same system share some key components but significantly differ in their breadth and resolution (Auffray et al., 2020; Burger et al., 2017; Krakauer et al., 2011; Würstle et al, 2014). The second factor concerns the intrinsic relationship between models and experimental data. The specificity of mechanistic models requires that our assumptions about the system be articulated in precise form (Stumpf, 2020), almost always with much more precision than afforded by available data. To date, the most common practice in the field is to optimize model parameters to a chosen dataset using frequentist minimization or optimization methods and then draw conclusions about biological mechanisms based on a vector of optimized parameter values (Mitra et al., 2019; Villaverde et al., 2014). However, whether the available data can inform a model parameter, or even the model structure, is not routinely considered (Lema-Perez et al., 2019; Raue et al., 2014). Mechanistic representations for large-scale cellular processes, including cell population dynamics and growth modeling, suffer from both issues. Multiple hypotheses and parameter fits are possible to explain the same existing data, decreasing the interpretive power of a model to draw mechanistic, explanatory conclusions about the underlying system.

This need for a rigorous understanding of the relationship between mathematical models and available data in the context of a given cellular process is exacerbated in fields where available data is limited and multiple hypotheses are concurrently proposed. For example, the need to investigate cell population dynamics and the underlying growth and stasis mechanisms therein is high. Tumor heterogeneity is hypothesized to be the major cause of tumor treatment resistance and poor outcome in cancer patients (Marusyk et al., 2012). Numerous experiments and observations in the field in the past few decades led to the cancer stem cell (CSC) paradigm, where rare stem-like populations can regenerate a tumor if even one cell remains after treatment (Batlle and Clevers, 2017; Clarke et al., 2006). In this case, tumor heterogeneity is attributed to the ability of the CSC to self-renew and differentiate into every tumor cell phenotype required for population equilibrium. On the other hand, more recent approaches have indicated movement up and down the hierarchy of differentiation in both physiologically ‘normal’ cells as well as cancer cells (Batlle and

Clevers, 2017), with phenotypic transitions across the stem cell and more differentiated niches (Gupta et al., 2011). Such movement across the hierarchy can occur based on signaling that a cell receives based on its proximity to the stem cell or CSC niche (Batlle and Clevers, 2017), and/or feedback mechanisms maintaining stem and non-stem population equilibrium (Lander et al., 2009).

With these fundamental questions in mind, we study tumor growth mechanisms in small cell lung cancer (SCLC), where relapse after treatment is extremely common and the survival prognosis is poor. SCLC comprises ~15% of all lung cancer cases worldwide and results in ~200,000 deaths annually with a 5-year survival rate of less than 10% (Rudin et al., 2021). Initially, chemotherapy response rates are high at 60-70%, but nearly all patients relapse and succumb to the disease within two years of diagnosis (Rudin et al., 2021). Intratumoral heterogeneity is hypothesized to be the main contributor to the natural history of this disease and its morbidity and mortality (Rudin et al., 2019, 2021; Zeman et al., 2017). SCLC tumors comprise a mix of functionally distinct subtypes of interacting cells (Calbo et al., 2011; Lim et al., 2017; Williamson et al., 2016), but investigations into the role of clonal outgrowth versus phenotypic transitions during development of treatment resistance have only been performed in the past six years (Gay et al., 2021; Polley et al., 2016; Rudin et al, 2019; Udyavar et al., 2017; Wooten et al., 2019). To date, it is recognized that neuroendocrine (NE) SCLC subtypes transition to non-neuroendocrine (Non-NE) SCLC subtypes over time or upon treatment (Gay et al., 2021; Ireland et al., 2020; Lim et al., 2017; Stewart et al., 2020), but the role of each individually recognized SCLC subtype, studied in varying experimental models (Rudin et al., 2019) has not been elucidated during the tumor-growth process. Multiple groups have proposed nonoverlapping models of tumor growth mechanisms, leading to variable conclusions specific to the experimental model used (Gay et al., 2021; Ireland et al., 2020; Jahchan et al., 2016; Lim et al., 2017; Mollaoglu et al., 2017; Stewart et al., 2020). Notably, the contribution of CSC-like behavior in a hierarchical phenotypic heterogeneity perspective vs. tumor cell plasticity in a less hierarchical framework has not been fully explored or elucidated.

To attain our goals, we introduce a framework that leverages Bayesian model selection, enabling the researcher to assess the inclusion or exclusion of species and behaviors in a mechanistic model, and Bayesian principles, revealing the extent to which data can inform model parameters. We integrate the most suitable available datasets and published theories of SCLC cellular biology to enumerate mechanistic hypotheses for SCLC tumor growth, resulting in thousands of candidate population dynamics models to test via nested sampling. Fitting all candidate mathematical models to tumor steady-state data, we apply the principles of model selection and model averaging for a principled and comprehensive assessment of SCLC mechanistic hypotheses, including tumor initiation, tumor subtype composition, cell-cell interactions, and phenotypic transitions within and outside of a differentiation hierarchy. We estimate the likelihood 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 systems, and as such it represents a paradigm shift from the common practice to formulate *one* mechanistic model of a cellular process and instead propose a move toward a probabilistic understanding of cellular behaviors.

RESULTS

SCLC tumor growth data integration and subtype deconvolution

We first gathered data to build a mechanistic hypothesis exploration space for SCLC tumor growth dynamics. We use multiple datasets (Barretina et al., 2012; Lim et al., 2017; Mollaoglu et al., 2017; Wooten et al., 2019) with the goal of identifying consensus behavior of SCLC and providing a unifying model supported by SCLC data more broadly. Throughout this manuscript, we refer to experimentally observed SCLC phenotypic subtypes as "subtypes". Each subtype comprises a subpopulation within a tumor identified by differential expression of transcriptional regulators. As shown schematically in **Figure 1A**, subtypes broadly fall under one of two characterizations, neuroendocrine (NE) or non-neuroendocrine (Non-NE) (**Figure 1A(1)**). Subtypes within the NE category include SCLC-A, positive for Achaete Scute Homologue 1 (ASCL1), and SCLC-N, positive for neurogenic differentiation factor 1 (NeuroD1), while Non-NE subtypes include SCLC-P, positive for POU class 2 homeobox 3 (POU2F3), and SCLC-Y, positive for Yes-associated protein 1 (YAP1) (Rudin et al., 2019) (**Figure 1A(3)**). Some of these subtypes have been experimentally verified to coexist within a tumor population, while others have only been identified in isolation (Baine et al., 2020; Gay et al., 2021; Ireland et al., 2020; Lim et al., 2017). We have also determined the existence of a fifth subtype, itself a subcategory of SCLC-A, thus designated SCLC-A2 (**Figure 1A(4)**). SCLC-A2 is positive for ASCL1 but also Hes1, a Notch pathway target that typically represses NE differentiation (Wooten et al., 2019). Other experiments have led to additional proposed phenotypic subtypes (Chan et al., 2021; Simpson et al., 2020; Gay et al., 2021; Williamson et al., 2016), but for this investigation we focus on canonical SCLC subtypes identified in (Rudin et al., 2019). It has been shown that Hes1-negative (NE) cells may undergo a transition toward Hes1-positive (Non-NE) identities, (Lim et al., 2017) (**Figure 1A(1)**). Studies *in vitro* and *in vivo* have suggested that Non-NE subtype(s) support growth of NE subtypes (Lim et al., 2017) (**Figure 1A(2)**), including vasculogenic

mimicking SCLC cells having such supportive effects (Williamson et al., 2016), while that the presence of NE subtypes has a dampening effect on Non-NE growth (Mollaoglu et al., 2017). In separate model systems, SCLC-A subtype cells have been shown to transition to the SCLC-N subtype and from SCLC-N to SCLC-Y (Ireland et al., 2020) (**Figure 1A(5)**). It has recently been shown that Hes1-positive (Non-NE) cells have upregulated YAP1 (Shue et al., 2022) (**Figure 1A(2)**), and as such could be considered the SCLC-Y subtype. These conclusions and remaining questions are summarized in **Table 1**. For a comprehensive review of SCLC literature, we refer the reader to (Rudin et al., 2019).

We gathered RNA-sequencing data from experiments used to draw these conclusions to explore whether a unifying mechanistic model of SCLC tumor growth could be formulated. These datasets include two genetically-engineered mouse models (GEMMs), the triple-knockout (TKO) model ($p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017)), and the RPM model ($Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-Myc^{T58A} (Mollaoglu et al., 2017)). Data from each model provide multiple replicates of tumors from two genetic backgrounds and two unique oncogenic processes: replicates of data from the same system are a requirement for the Bayesian model-fitting process (see **Methods**). We also used publicly available SCLC human cell line data, a subset of the cell lines in the Cancer Cell Line Encyclopedia (CCLE) (Barretina et al., 2012), as these provide the more translationally relevant representation of SCLC. Originally derived from human tumors, each cell line has a different genetic background and different oncogenic history, and therefore we have only one replicate per CCLE cell line. To approximate multiple replicates of an SCLC tumor system, we use cell lines of similar steady state composition (**Figure 1B,** right). We previously grouped publicly available SCLC cell line data into clusters that align with the different subtypes, and for the analysis herein we use cell lines made up largely of the SCLC-A subtype determined in (Wooten et al., 2019). We previously used CIBERSORT (Newman et al., 2015) to deconvolute the RNA sequencing data of these GEMMs and majority-SCLC-A cell lines and assign subtype populations (**Figure 1B**; **File S1**). The different datasets represented in **Figure 1B** demonstrate differing SCLC tumor makeup, according to the experimental model employed in the study.

Population dynamics modeling of SCLC to explore biological hypotheses

In order 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 (Harris et al., 2019). Population dynamics models employ a mathematical description of the dynamics within and between heterogeneous subpopulations in an overall population (Charlebois

and Balázsi, 2019; Metzcar et al., 2019). With such models, researchers can mathematically simulate population growth over time and investigate growth dynamics inherent in the simulations (**Box 1**).

In our attempt to build a mathematical model of SCLC that could unify nonoverlapping conclusions and hypotheses in the field (**Figure 1**, **Table 1**), we came across several levels of bias. For example, phenotypic transitions were studied in Ireland et al. (2020), but not cell-cell interactions, which themselves were studied in Lim et al. (2017). In addition, to build one mathematical model it would be necessary to choose one model topology (subtype makeup of the tumor population) as well as an initiating subtype (cell(s) of tumor origin), which have not been demonstrated experimentally but have been hypothesized in Wooten et al. (2019), Ireland et al. (2020), and Ouadah et al. (2019). A common approach to address such a problem is to informally choose one interpretation of existing data to build the model, such as combining conclusions from separate investigations, or selecting one out of several possible conclusions if findings in the literature conflict. These are legitimate model-building methods, but these methods introduce bias because other interpretations may be plausible. To avoid the bias that comes from selecting a single set of interpretations on which to build a mathematical model, we instead combine Bayesian statistics and multimodel inference (Burnham and Anderson, 2002; Fragoso et al., 2017), for which we could develop multiple candidate models as part of the investigative process. We developed these candidates using HypBuilder, a tool for generating mechanistic hypotheses of kinetic processes (see **Methods**), and PySB, a modeling framework for dynamic biochemical processes (Lopez et al., 2013). These multiple model candidates comprise between 9 and 13 actions and interactions per cell subtype, with an action or interaction available to a subtype in the model if it has been shown or hypothesized in the literature (**Table 1**).

We interpret outstanding questions in the SCLC field - tumor topology, initiation, potential subtype behaviors - as features (model variables) in candidate models (**Table 1**). In the following, we define *model variables* as representations of qualitative action in the model (e.g., "subtype A cell division") whose rate is denoted by a *kinetic parameter* (e.g., "subtype A division rate"). To fully account for possibilities of tumor subtype makeup, or topology, we included all potential cell subtype combinations (A, A2, N, Y) with between two and four subtypes in each model (**Figure 2A, Figure S1G**). While the SCLC-P subtype is known to exist (**Figure 1A(3)**), the data used for our analyses does not include subtype P (**Figure 1B**). Growth supportive effects and transition-inducing effects (**Figure 2B**) (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 (Wooten et al., 2019), 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 2B**, left).

To compare a hierarchical system, where a CSC can (re)populate a tumor, and non-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 (Ireland et al., 2020; Lim et al., 2017), and models with reversible transitions, i.e., high plasticity (**Figure 2C**). 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 2C**, **Figure S1G**). 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.

To explore the space of tumor growth mechanistic hypotheses, we employ multimodel inference (MMI) in the context of Bayesian statistics (**Box 2**). There have been biological investigations using MMI approaches (Lu et al., 2021; van de Schoot et al., 2021) but, to our knowledge, this work is the first application of Bayesian MMI to cell population dynamics models. MMI using information theoretic methods (Burnham and Anderson, 2002) is subject to issues we attempt to mitigate using our Bayesian approach, as described in **Box 3**. Our application of MMI includes assessing the posterior probability of hypotheses from **Table 1**, denoted as different model variables in our candidate models. This use of MMI 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 (van de Schoot et al., 2021). We apply a Bayesian approach to MMI here to assess model variables that would be mechanistically meaningful for tumor growth dynamics, and whether these are likely to be representative of the SCLC system based on the data. This assessment of population dynamics models, and specifically the Bayesian probability that biologically relevant model features are supported by the data, is 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.

Table 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 (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017)

Tumor composition
• Neuroendocrine-classified (NE) subtypes are A (ASCL1+) and N (NEUROD1+), with non-neuroendocrine-classified (Non-NE) subtype Y (YAP1+) and P (POU2F3+).[a]
• Previous work predicted subtype A2, and SCLC-A, SCLC-A2 and SCLC-N have been seen to express ASCL1.[a,b,c,d] SCLC-A2 expresses Hes1.[d]
• Hes1-positive TKO tumor cells (Non-NE) have YAP1 upregulated. [e]
• 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.[a,d]
• Some subtype combinations have been experimentally verified and others have been predicted using CIBERSORT.[a,c,d]
• We do not see SCLC-P in our subtype deconvolution. [d]
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. [b]
• *Ex vivo* culturing of *in situ* RPM tumors results in histologic and transcriptional phenotypic changes from NE to Non-NE gene expression over time. [c]
• Transitions between A and A2, A2 and N have not been studied.
• No evidence of SCLC-Y transition to NE identity.[b]
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. [b]
• NE cells suppress Non-NE cell division. [f]
• Application of conditioned media or of isolated exosomes results in a morphological change in TKO-derived NE cell line KP3.*

a) Rudin CM, et al. *Nat. Rev. Cancer*. 2019

b) Lim, J et al. *Nature.* 2017

c) Ireland, A et al. *Cancer Cell.* 2020

d) Wooten, Groves et al. *PLoS Comput. Biol.* 2019

e) Shue, YT et al. *Nat. Commun*. 2022

f) Jahchan NS, et al. *Cell Rep.* 2016

*) *personal communication (Alissa Weaver, Vanderbilt University)*

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

To assess how many of the 5,891 candidate models are supported by experimental data, we performed model fitting to data and calculation of marginal likelihood using nested sampling (Feroz and Hobson,

2008; Feroz et al., 2009, 2019; Skilling, 2004). We use nested sampling for model selection because the marginal likelihood, which nested sampling calculates, enables us to carry out a rigorous Bayesian analysis (see **Box 3**). 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 interpretations of tumor growth mechanisms. The average fitting time for each model was ~19 wall-clock hours, thus necessitating high-performance computing for a complete parameter space exploration of the candidate models.

To ensure that we are exhaustively exploring a biologically relevant posterior probability space, we incorporated prior knowledge about mechanistic behavior of tumor populations (Chen et al., 2016; Fuentes-Garí et al., 2015; Greene et al., 2015; Gupta et al., 2011; Wang et al., 2014; Zhou et al., 2014) and cellular behaviors likely to exist in SCLC (Ireland et al., 2020; Jahchan et al., 2016; Lim et al., 2017; Mollaoglu et al., 2017). We therefore expect that all plausible SCLC tumor growth mechanisms are represented in our 5,891-model hypothesis space to the best of our knowledge (see **Box 2** for a description of Bayesian statistics in the context of MMI).

In a Bayesian model selection approach, a more likely model comprises a higher proportion of the probability of the candidate model space (**Figure 3A-C**). After nested sampling, our results indicate the highest-scoring model for each dataset is ~10^{19} times more likely than the lowest-scoring model, and ~10^{3} times more likely than the median scoring model. For reference, the smallest comparison between models that is considered significant is $10^{1/2}$ (Kass and Raftery, 1995) (see **Methods**). 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 (**Figure 3A-C**). We therefore turned to 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, (**Figure 4A-C**, orange) we also calculate a "relative likelihood confidence interval," as discussed in Burnham and Anderson (2002) (see **Methods**). 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 > $10^{1/2}$ (see above). 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 (~1% or less, **Figure 3D**).

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

variable, we wanted to investigate how the data available can inform model variables. If data does not inform a model variable and its 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.

Most-likely 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 depart from the traditional single-model mechanism approach, and instead employ 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 (Hinne et al., 2020). In BMA, each model is weighted by its posterior model evidence (Fragoso et al., 2017) and the model variables within each model receive an averaged likelihood (van de Schoot et al., 2021) (see **Boxes 2, 3**).

However, before we could investigate model-averaged probabilities, we found that the fitted parameter distribution outcome was dependent on choice of initiating subtype (**Figure S4**). We, therefore, turned to the literature to impose stricter constraints about initial subtype conditions, to investigate model variable probabilities when kinetic parameter posterior distributions were constrained. Reports link NE SCLC subtypes and long-term tumor propagation (Lim et al., 2017; Jahchan et al., 2016) and, in particular, cells of subtype A (Ireland et al., 2020). We thus used only candidate models those with a starting 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 4A**; model topology probabilities without filtering by initiating subtype are shown in **Figure S5A**).

We perform BMA across all models for each dataset. As shown in **Figure 4A**, 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 (Skilling, 2004). 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 4A**). 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 GEM (**Figure 1B**). However, the SCLC-A cell line data also has probability that falls in the A, N, and A2 topology (model 4 in **Figure 4A**) – 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.

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

After establishing 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 4A**, blue) for the TKO GEMM data, along with the four-subtype topology, are compared in **Figure 4B** (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 subpopulation in the tumor that is longer-lived, and will only slowly transition to Y. In the topology with A, A2, and Y (**Figure 4A**, 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 S5B**), suggesting that N is a shorter-lived intermediate in the A-to-N-to-Y transition. This result aligns with previous experiments (Ireland et al., 2020) 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 4A**, red) for RPM-fitted models, as well as the four-subtype topology (**Figure 4B**, 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 S5B**). 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 4A**, green) for the SCLC-A cell line data and the four-subtype topology (**Figure 4B**, 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.

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 BMA and MMI to work toward a unifying model of SCLC tumor growth based on multiple datasets.

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

We have considered candidate models (**Figure 3**), model topologies (**Figure 4A**), and kinetic parameters (**Figure 4B**) to explore tumor growth mechanisms in SCLC. There is compelling experimental evidence for multi-subtype tumor composition, which implies multiple potential growth mechanisms (Gay et al., 2021; Ireland et al., 2020; Lim et al., 2017; Mollaoglu et al., 2017; Stewart et al., 2020). We therefore focused on model topologies 1, 2, and 4, which are three-subtype topologies with detectable probability (> 1%) (**Figure 4A**), 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 (see **Table 1**). 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 variable (see **Box 2** and **Methods**). Despite different posterior probability values (**Figure 5A**), the probabilities of model variables across datasets were similar in their trends: across all three-subtype topologies, phenotypic transition probabilities were all more than $^1/_2$ (**Figure 5A**, red squares). While some probability values were poorly informed (light red), (probability between $^1/_2$ and $^2/_3$), more were informed by the data (deep red) ($^2/_3$ or more). Conversely, probabilities of Non-NE effects on the growth or transitions were all less than $^1/_2$ (**Figure 5A**, blue squares). Some probability values were poorly informed, (light blue) (between $^1/_3$ and $^1/_2$) and others were informed (deep blue) ($^1/_3$ or less) with the addition of data.

Table 2. Probabilities after hypothesis exploration using Bayesian multimodel inference. 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 (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) representative of SCLC-A tumors (Wooten et al., 2019))

<u>Informed high posterior probabilities</u>

- 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: ~1.5 to 4.
- Transitions between N and A2 (N-to-A2, A2-to-N) had posterior probability 66%
- Posterior odds: ~2

<u>Informed low posterior probabilities</u>

- 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: ~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.

Overall, the data suggests that Non-NE effects on transition rates of N-to-Y, or A2-to-Y, are unlikely, (**Figure 5B-D**, deep blue) regardless of whether "Non-NE" defines only the Y subtype, or both A2 and Y are Non-NE (**Figure 2B**). 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 5A**), 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 (**Figure 5B-D**, 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 S1G**) 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 variables, 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 (**Figure 5B-D**, light blue, light red).

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 5E**). Briefly, phenotypic transition probabilities were chosen from the models least informed by the data in an attempt to make conservative predictions (see **Methods**). Model-averaged parameter rates were visually compared (**Figure 5F**) 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 4B**). Values from the consolidated probabilities (**Figure 5E**) are those reported in **Table 2**, along with posterior odds to compare one hypothesis to its opposite.

Taken together, these results provide insight not only into what model 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 5E**). 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.

DISCUSSION

In this work, we integrated conclusions, theories, and hypotheses from SCLC literature to develop aspects of a population dynamics model – initiating subtype, tumor subtype makeup, subtype phenotypic transitions, and cell-cell interactions affecting growth and transitions (**Figure 1, Table 1**) – that we could then view as mechanistic hypotheses to include or exclude from thousands of candidate models (**Figure 2**) to which to apply multimodel inference. Using nested sampling and model selection, we fit all candidate models to data and determined the marginal likelihood that each model best fits the data, which we then converted to posterior probability (**Figure 3**). Using model averaging, we aggregated the posterior probabilities of groups of candidate models representing single hypotheses. Investigating candidate model topologies, we determined that varying two- and three-subtype topologies were optimal for representing each dataset (**Figure 4a**), and assessing growth and transition rates across these topologies, determined that the presence of N and A2 altered phenotypic transitions rates (**Figure 4b**). Estimating the likelihood of each mechanistic hypothesis given the data, we saw overall that all datasets support the high likelihood of non-hierarchical phenotypic transitions, indicating SCLC subtype plasticity, and less likely cell-cell interaction that affect these transitions (**Figure 5**). Comparing between prior and posterior probabilities, we saw that while aspects of the model such as phenotypic transitions are well informed by the data, other aspects, such as tumor initiation and growth rate effects, are not informed by the data.

Here we focus on the aspects of growth in the SCLC primary tumor, and thus do not address invasion or metastasis. The invasive or metastatic potential of the SCLC tumor is known to be increased by Non-NE subtypes (Kwon et al., 2015; Williamson et al., 2016), which could be included in the model along with other aspects of tumor invasion. 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 are

limited in our ability to use RNA sequencing subtyping data from primary human tumors, an important detail in the context of recent studies providing such data (Chan et al., 2021; Simpson et al., 2020; Stanzione et al., 2022; Stewart et al., 2020). We focus on non-human GEMM data due to the Bayesian model-fitting process: optimal fitting requires multiple biological replicates to provide a distribution of data points that account for stochastic variation in the system. Samples from each GEMM used here represent replicates of tumors from identical backgrounds, since each tumor stems from the same genetic modification leading to oncogenesis. We cannot use any human data similarly, as each tumor has a unique background. We approximate multiple human tumor replicates via cell lines of similar steady state composition (**Figure 1B**), and in this way consider that we have attained some translational relevance. Future work requires assessment of which patient primary samples may approximate replicates and whose data can thus be used for fitting our model(s) using the paradigm described herein. With the wealth of primary human tumor data now available, additional translational relevance and fitting our models to this data is an ongoing effort.

We note that our candidate models only included encoded cell-cell interactions between subtype Y and others, or between subtypes A2 and Y together and others (**Figure 2B**), while our analysis comparing rates across model topologies suggest that cell-cell interactions may occur between N and/or A2 and other subtypes (**Figure 4B**). Interestingly, while we find that N or A2 cell-cell interactions may affect the rate of phenotypic transitions (**Figure 4B**), the likelihood of Y or A2 and Y affecting phenotypic transitions is low (**Figure 5A-D**). Cell-cell interactions where N or A2 affect tumor growth dynamics have not previously been studied and we did not include them among candidate models, but our analysis predicts that these may be interactions to further probe experimentally.

The experimental data used for this analysis supports two-subtype topologies as higher-probability candidates. This is not surprising, because nested sampling prioritizes simpler models. However, mounting experimental evidence supports multi-subtype tumor composition (Gay et al., 2021; Ireland et al., 2020; Lim et al., 2017; Mollaoglu et al., 2017; Stewart et al., 2020). In fact, previous work from our labs suggests that a tumor's 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 (Harris et al., 2019; Hayford et al., 2021; Paudel et al., 2018). We therefore hypothesize that with additional data, such as timecourse or perturbation data, the topology likelihood may be shifted toward three- and 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 (Chaffer et al., 2011; Gupta et al., 2011; Pisco et al., 2013; Yang et al., 2012) and changes in the proportion of phenotypic subpopulations after a perturbation. Often this perturbation is the application of drug treatment (Hata et al., 2016; Ku et al.,

2017; Mu et al., 2017; Risom et al., 2018; Shaffer et al., 2018; Zou et al., 2017), but may also be changes in microenvironment or related factors (Ireland et al., 2020; Watson et al., 2018).

The results presented here provide strong evidence for phenotypic plasticity in SCLC tumors, based on the higher likelihood for most phenotypic transitions tested, regardless of differentiation hierarchy. Such plasticity and ability to undergo any phenotypic transition align less with the CSC paradigm, where only rare cells have the capacity to undergo transitions toward all subtypes, and more with theories supporting lineage plasticity where most cell types may undergo phenotypic transitions to maintain population equilibrium, not only in cancer but in physiological tissues as well. We hypothesize that the ability to reach different phenotypic states other than those apparent in an untreated tumor sample is a reason for the treatment resistance seen in nearly all SCLC cases. With a more plastic and less CSC-based phenotypic equilibrium, instead of rare remaining stem-like populations leading to regenerate a tumor after treatment, 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, here we make headway by predicting that plasticity is highly likely, and extremely important in the growth and evolution of the SCLC tumor.

With movement across the differentiation hierarchy (Batlle and Clevers, 2017) due to phenotypic transitions likely based on signaling and/or feedback mechanisms maintaining population equilibrium (Lander et al., 2009), additional data from perturbing the experimental models used in this analysis and time-course experiments of increase, decrease, and regulation of tumor subpopulations will be extremely important to collect. It is of particular interest as well 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 may be 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.

We believe a paradigm shift is needed for the field of modeling in systems biology, moving from single, built-to-purpose models to a more unbiased approach, investigating the optimal model(s) with identifiable parameters for a particular dataset. This can help steer investigators away from informal reasoning that may lead to incomplete models of mechanistic processes. 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. Not only can hypotheses be compared in terms of their likelihoods, but the change from prior to posterior probability represents what aspects of a

system the data at hand is able to inform. Understanding which hypotheses the data informs 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.

METHODS

CIBERSORT deconvolution of RNA sequencing data

Data from two GEMM models provide multiple replicates of tumors from two genetic backgrounds: one from $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ (triple-knockout, or TKO) GEMM tumors (Lim et al., 2017), and another from $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-Myc^{T58A} (RPM) GEMM tumors (Mollaoglu et al., 2017). We also used publicly available SCLC cell line data as these provide the more translationally relevant representation of human tumors. 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 Wooten et al. (2019), 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. Results used for this publication can be found as **File S1**.

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. PySB is a rule-based modeling language, where one will encode

$$A \xrightarrow{k} 2A, \quad k = 0.469 \tag{1}$$

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

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

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 S1**), 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 are then provided to the Multinest algorithm to perform nested sampling.

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 automatically 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 {Github link} 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 (Burnham and Anderson, 2002). However, prior knowledge enables us to determine that all 5,891 models represent all possibilities with regard to outstanding SCLC 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 S2**). Correspondingly, a probabilistic representation of best-fitting rates for each model is returned by the Multinest algorithm (posterior marginal distribution) (e.g., **Figure 4B, 5E**; **Figure S4, S5B**).

Parameter estimation and evidence calculation by nested sampling

As noted in Eqs. (1) and (2), 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, (Mitra et al., 2019; van Ravenzwaaij et al., 2018) with Bayesian methods utilizing a prior rate parameter distribution, $P(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ represents the set of n parameters $\{\theta_1, \theta_2, \dots, \theta_n\}$, and a likelihood function to assess a parameter set

$$\mathcal{L}(\boldsymbol{\theta}_i) = P(D|\boldsymbol{\theta}_i), \tag{3}$$

where $\boldsymbol{\theta}_i$ is the i^{th} parameter set and D represents the data being used for fitting. Parameter set $\boldsymbol{\theta}_i$ is scored via the likelihood function $\mathcal{L}(\boldsymbol{\theta}_i)$ and optimization continues, moving toward better-scoring parameter sets until an optimal score is reached.

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, $P(\boldsymbol{\theta})$, and a likelihood (Eq. 3) the posterior probability can be calculated via Bayes' Theorem,

$$P(\boldsymbol{\theta}_i|D) = \frac{P(D|\boldsymbol{\theta}_i)P(\boldsymbol{\theta}_i)}{\int_{\boldsymbol{\theta}} P(D|\boldsymbol{\theta})P(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{4}$$

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 (Skilling, 2004).

To perform nested sampling, we utilize the Multinest algorithm (Feroz and Hobson, 2008; Feroz et al., 2009, 2019). 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, $P(\boldsymbol{\theta})$, 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 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 (Eq. 3). 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 importance nested sampling evidence value, because multiple importance nested sampling runs (multiple 'replicates') for the same candidate model and prior parameters returned more consistent evidence values than 'vanilla' nested sampling (data not shown) (Feroz et al., 2019).

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 (Kruschke, 2011):

$$\nu = \alpha + \beta = \frac{\mu(1-\mu)}{\sigma^2} - 1, \text{where } \nu = (\alpha+\beta) > 0 \text{ (meaning } \sigma^2 < \mu(1-\mu)) \tag{5}$$

$$\alpha = \mu\nu = \mu\left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right), \text{if } \sigma^2 < \mu(1-\mu) \tag{6}$$

$$\beta = (1-\mu)\nu = (1-\mu)\left(\frac{\mu(1-\mu)}{\sigma^2} - 1\right), \text{if } \sigma^2 < \mu(1-\mu). \tag{7}$$

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

$$\sum_{s}^{S-1} \begin{cases} log\left(\frac{x_{s_{data}}^{\alpha_{s_{sim}}-1}\left(1-x_{s_{data}}\right)^{\beta_{s_{sim}}-1}}{\mathrm{B}(\alpha,\beta)}\right) & \text{if } s \in \text{ simulated model topology} \\ log\left(\lambda e^{-\lambda x_{s_{data}}}\right), \text{where } \lambda = \frac{1}{\sigma_{s_{data}}}, & \text{otherwise} \end{cases} \tag{8}$$

where $\mathcal{S}$ is the set of subtypes, $\alpha_{s_{sim}}$ and $\beta_{s_{sim}}$ indicate the calculation of α (Eq. 6) and β (Eq. 7) 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, $\mathrm{B}(\alpha, \beta)$ 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 ($\log(\lambda e^{-\lambda x_{s_{data}}})$) as part of the log-likelihood (Eq. 8) 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. 8, above) and the exponential log likelihood (Eq. 8, 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 (~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 (**Figures 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.

Candidate model prior and posterior probabilities and confidence interval calculation

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 underlying SCLC tumor system,

$$P(M_k) = \frac{1}{|\mathcal{M}|}, \tag{9}$$

Where $\mathcal{M}$ is the set of all candidate models. With the model evidence, or marginal likelihood, $P(D|M_k)$ estimated by Multinest, (Feroz and Hobson, 2008; Feroz et al., 2009, 2019) the posterior probability per model can be calculated as

$$P(M_k|D) = \frac{P(D|M_k)P(M_k)}{\sum_j P(D|M_j)P(M_j)}. \tag{10}$$

With a posterior probability per model, (**Figure 3**) 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 (Burnham and Anderson, 2002) (**Figure 3**, orange). Using this confidence interval results in ~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 (Burnham and Anderson, 2002). 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 $\frac{P(D|M_{highest})}{P(D|M_i)} >$ cutoff. Burnham and Anderson (Burnham and Anderson, 2002) denote such a method as a "relative likelihood confidence interval" and discuss its support by statistical theory (Azzalini, 1996; Berger and Wolpert, 1984; Edwards, 1992; Royall, 1997), noting that it is uncommonly found in the model selection literature. We used a cutoff of $10^{1/2}$, the lowest Bayes Factor at which a difference may be determined (Kass and Raftery, 1995). Even with this permissive cutoff, the relative likelihood CI includes only tens of models, an even greater decrease from the initial number of candidates.

Prior and posterior probabilities per hypothesis being investigated

Each hypothesis has an assigned prior probability based on our prior expectations. For all hypotheses, we took an approach where we considered each hypothesis as equally likely compared to competing hypotheses. For the inclusion of most model variables, this was a prior probability of 50%, where it is 50% likely the model variable is part of a model that is the best representation of the tumor system, and 50% likely that same variable 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 S2-S4**, 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 4A**). For model initiating subtype hypotheses, (**Figure S1G**) with 15 potential combinations of initiating subtypes, each initiating subtype combination has a 6.67% prior probability.

Each candidate model can then be assigned a prior probability conditional on the hypothesis being considered, $P(M_k|H_i)$, where M_k is the kth candidate model and H_i is the hypothesis being considered. The

calculation of $P(M_k|H_i)$ is based on the number of candidate models that fall under the hypothesis being considered,

$$P(M_k|H_i) = \frac{P(H_i)}{|\{M_j \in M, H_i\}|}, \quad (11)$$

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 the model variable 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 variable.

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

$$P(M_k|D, H_i) = \frac{P(D|M_k,H_i)P(M_k|H_i)}{\sum_j P(D|M_j,H_i)P(M_j|H_i)}, \quad (12)$$

Where $P(D|M_k, H_i)$ is the Bayesian model evidence (marginal likelihood) for $Model_k$.

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

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{\sum_j P(D|H_j)P(H_j)}, \quad (13)$$

Where $P(D|H_i)$ is the marginal likelihood of H_i over all models to which it applies, $\{M_j \in M, H_i\}$. According to (Kass & Raftery, 1994), this can be calculated as

$$P(D|H_i) = \sum_k P(D|M_k, H_i)P(M_k|H_i), \quad (14)$$

with a summation instead of an integral because each model has a discrete prior probability as calculated in Eq. (11).

Using the results of Eq. (14) in Eq. (13), we then calculate the posterior probability for each hypothesis, pictured in **Figure 5A-D** and noted in **Table 2**. In this way, we can use Bayesian calculation rather than parameter importance analysis (Galipaud et al., 2014; 2017) to determine the posterior probability of each model variable. This also enables us to avoid bias in considering models with and without certain model variables, if an uneven number of candidate models contain a model variable *vs.* do not contain the variable (Garthwaite and Mubwandarikwa, 2010).

Posterior odds per hypothesis being investigated

All model variables begin with a prior probability of 0.5. With equal prior probabilities across all model variable hypotheses, the posterior odds represented by $\frac{posterior\ probability}{(1-posterior\ probability)}$ is equivalent to the Bayes Factor (see **Box 2**). Therefore, calculation of the posterior odds and the Bayes Factors for each model variable are equivalent.

A posterior probability of model variable inclusion 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 (Kass and Raftery, 1995). 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 variable considered. We consider probabilities between 0.333 and 0.667 to not have been significantly informed by the data.

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) (**Figure 4B, 5E**; **Figure S4, S5B**). Since each candidate model is assigned a posterior probability as in Equation (11), 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.

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.

Generating a consolidated model of the SCLC tumor

A hypothesis (model variable) 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 5E**). 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 variables 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.

CONTRIBUTIONS

Conceptualization, S.P.B, L.A.H., and C.F.L; methodology, S.P.B., L.A.H., and M.A.K.; software, S.P.B., L.A.H., and M.A.K.; formal analysis and investigation, S.P.B.; resources, C.F.L., J.S. and V.Q.; data curation, S.P.B.; writing – original draft, S.P.B., C.F.L.; writing – review & editing, S.P.B., L.A.H., M.A.K., J.S., V.Q., C.F.L.; visualization, S.P.B.; supervision, V.Q., L.A.H., and C.F.L; funding acquisition, V.Q. and C.F.L.

ACKNOWLEDGMENTS

The authors would like to thank Sarah Groves, Michael Irvin, and Christine Lovly for insightful conversations and critical feedback on this work. This work was supported by the following funding sources: S.P.B. was supported by the National Institutes of Health (NIH) [T32GM007347 and T32LM012412] and the National Cancer Institute (NCI) [F30CA247078]; C.F.L. was supported by the National Science Foundation (NSF) [MCB 1411482] and NSF CAREER Award [MCB 1942255]; C.F.L. and V.Q. were supported by the National Institutes of Health (NIH) [U54-CA217450 and U01-CA215845].

AVAILABILITY

All the code to reproduce the data analysis and the figures is open source and can be found in this GitHub repository: https://github.com/LoLab-VU/Bayes-MMI. This repository contains all the source code for each step in the analysis.

Files needed to reproduce the figures without running the data analysis prior can be found in this DropBox folder: https://www.dropbox.com/sh/4fqzpvu9hgyjicm/AABdfFlCenEuiOPgiH0TT-xqa?dl=0. Download of these files and placement in the corresponding directories in the GitHub repository enables reproduction of Figures 3-5.

REFERENCES

Auffray, C., Noble, D., Nottale, L., & Turner, P. (2020). Progress in integrative systems biology, physiology and medicine: towards a scale-relative biology. *Eur. Phys. J. A*, *56*, 88. https://doi.org/10.1140/epja/s10050-020-00090-3

Azzalini, Adelchi. (1996). *Statistical inference : based on the likelihood* (Chapman and Hall/CRC)

Baine, M. K., Hsieh, M. S., Lai, W. V., Egger, J. v., Jungbluth, A. A., Daneshbod, Y., Beras, A., Spencer, R., Lopardo, J., Bodd, F., Montecalvo, J., Sauter, J. L., Chang, J. C., Buonocore, D. J., Travis, W. D., Sen, T., Poirier, J. T., Rudin, C. M., & Rekhtman, N. (2020). SCLC Subtypes Defined by ASCL1, NEUROD1, POU2F3, and YAP1: A Comprehensive Immunohistochemical and Histopathologic Characterization. *J. Thorac. Oncol. 15*(12), 1823–1835. https://doi.org/10.1016/j.jtho.2020.09.009

Baker, R. E., Peña, J. M., Jayamohan, J., & Jérusalem, A. (2018). Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biol. Lett.*, *14*: 20170660. https://doi.org/10.1098/RSBL.2017.0660

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., Wilson, C. J., Lehár, J., Kryukov, G. v., Sonkin, D., Reddy, A., Liu, M., Murray, L., Berger, M. F., Monahan, J. E., Morais, P., Meltzer, J., Korejwa, A., Jané-Valbuena, J., … Garraway, L. A. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature 483*(7391), 603–607. https://doi.org/10.1038/nature11003

Batlle, E., & Clevers, H. (2017). Cancer stem cells revisited. *Nat. Med. 23*(10), 1124–1134. https://doi.org/10.1038/nm.4409

Berger, J.O. and Wolpert, R.L. (1984) The Likelihood Principle. Lecture Notes, Monograph Series, Vol. 6. (Institute of Mathematical Statistics, Hayward)

Burger, G. A., Danen, E. H. J., & Beltman, J. B. (2017). Deciphering epithelial-Mesenchymal Transition Regulatory Networks in Cancer through Computational Approaches. *Front. Oncol.*, *7*:162. https://doi.org/10.3389/fonc.2017.00162

Burnham, K. P., & Anderson, D. R. (2002). Model Selection and Multimodel Inference: a Practical Information-theoretic Approach. In 2nd ed. (Springer, New York)

Calbo, J., van Montfort, E., Proost, N., van Drunen, E., Beverloo, H. B., Meuwissen, R., & Berns, A. (2011). A Functional Role for Tumor Cell Heterogeneity in a Mouse Model of Small Cell Lung Cancer. *Cancer Cell*, *19*(2), 244–256. https://doi.org/10.1016/J.CCR.2010.12.021

Chaffer, C. L., Brueckmann, I., Scheel, C., Kaestli, A. J., Wiggins, P. A., Rodrigues, L. O., Brooks, M., Reinhardt, F., Suc, Y., Polyak, K., Arendt, L. M., Kuperwasser, C., Bierie, B., & Weinberg, R. A. (2011). Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proc. Natl. Acad. Sci. U.S.A.*, *108*(19), 7950–7955. https://doi.org/10.1073/pnas.1102454108

Chan, J. M., Quintanal-Villalonga, Á., Gao, V. R., Xie, Y., Allaj, V., Chaudhary, O., Masilionis, I., Egger, J., Chow, A., Walle, T., Mattar, M., Yarlagadda, D. V. K., Wang, J. L., Uddin, F., Offin, M., Ciampricotti, M., Qeriqi, B., Bahr, A., de Stanchina, E., … Rudin, C. M. (2021). Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell*, *39*(11), 1479-1496.e18. https://doi.org/10.1016/J.CCELL.2021.09.008

Charlebois, D. A., & Balázsi, G. (2019). Modeling cell population dynamics. *In Silico Biol. 13*(1–2), 21–39. https://doi.org/10.3233/ISB-180470

Chen, X., Wang, Y., Feng, T., Yi, M., Zhang, X., & Zhou, D. (2016). The overshoot and phenotypic equilibrium in characterizing cancer dynamics of reversible phenotypic plasticity. *J. Theor. Biol.*, *390*, 40–49. https://doi.org/10.1016/J.JTBI.2015.11.008

Clarke, M. F., Dick, J. E., Dirks, P. B., Eaves, C. J., Jamieson, C. H. M., Jones, D. L., Visvader, J., Weissman, I. L., & Wahl, G. M. (2006). Cancer stem cells - Perspectives on current status and future directions: AACR workshop on cancer stem cells. *Cancer Res.*, *66*(19), 9339–9344. https://doi.org/10.1158/0008-5472.CAN-06-3126

Dormann, C. F., Calabrese, J. M., Guillera-Arroita, G., Matechou, E., Bahn, V., Bartoń, K., Beale, C. M., Ciuti, S., Elith, J., Gerstner, K., Guelat, J., Keil, P., Lahoz-Monfort, J. J., Pollock, L. J., Reineking, B., Roberts, D. R., Schröder, B., Thuiller, W., Warton, D. I., … Hartig, F. (2018). Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. *Ecol. Monogr.*, *88*(4), 485–504. https://doi.org/10.1002/ecm.1309

Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Brief. Bioinformatics*, *21*(2), 553–565. https://doi.org/10.1093/BIB/BBZ016

Edwards, A. W. F. (1992). *Likelihood (Expanded Edition)* (Johns Hopkins University Press)

Feroz, F., & Hobson, M. P. (2008). Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. *Mon. Notices Royal Astron. Soc.*, *384*(2), 449–463. https://doi.org/10.1111/J.1365-2966.2007.12353.X

Feroz, F., Hobson, M. P., & Bridges, M. (2009). MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Notices Royal Astron. Soc.*, *398*(4), 1601–1614. https://doi.org/10.1111/J.1365-2966.2009.14548.X

Feroz, F., Hobson, M. P., Cameron, E., & Pettitt, A. N. (2019). Importance Nested Sampling and the MultiNest Algorithm. *Open J. Astrophysics*, *2*(1), 11120. https://doi.org/10.21105/ASTRO.1306.2144

Fragoso, T. M., Bertoli, W., & Louzada, F. (2017). Bayesian Model Averaging: A Systematic Review and Conceptual Classification. *Int. Stat. Rev.*, *86*(1), 1–28. https://doi.org/10.1111/INSR.12243

Fuentes-Garí, M., Misener, R., García-Munzer, D., Velliou, E., Georgiadis, M. C., Kostoglou, M., Pistikopoulos, E. N., Panoskaltsis, N., & Mantalaris, A. (2015). A mathematical model of subpopulation kinetics for the deconvolution of leukaemia heterogeneity. *J. R. Soc. Interface*, 12: 20150276. https://doi.org/10.1098/RSIF.2015.0276

Galipaud, M., Gillingham, M. A. F., David, M., & Dechaume-Moncharmont, F.-X. (2014). Ecologists overestimate the importance of predictor variables in model averaging: a plea for cautious interpretations. *Methods Ecol. Evol.*, *5*(10), 983–991. https://doi.org/10.1111/2041-210X.12251

Galipaud, M., Gillingham, M. A. F., & Dechaume-Moncharmont, F. X. (2017). A farewell to the sum of Akaike weights: The benefits of alternative metrics for variable importance estimations in model selection. *Methods Ecol. Evol.*, *8*(12), 1668–1678. https://doi.org/10.1111/2041-210X.12835

Garthwaite, P. H., & Mubwandarikwa, E. (2010). Selection of Weights for Weighted Model Averaging. *Aust. N. Z. J. Stat*, *52*(4), 363–382. https://doi.org/10.1111/j.1467-842X.2010.00589.x

Gay, C. M., Stewart, C. A., Park, E. M., Diao, L., Groves, S. M., Heeke, S., Nabet, B. Y., Fujimoto, J., Solis, L. M., Lu, W., Xi, Y., Cardnell, R. J., Wang, Q., Fabbri, G., Cargill, K. R., Vokes, N. I., Ramkumar, K., Zhang, B., della Corte, C. M., … Byers, L. A. (2021). Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. *Cancer Cell*, *39*(3), 346-360.e7. https://doi.org/10.1016/J.CCELL.2020.12.014

Greene, J. M., Levy, D., Fung, K. L., Souza, P. S., Gottesman, M. M., & Lavi, O. (2015). Modeling intrinsic heterogeneity and growth of cancer cells. *J. Theor. Biol.*, *367*, 262–277. https://doi.org/10.1016/J.JTBI.2014.11.017

Gupta, P. B., Fillmore, C. M., Jiang, G., Shapira, S. D., Tao, K., Kuperwasser, C., & Lander, E. S. (2011). Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. *Cell*, *146*(4), 633–644. https://doi.org/10.1016/J.CELL.2011.07.026

Harris, L. A., Beik, S., Ozawa, P. M. M., Jimenez, L., & Weaver, A. M. (2019). Modeling heterogeneous tumor growth dynamics and cell–cell interactions at single-cell and cell-population resolution. *Curr. Opin. Syst. Biol.*, *17*, 24–34. https://doi.org/10.1016/J.COISB.2019.09.005

Hata, A. N., Niederst, M. J., Archibald, H. L., Gomez-Caraballo, M., Siddiqui, F. M., Mulvey, H. E., Maruvka, Y. E., Ji, F., Bhang, H. C., Radhakrishna, V. K., Siravegna, G., Hu, H., Raoof, S., Lockerman, E., Kalsy, A., Lee, D., Keating, C. L., Ruddy, D. A., Damon, L. J., … Engelman, J. A. (2016). Tumor cells can follow distinct evolutionary paths to become resistant to epidermal growth factor receptor inhibition. *Nat. Med.*, *22*(3), 262–269. https://doi.org/10.1038/nm.4040

Hayford, C. E., Tyson, D. R., Jack Robbins, C., Frick, P. L., Quaranta, V., & Harris, L. A. (2021). An in vitro model of tumor heterogeneity resolves genetic, epigenetic, and stochastic sources of cell state variability. *PLoS Biol.*, *19*(6). https://doi.org/10.1371/JOURNAL.PBIO.3000797

Hinne, M., Gronau, Q. F., van den Bergh, D., & Wagenmakers, E.-J. (2020). A Conceptual Introduction to Bayesian Model Averaging. *Adv. Meth. Pract. Psychol. Sci.*, *3*(2), 200–215. https://doi.org/10.1177/2515245919898657

Ireland, A. S., Micinski, A. M., Kastner, D. W., Guo, B., Wait, S. J., Spainhower, K. B., Conley, C. C., Chen, O. S., Guthrie, M. R., Soltero, D., Qiao, Y., Huang, X., Tarapcsák, S., Devarakonda, S., Chalishazar, M. D., Gertz, J., Moser, J. C., Marth, G., Puri, S., … Oliver, T. G. (2020). MYC Drives Temporal Evolution of Small Cell Lung Cancer Subtypes by Reprogramming Neuroendocrine Fate. *Cancer Cell*, *38*(1), 60-78.e12. https://doi.org/10.1016/j.ccell.2020.05.001

Jahchan, N. S., Lim, J. S., Bola, B., Morris, K., Seitz, G., Tran, K. Q., Xu, L., Trapani, F., Morrow, C. J., Cristea, S., Coles, G. L., Yang, D., Vaka, D., Kareta, M. S., George, J., Mazur, P. K., Nguyen, T., Anderson, W. C., Dylla, S. J., … Sage, J. (2016). Identification and Targeting of Long-Term Tumor-Propagating Cells in Small Cell Lung Cancer. *Cell Rep.*, *16*(3), 644–656. https://doi.org/10.1016/J.CELREP.2016.06.021

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *J. Am. Stat. Assoc.*, *90*(430), 773–795. https://doi.org/10.2307/2291091

Kaur, G., Reinhart, R. A., Monks, A., Evans, D., Morris, J., Polley, E., & Teicher, B. A. (2016). Bromodomain and hedgehog pathway targets in small cell lung cancer. *Cancer Lett.*, *371*(2), 225–239. https://doi.org/10.1016/J.CANLET.2015.12.001

Krakauer, D. C., Collins, J. P., Erwin, D., Flack, J. C., Fontana, W., Laubichler, M. D., Prohaska, S. J., West, G. B., & Stadler, P. F. (2011). The challenges and scope of theoretical biology. *J. Theor. Biol.*, *276*(1), 269–276. https://doi.org/10.1016/j.jtbi.2011.01.051

Kruschke, J. (2011). Doing Bayesian data analysis: a tutorial with R and BUGS (Burlington, MA: Academic Press)

Ku, S. Y., Rosario, S., Wang, Y., Mu, P., Seshadri, M., Goodrich, Z. W., Goodrich, M. M., Labbé, D. P., Gomez, E. C., Wang, J., Long, H. W., Xu, B., Brown, M., Loda, M., Sawyers, C. L., Ellis, L., & Goodrich, D. W. (2017). Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science*, *355*(6320), 78-83. https://doi.org/10.1126/SCIENCE.AAH4199

Kwon, M.-C., Proost, N., Song, J.-Y., Sutherland, K. D., Zevenhoven, J., & Berns, A. (2015). Paracrine signaling between tumor subclones of mouse SCLC: a critical role of ETS transcription factor Pea3 in facilitating metastasis. *Genes Dev.* 29: 1587-1592. https://doi.org/10.1101/gad.262998.115

Lander, A. D., Gokoffski, K. K., Wan, F. Y. M., Nie, Q., & Calof, A. L. (2009). Cell Lineages and the Logic of Proliferative Control. *PLoS Biol.*, *7*(1), e1000015. https://doi.org/10.1371/JOURNAL.PBIO.1000015

Lema-Perez, L., Muñoz-Tamayo, R., Garcia-Tirado, J., & Alvarez, H. (2019). On parameter interpretability of phenomenological-based semiphysical models in biology. *Inform. Med. Unlocked*, *15*, 100158. https://doi.org/10.1016/J.IMU.2019.02.002

Lim, J. S., Ibaseta, A., Fischer, M. M., Cancilla, B., O'Young, G., Cristea, S., Luca, V. C., Yang, Di., Jahchan, N. S., Hamard, C., Antoine, M., Wislez, M., Kong, C., Cain, J., Liu, Y. W., Kapoun, A. M., Garcia, K. C., Hoey, T., Murriel, C. L., & Sage, J. (2017). Intratumoural heterogeneity generated by Notch signalling promotes small-cell lung cancer. *Nature*, *545*(7654), 360–364. https://doi.org/10.1038/nature22323

Lopez, C. F., Muhlich, J. L., Bachman, J. A., & Sorger, P. K. (2013). Programming biological models in Python using PySB. *Mol. Syst. Biol.*, *9*(1), 646. https://doi.org/10.1038/MSB.2013.1

Lu, M., Jolly, M. K., Levine, H., Onuchic, J. N., & Ben-Jacob, E. (2013). MicroRNA-based regulation of epithelial–hybrid–mesenchymal fate determination. *Proc. Natl. Acad. Sci. U.S.A.*, *110*(45), 18144–18149. https://doi.org/10.1073/PNAS.1318192110

Lu, Z., & Lou, W. (2022). Bayesian approaches to variable selection: A comparative study from practical perspectives. *Int. J. Biostat.*, 18(1): 83-108. https://doi.org/10.1515/ijb-2020-0130

Marusyk, A., Almendro, V., & Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer 2012 12*(5), 323–334. https://doi.org/10.1038/nrc3261

Metzcar, J., Wang, Y., Heiland, R., & Macklin, P. (2019). A Review of Cell-Based Computational Modeling in Cancer Biology. *JCO Clin. Cancer Inform.*, *3*, 1–13. https://doi.org/10.1200/cci.18.00069

Mitra, E. D., & Hlavacek, W. S. (2019). Parameter estimation and uncertainty quantification for systems biology models. *Curr. Opin. Syst. Biol.*, *18*, 9–18. https://doi.org/10.1016/j.coisb.2019.10.006

Mollaoglu, G., Guthrie, M. R., Böhm, S., Brägelmann, J., Can, I., Ballieu, P. M., Marx, A., George, J., Heinen, C., Chalishazar, M. D., Cheng, H., Ireland, A. S., Denning, K. E., Mukhopadhyay, A., Vahrenkamp, J. M., Berrett, K. C., Mosbruger, T. L., Wang, J., Kohan, J. L., … Oliver, T. G. (2017). MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. *Cancer Cell*, *31*(2), 270–285. https://doi.org/10.1016/J.CCELL.2016.12.005

Mu, P., Zhang, Z., Benelli, M., Karthaus, W. R., Hoover, E., Chen, C. C., Wongvipat, J., Ku, S. Y., Gao, D., Cao, Z., Shah, N., Adams, E. J., Abida, W., Watson, P. A., Prandi, D., Huang, C. H., de Stanchina, E., Lowe, S. W., Ellis, L., … Sawyers, C. L. (2017). SOX2 promotes lineage plasticity and antiandrogen resistance in TP53-and RB1-deficient prostate cancer. *Science*, *355*(6320) : 84-88. https://doi.org/10.1126/SCIENCE.AAH4307

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., Hoang, C. D., Diehn, M., & Alizadeh, A. A. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods 2015 12*(5), 453–457. https://doi.org/10.1038/nmeth.3337

Ouadah, Y., Rojas, E. R., Riordan, D. P., Capostagno, S., Kuo, C. S., & Krasnow, M. A. (2019). Rare Pulmonary Neuroendocrine Cells Are Stem Cells Regulated by Rb, p53, and Notch. *Cell*, *179*(2), 403-416.e23. https://doi.org/10.1016/J.CELL.2019.09.010

Paudel, B. B., Harris, L. A., Hardeman, K. N., Abugable, A. A., Hayford, C. E., Tyson, D. R., & Quaranta, V. (2018). A Nonquiescent "Idling" Population State in Drug-Treated, BRAF-Mutated Melanoma. *Biophys. J.*, 114(6): 1499-1511. https://doi.org/10.1016/j.bpj.2018.01.016

Pettengill, O. S., Sorenson, G. D., Wurster-Hill, D. H., Curphey, T. J., Noll, W. W., Gate, C. C., & Maurer, L. H. (1980). Isolation and growth characteristics of continuous cell lines from small-cell carcinoma of the lung. *Cancer*, *45*(5): 906-918. https://doi.org/10.1002/1097-0142(19800301)45:5%3C906::AID-CNCR2820450513%3E3.0.CO;2-H

Pisco, A. O., Brock, A., Zhou, J., Moor, A., Mojtahedi, M., Jackson, D., & Huang, S. (2013). Non-Darwinian dynamics in therapy-induced cancer drug resistance. *Nat. Commun. 2013 4*(1), 1–11. https://doi.org/10.1038/ncomms3467

Polley, E., Kunkel, M., Evans, D., Silvers, T., Delosh, R., Laudeman, J., Ogle, C., Reinhart, R., Selby, M., Connelly, J., Harris, E., Fer, N., Sonkin, D., Kaur, G., Monks, A., Malik, S., Morris, J., & Teicher, B. A. (2016). Small Cell Lung Cancer Screen of Oncology Drugs, Investigational Agents, and Gene and microRNA Expression. *J. Natl. Cancer Inst.*, *108*(10), djw122. https://doi.org/10.1093/JNCI/DJW122

Raue, A., Karlsson, J., Saccomani, M. P., Jirstrand, M., & Timmer, J. (2014). Comparison of approaches for parameter identifiability analysis of biological systems. *Bioinformatics*, *30*(10), 1440–1448. https://doi.org/10.1093/BIOINFORMATICS/BTU006

Risom, T., Langer, E. M., Chapman, M. P., Rantala, J., Fields, A. J., Boniface, C., Alvarez, M. J., Kendsersky, N. D., Pelz, C. R., Johnson-Camacho, K., Dobrolecki, L. E., Chin, K., Aswani, A. J., Wang, N. J., Califano, A., Lewis, M. T., Tomlin, C. J., Spellman, P. T., Adey, A., … Sears, R. C. (2018). Differentiation-state plasticity is a targetable resistance mechanism in basal-like breast cancer. *Nat. Commun.*, *9*(3815), 1–17. https://doi.org/10.1038/s41467-018-05729-w

Royall, R. M. (1997). *Statistical evidence: a likelihood paradigm* (Chapman and Hall/CRC)

Rudin, C. M., Brambilla, E., Faivre-Finn, C., & Sage, J. (2021). Small-cell lung cancer. *Nat. Rev. Dis. Primers*, *7*, 3. https://doi.org/10.1038/S41572-020-00235-0

Rudin, C. M., Poirier, J. T., Byers, L. A., Dive, C., Dowlati, A., George, J., Heymach, J. v., Johnson, J. E., Lehman, J. M., MacPherson, D., Massion, P. P., Minna, J. D., Oliver, T. G., Quaranta, V., Sage, J., Thomas, R. K., Vakoc, C. R., & Gazdar, A. F. (2019). Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat. Rev. Cancer 2019 19*(5), 289–297. https://doi.org/10.1038/s41568-019-0133-9

Shaffer, S. M., Dunagin, M. C., Torborg, S. R., Torre, E. A., Emert, B., Krepler, C., Beqiri, M., Sproesser, K., Brafford, P. A., Xiao, M., Eggan, E., Anastopoulos, I. N., Vargas-Garcia, C. A., Singh, A., Nathanson, K. L., Herlyn, M., & Raj, A. (2018). Rare cell variability and drug-induced reprogramming as a mode of cancer drug resistance. *Nature 546:* 431-435. https://doi.org/10.1038/nature22794

Shue, Y. T., Drainas, A. P., Li, N. Y., Pearsall, S. M., Morgan, D., Sinnott-Armstrong, N., Hipkins, S. Q., Coles, G. L., Lim, J. S., Oro, A. E., Simpson, K. L., Dive, C., & Sage, J. (2022) A conserved YAP/Notch/REST network controls the neuroendocrine cell fate in the lungs. *Nat. Commun.* 13, 2690. https://doi.org/10.1038/s41467-022-30416-2

Sirzén, F., Zhivotovsky, B., Nilsson, A., Bergh, J., & Lewensohn, R. (1998). 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), 1–13. https://doi.org/10.1016/S0169-5002(98)00066-X

Skilling, J. (2004). Nested Sampling. *AIP Conf. Proc.*, *735*, 395. https://doi.org/10.1063/1.1835238

Stewart, C. A., Gay, C. M., Xi, Y., Sivajothi, S., Sivakamasundari, V., Fujimoto, J., Bolisetty, M., Hartsfield, P. M., Balasubramaniyan, V., Chalishazar, M. D., Moran, C., Kalhor, N., Stewart, J., Tran, H., Swisher, S. G., Roth, J. A., Zhang, J., de Groot, J., Glisson, B., … Byers, L. A. (2020). Single-cell analyses reveal increased intratumoral heterogeneity after the onset of therapy resistance in small-cell lung cancer. *Nat. Cancer 1*(4), 423–436. https://doi.org/10.1038/s43018-019-0020-z

Stumpf, M. P. H. (2020). Multi-model and network inference based on ensemble estimates: avoiding the madness of crowds. *J. R. Soc. Interface*, *17*: 20200419. https://doi.org/10.1098/RSIF.2020.0419

Turner, C., & Kohandel, M. (2012). Quantitative approaches to cancer stem cells and epithelial–mesenchymal transition. *Semin. Cancer Biol.*, *22*(5–6), 374–378. https://doi.org/10.1016/J.SEMCANCER.2012.04.005

Udyavar, A. R., Wooten, D. J., Hoeksema, M., Bansal, M., Califano, A., Estrada, L., Schnell, S., Irish, J. M., Massion, P. P., & Quaranta, V. (2017). Novel Hybrid Phenotype Revealed in Small Cell Lung

Cancer by a Transcription Factor Network Model That Can Explain Tumor Heterogeneity. *Cancer Res.*, *77*(5), 1063–1074. https://doi.org/10.1158/0008-5472.CAN-16-1467

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nat. Rev. Methods Primers 2021 1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2018). A simple introduction to Markov Chain Monte–Carlo sampling. *Psychon. Bull. Rev.*, *25*(1), 143–154. https://doi.org/10.3758/S13423-016-1015-8

Villaverde, A. F., & Banga, J. R. (2014). Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *J. R. Soc. Interface*, *11*: 20130505. https://doi.org/10.1098/RSIF.2013.0505

Wang, W., Quan, Y., Fu, Q., Liu, Y., Liang, Y., Wu, J., Yang, G., Luo, C., Ouyang, Q., & Wang, Y. (2014). Dynamics between Cancer Cell Subpopulations Reveals a Model Coordinating with Both Hierarchical and Stochastic Concepts. *PLoS One*, *9*(1), e84654. https://doi.org/10.1371/JOURNAL.PONE.0084654

Watson, S. S., Dane, M., Chin, K., Jonas, O., Gray, J. W., & Korkola, J. E. (2018). Microenvironment-Mediated Mechanisms of Resistance to HER2 Inhibitors Differ between HER2+ Breast Cancer Subtypes. *Cell Syst.*, *6*, 329–342. https://doi.org/10.1016/j.cels.2018.02.001

Williamson, S. C., Metcalf, R. L., Trapani, F., Mohan, S., Antonello, J., Abbott, B., Leong, H. S., Chester, C. P. E., Simms, N., Polanski, R., Nonaka, D., Priest, L., Fusi, A., Carlsson, F., Carlsson, A., Hendrix, M. J. C., Seftor, R. E. B., Seftor, E. A., Rothwell, D. G., … Dive, C. (2016). Vasculogenic mimicry in small cell lung cancer. *Nat. Commun. 2016 7*(1), 1–14. https://doi.org/10.1038/ncomms13322

Wooten, D. J., Groves, S. M., Tyson, D. R., Liu, Q., Lim, J. S., Albert, R., Lopez, C. F., Sage, J., & Quaranta, V. (2019). Systems-level network modeling of Small Cell Lung Cancer subtypes identifies master regulators and destabilizers. *PLoS Comput. Biol. 15*(10), e1007343. https://doi.org/10.1371/JOURNAL.PCBI.1007343

Würstle, M. L., Zink, E., Prehn, J. H. M., & Rehm, M. (2014). From computational modelling of the intrinsic apoptosis pathway to a systems-based analysis of chemotherapy resistance: achievements, perspectives and challenges in systems medicine. *Cell Death Dis.*, *5*(5), e1258–e1258. https://doi.org/10.1038/cddis.2014.36

Yang, G., Quan, Y., Wang, W., Fu, Q., Wu, J., Mei, T., Li, J., Tang, Y., Luo, C., Ouyang, Q., Chen, S., Wu, L., Hei, T. K., & Wang, Y. (2012). Dynamic equilibrium between cancer stem cells and non-stem cancer cells in human SW620 and MCF-7 cancer cell populations. *Br. J. Cancer*, *106*(9), 1512–1519. https://doi.org/10.1038/bjc.2012.126

Zeman, K. G., Brzezniak, C. E., & Carter, C. A. (2017). Recalcitrant small cell lung cancer: the argument for optimism. *J. Thorac. Dis.*, *9*(3), e295–e296. https://doi.org/10.21037/JTD.2017.03.40

Zhou, J. X., Pisco, A. O., Qian, H., & Huang, S. (2014). Nonequilibrium Population Dynamics of Phenotype Conversion of Cancer Cells. *PLoS One*, *9*(12), e110714. https://doi.org/10.1371/JOURNAL.PONE.0110714

Zou, M., Toivanen, R., Mitrofanova, A., Floch, N., Hayati, S., Sun, Y., le Magnen, C., Chester, D., Mostaghel, E. A., Califano, A., Rubin, M. A., Shen, M. M., Abate-Shen, C., & Shen, M. (2017). Transdifferentiation as a Mechanism of Treatment Resistance in a Mouse Model of Castration-Resistant Prostate Cancer. *Cancer Discov.* *7*(7), 737–749. https://doi.org/10.1158/2159-8290.CD-16-1174

Main Figures

Figure 1: Conclusions and hypotheses from the literature, along with experimental data from mouse and human tumors, enables building a mechanistic hypothesis exploration space for SCLC tumor growth and development. 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. TKO, $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) that we previously assigned as representative of tumors made up largely of the SCLC-A subtype (Wooten et al., 2019)).

A. Synthesis of what is currently known about SCLC subtypes, which have been divided into two overall phenotypes, neuroendocrine and non-neuroendocrine, and then further classified into subtypes based on transcription factor expression (SCLC-P not shown). (1) Neuroendocrine SCLC cells, which do not express Hes1, transition into non-neuroendocrine 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) Predicted subtype with transcriptional signature intermediate between neuroendocrine and non-neuroendocrine, named SCLC-A2. (5) Phenotypic transitions occur in a hierarchical manner from SCLC-A to SCLC-N to SCLC-Y cells.
B. 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 RNA-sequencing data.

Figure 2: Exploring SCLC hypotheses via model selection requires building candidate models representing single or combinations of these hypotheses.

A. Model topologies constructed with 2+ subtypes, with number of combinations per number of subtypes.
B. Subtype effect schema, where there are different effectors between candidates and different affected cellular actions.
C. Potential transition schema, where all subtypes in topology must be accessible either as initiating subtypes or via transitions (top), unidirectional transitions are those that follow a hierarchy (middle), and bidirectional transitions must be symmetrical when present (middle/bottom).

D. Spectrum of tumor behavior, from the cancer stem cell hypothesis: one originating cell, hierarchical phenotypic transitions, to the plasticity hypothesis: multiple originating or early subtypes, transitions irrespective of hierarchy. An intermediate behavior is also possible: one originating cell, transitions without hierarchy.

Figure 3: Fitting to experimental data and assigning Bayesian evidence based on model fits separates candidate models into more and less likely. TKO, $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) representative of SCLC-A tumors (Wooten et al., 2019)).

(A) - (C). 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 101/2 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.

Figure 4: Most likely model topology varies across datasets, while transition parameter rates vary according to subtype presence in similar ways across datasets. TKO, $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) representative of SCLC-A tumors (Wooten et al., 2019)). (*) 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.

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 (A) and four-subtype topology per dataset.

Figure 5: Across datasets, multimodel inference indicates that bidirectional phentoypic transitions are likely, suggesting high SCLC phenotypic plasticity. Tsn, transition (e.g., subtype transition). TKO, $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) that we previously assigned as representative of tumors made up largely of the SCLC-A subtype (Wooten et al., 2019)). (*) 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.

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 probability, respectively.

(B)-(D). 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 (A). Top-scoring three-state topology for:

B. TKO dataset,
C. RPM dataset, and
D. SCLC-A cell line dataset.
E. Schematic of consolidated model behaviors, drawn from each dataset's high-probability three-subtype topology results ((B)-(D)). 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 (A)-(D).
F. Parameter fitting results (part of the nested sampling algorithm) for four-subtype topology models initiated by A +/- other subtypes, across datasets.

Figure 4: Most likely model topology varies across datasets, while transition parameter rates vary according to subtype presence in similar ways across datasets. TKO, $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012)

representative of SCLC-A tumors (Wooten et al., 2019)). (*) 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.

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 (A) and four-subtype topology per dataset.

Box 1: Population dynamics modeling and inter-subtype effects.

Box 2: Bayesian statistics and multimodel inference (MMI).

Box 3: Multiple methods for model selection / model averaging, from traditional (least informative) to Bayesian (most informative).

Main Tables

Table 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 (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017)

Table 2: Probabilities after hypothesis exploration using Bayesian multimodel inference. 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 (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) representative of SCLC-A tumors (Wooten et al., 2019))

Supplemental Figures

Figure S1: Cell fate rate, and rate parameter prior, calculations.

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 Box 1) (Harris et al., 2019).
B. Example calculation of division rate parameter prior for H841, representation of subtype Y, (see **Table S1**) 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 S1**) 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 S1**) 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 (A). 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 S2: 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 S3: Fitting models to data as part of nested sampling results in simulations that better fit the data 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**). 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 Note S1 for more detail.

Figure S4: Rate parameter posterior marginal distributions, after applying Bayesian 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 S5: Investigating probabilities and model-averaged parameter rates across datasets and within topologies. TKO, $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017); RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017); SCLC-A cell lines, a subset of SCLC cell lines from the CCLE (Barretina et al., 2012) representative of SCLC-A tumors (Wooten et al., 2019)).

A. Hypothesis assessment of model topologies per dataset, posterior probabilities based on all candidate models, with no filtering based on initiating subtype (see **Figure S3**). 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.

Supplemental Tables

Table S1: Existing data pertaining to SCLC intratumoral heterogeneity and communication used for rate parameter priors.

Table S2: Model variable posterior probabilities after hypothesis exploration using multimodel inference, TKO data in high-probability 3-subtype topology. TKO: $p53^{fl/fl}$;$Rb^{fl/fl}$;$p130^{fl/fl}$ tumors (Lim et al., 2017).

Table S3: Model variable posterior probabilities after hypothesis exploration using multimodel inference, RPM data in high-probability 3-subtype topology. RPM, $Rb1^{fl/fl}$;$Trp53^{fl/fl}$;Lox-Stop-Lox[LSL]-MycT58A tumors (Mollaoglu et al., 2017)

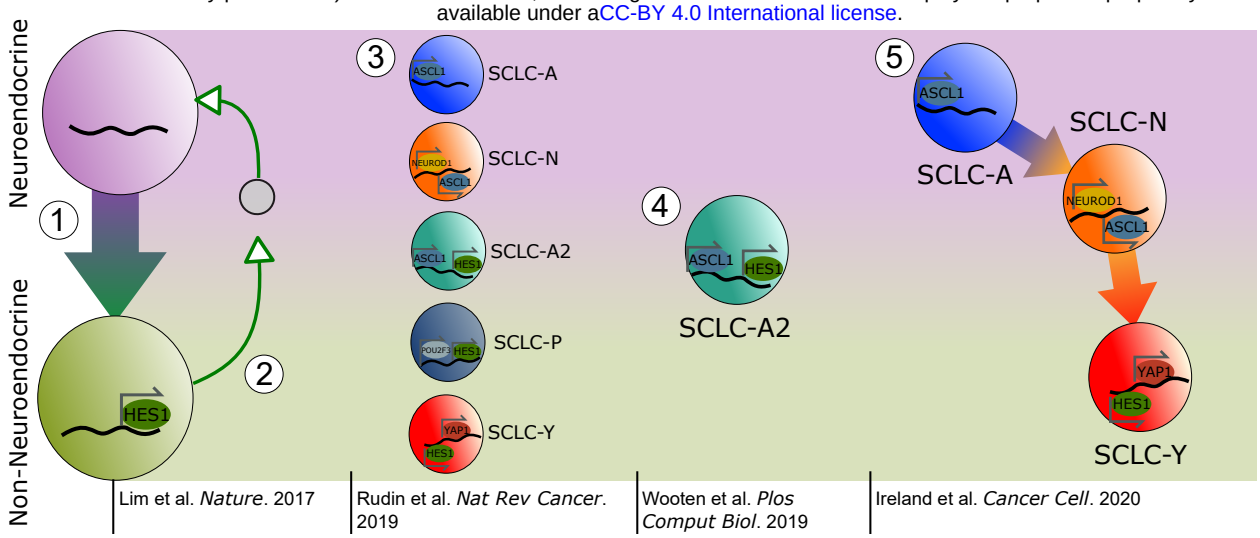
Table S4: Model variable posterior probabilities after hypothesis exploration using multimodel inference, SCLC-A cell line data in high-probability 3-subtype topology. SCLC-A cell lines, a

subset of SCLC cell lines from the CCLE (Barretina et al., 2012) representative of SCLC-A tumors (Wooten et al., 2019)).

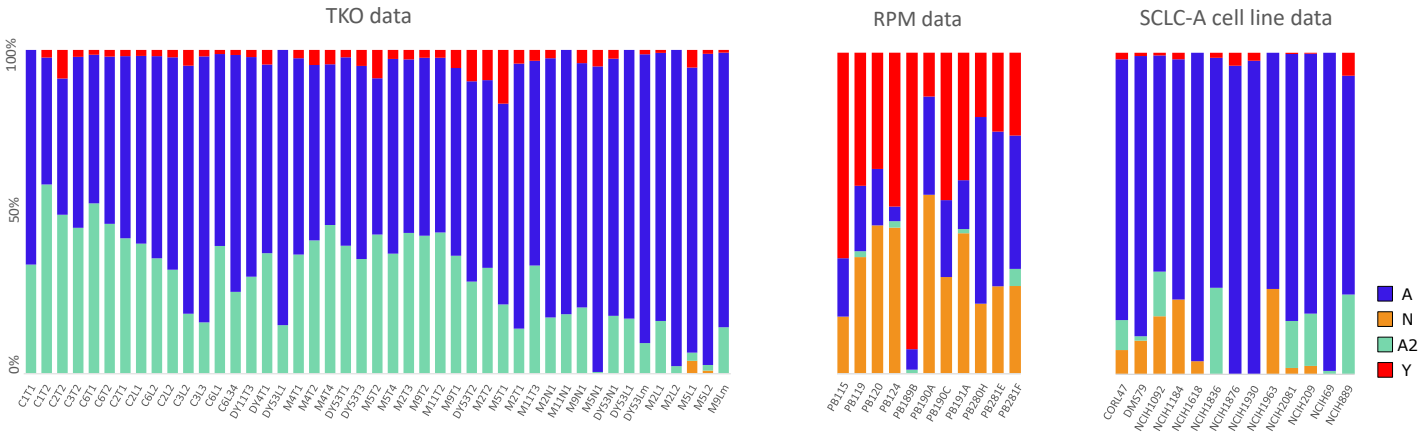
Supplemental Note

Note S1: Simulations using best-fitted parameters, as opposed to randomly-selected parameters from the prior distributions, replicate subtype proportions at steady state *update numbers when code finishes

A



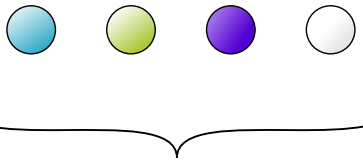
B



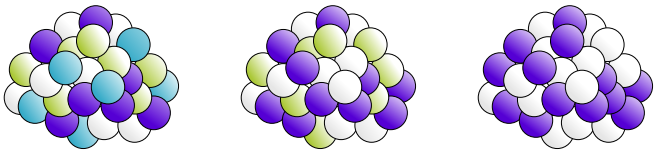
A

Candidates with varying subtype makeup

topologies



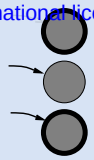
any of A, N, A2, Y



$\binom{4}{1} = 1$ $\binom{4}{3} = 4$ $\binom{4}{2} = 6$

C

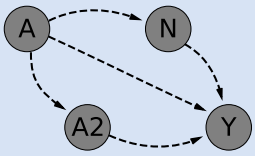
Candidates with varying initiation & transition schemes



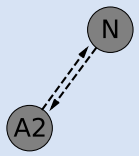
No transition, must initiate;

Not initiating, must have transition

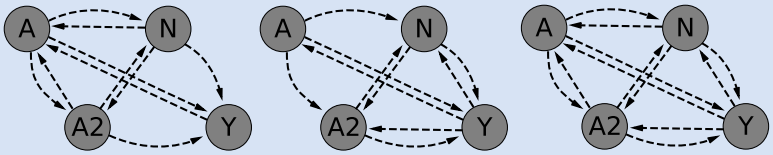
Can include transition and initiation



Unidirectional/
hierarchical transitions



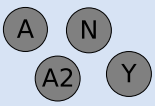
N, A2 transition always
bidirectional



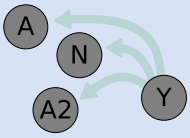
Bidirectional transitions require early (A, N, A2) symmetry,
late (N, A2, Y) symmetry, or both

B

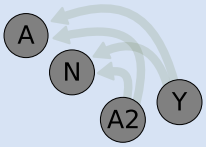
Candidates with varying effects and effectors



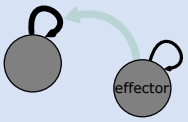
No effects



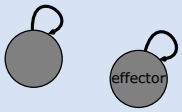
Y is the effector



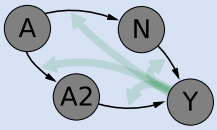
A2 and Y
are effectors



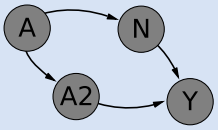
Effector
increases division



No effect
on division



Effector increases
transitions

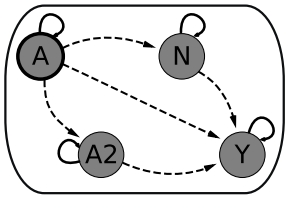


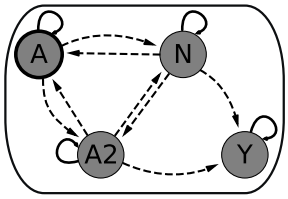
No effect on
transitions

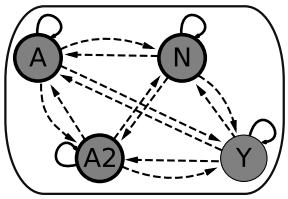
D

Topologies, effects, transition combinations lead to

5891 models* along potential hierarchy continuum



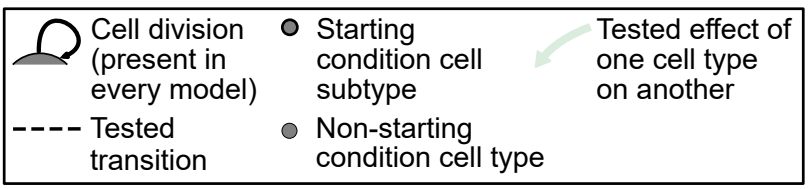


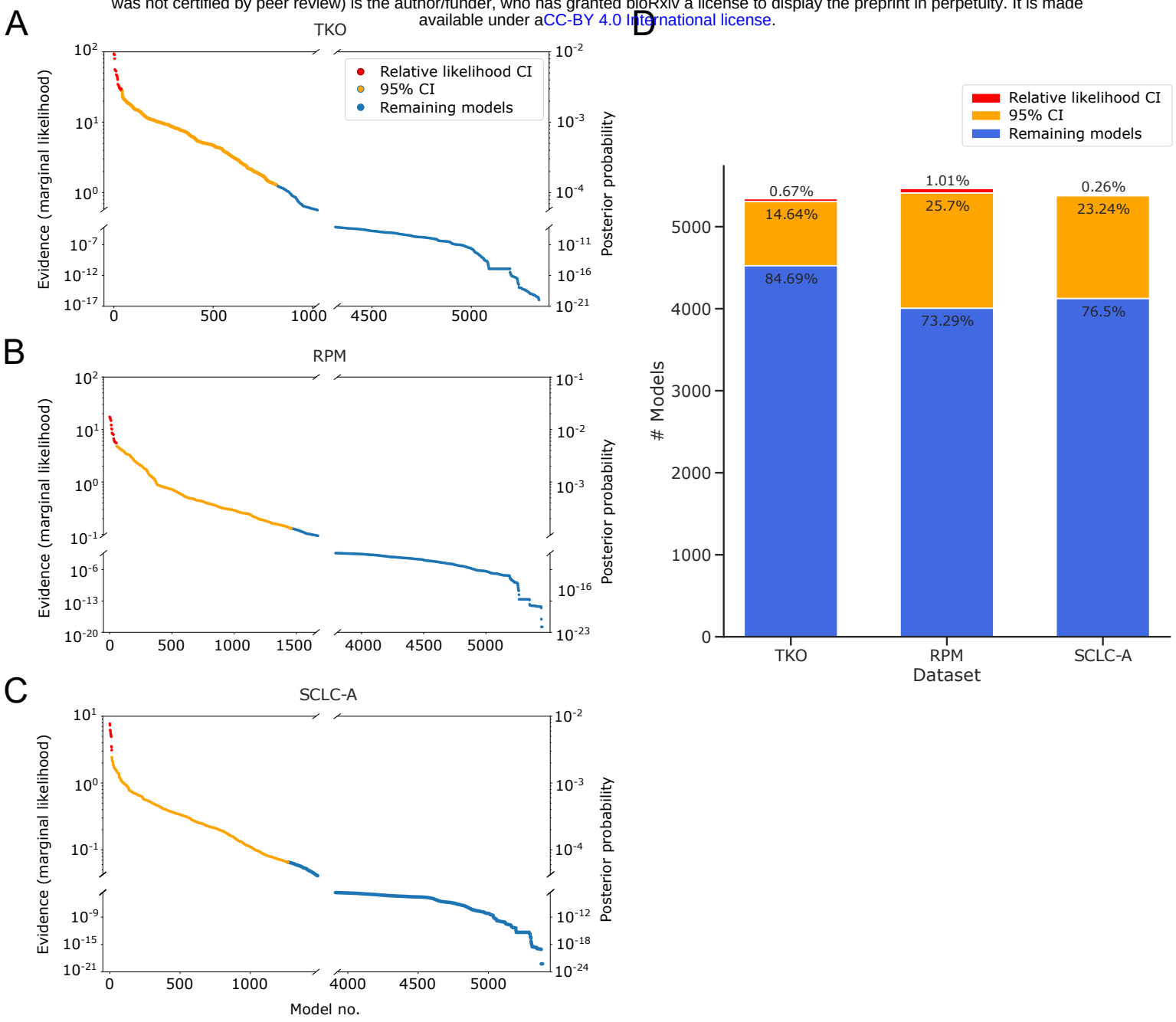


Cancer
stem cell
hypothesis

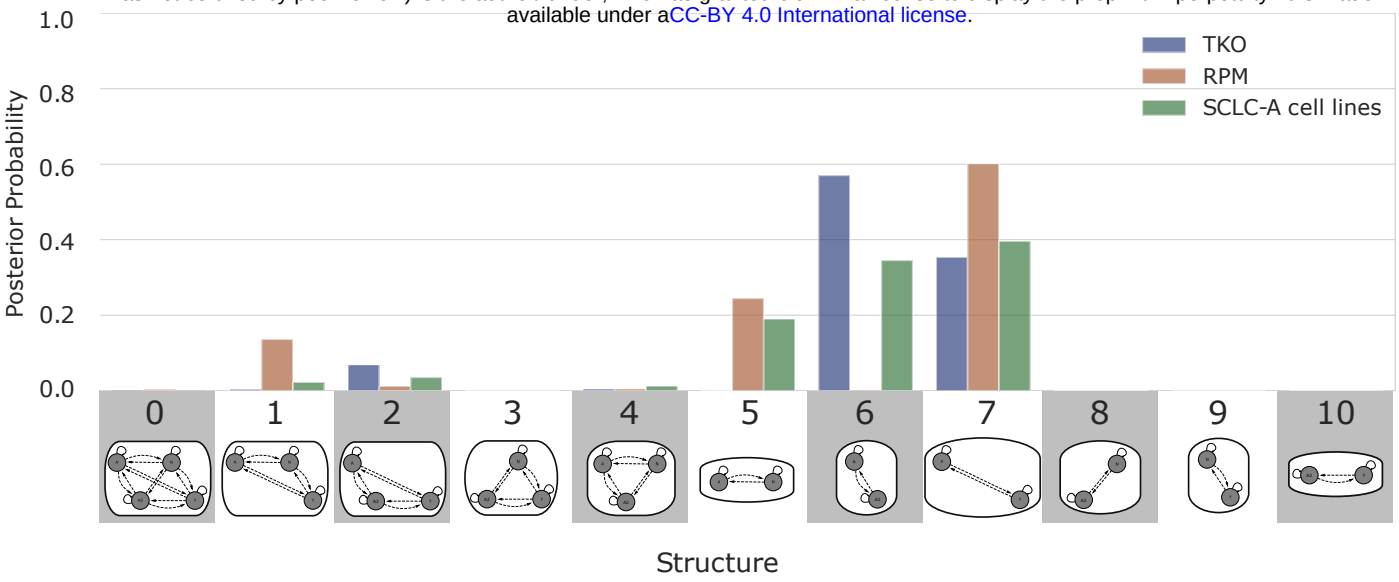
Plasticity
hypothesis

*only 4-subtype topology pictured in (C,D)

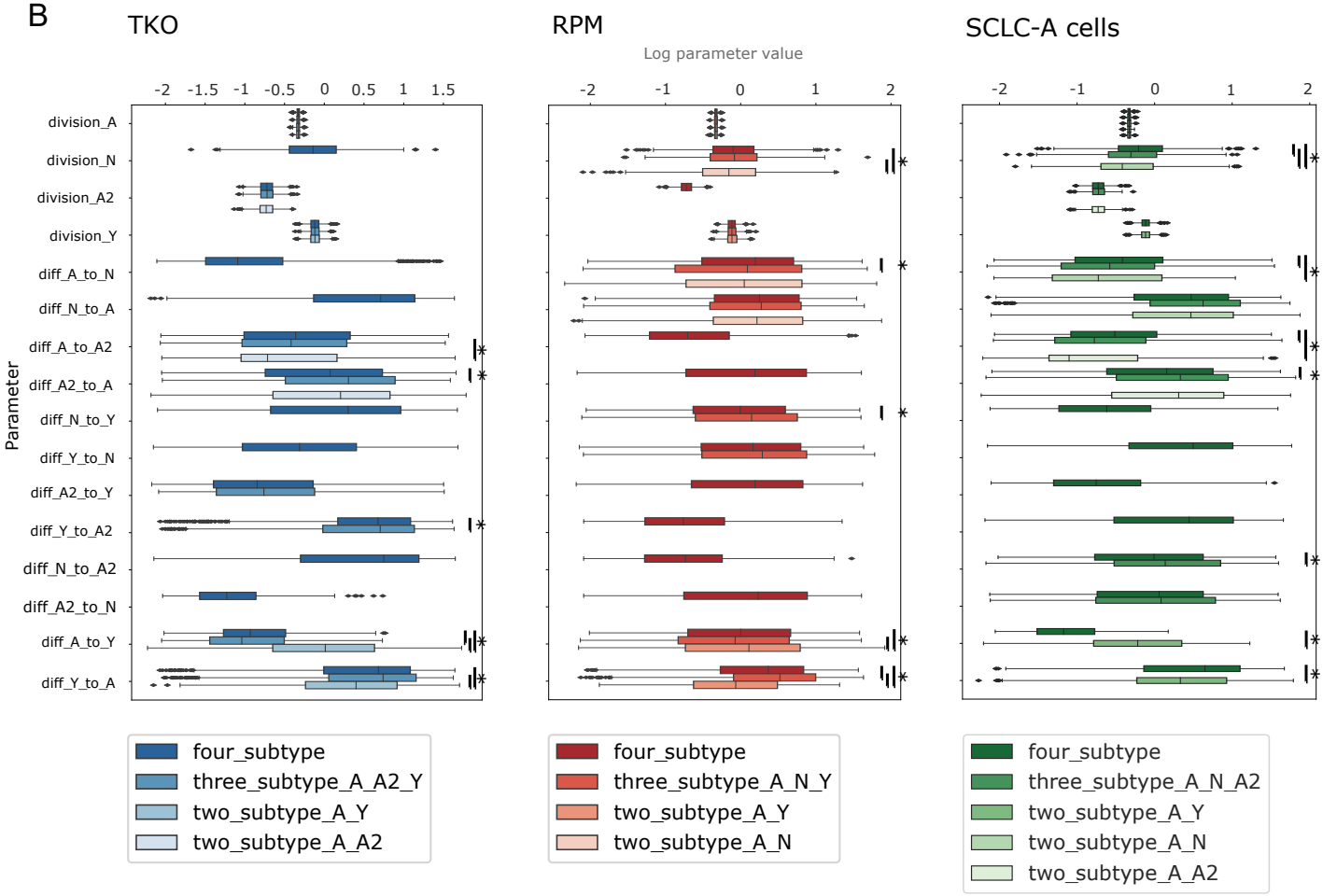


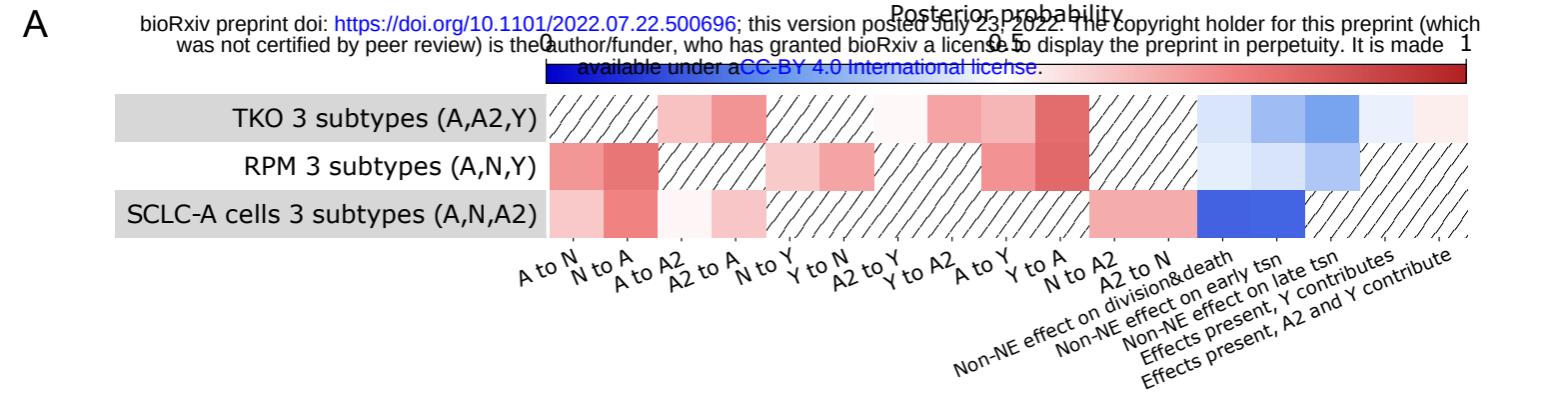


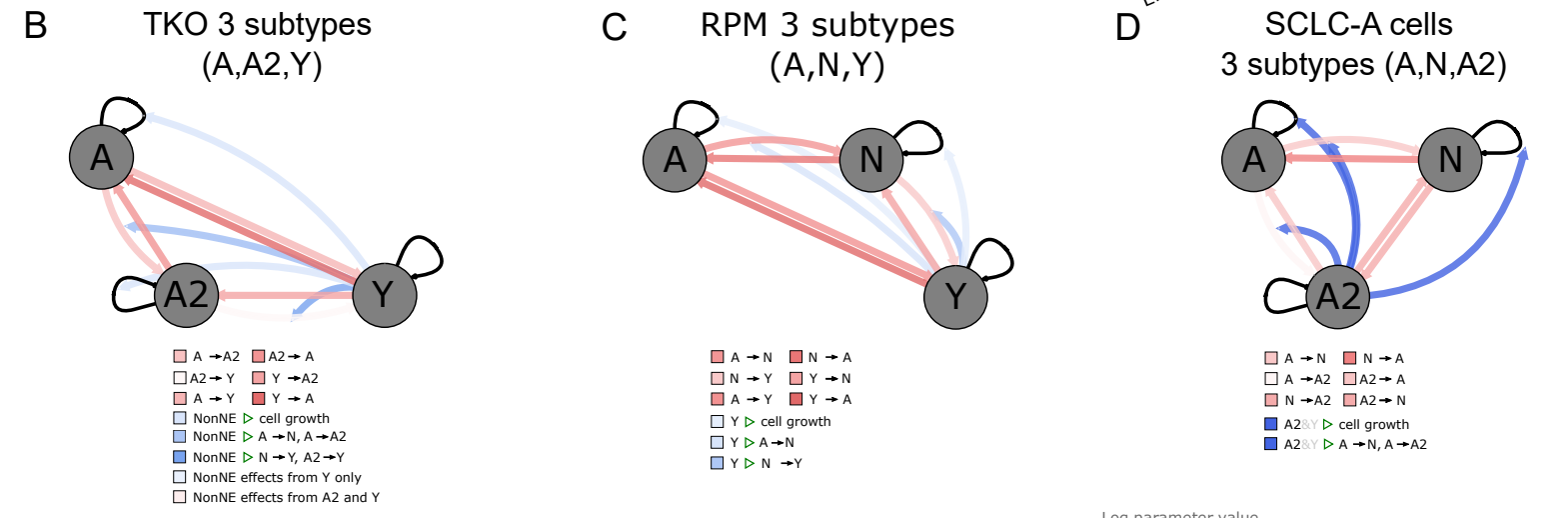
A

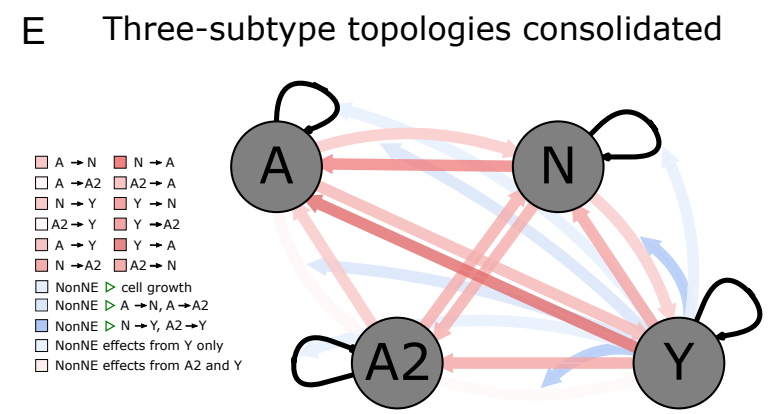


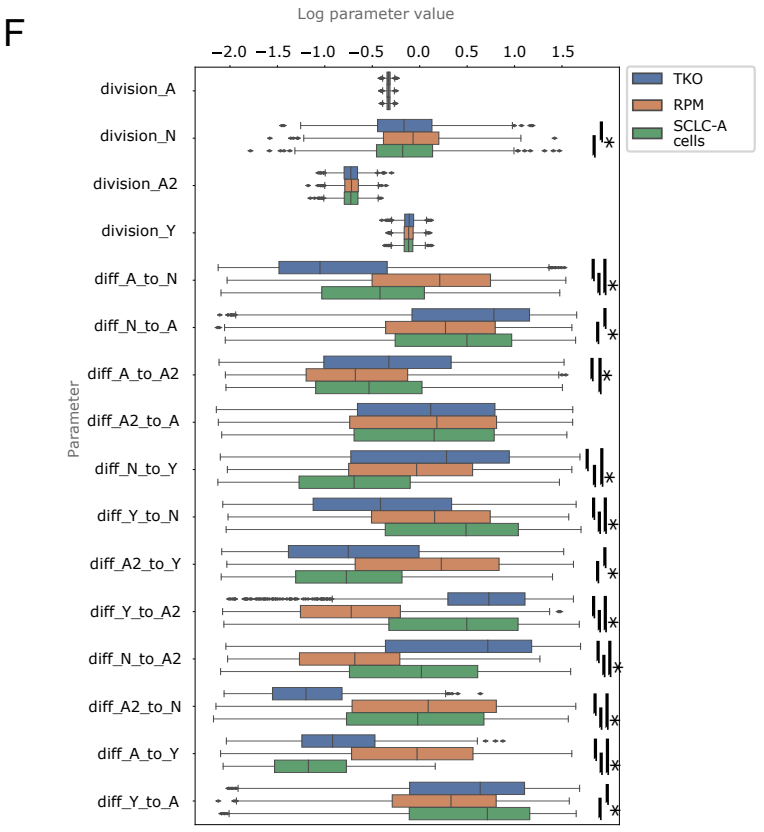
B







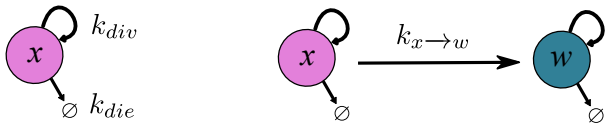




Box 1. Population dynamics modeling and inter-subtype effects

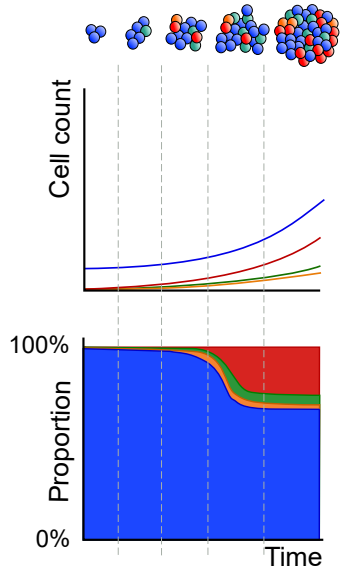
A population dynamics model represents behaviors of species over time, tracking the size of a cell population, rather than tracking individual cells. However, such a model can also include signaling and dependence beteween species (Harris et al., 2019). Cell population changes are represented as reactions, 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



$$x \xrightarrow{k_{div}} 2x$$
$$x \xrightarrow{k_{die}} \emptyset$$
$$x \xrightarrow{k_{x\to w}} w$$

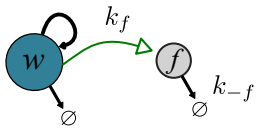
B



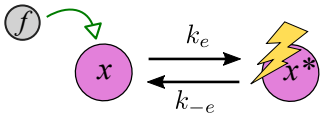
In our case, where species are cells of an SCLC tumor population, each is assigned a subtype identity (A, N, A2, or Y). The model simulates tumor growth over time, (B, top) calculating the increase in subpopulation amount (B, middle), from which tumor subtype proportion may be calculated (B, bottom) 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



$$w \xrightarrow{k_f} w + f$$
$$f \xrightarrow{k_{-f}} \emptyset$$



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

D

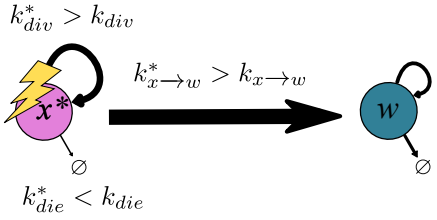
$$\frac{k_{-f}}{k_f} \equiv K_f^{eq} \qquad \frac{k_{-e}}{k_e} \equiv K_D$$

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

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



$$x \xrightarrow{k^*_{div}} 2x$$
$$x \xrightarrow{k^*_{die}} \emptyset$$
$$x \xrightarrow{k^*_{x\to w}} w$$

Box 2. Bayesian statistics and multimodel inference (MMI)

"Probability" in Bayesian statistics indicates the degree of belief or confidence in an event, or in a proposal that a statement may be true – that is, a hypothesis. The degree of belief in Bayesian statistics is affected by prior knowledge about the hypothesis combined with data. In this case, data can refer to measurements related to the hypothesis, synthesized into a likelihood that the hypothesis is true. Prior knowledge combined with data leads to the posterior probability, defined using Bayes' Theorem as

$$P(H|D) = \frac{P(D|H)P(H)}{\sum_{H\in\mathcal{H}} P(D|H)P(H)} \quad (1)$$

where $P(H|D)$ is the posterior probability of the hypothesis given the data D, $P(D|H)$ is the probability of observing the data D if the hypothesis H were true (also called the likelihood), and $P(H)$ is the prior probability of the hypothesis. Calculation of the posterior probability of the hypothesis $P(H|D)$ requires dividing the likelihood of one hypothesis times its prior, $P(H|D)P(H)$, by the sum of the likelihood times the prior probability for each hypothesis H within the set of all hypotheses $\mathcal{H}$ (shown in Eq. 1).

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:

$$BF = \frac{P(D|H_1)}{P(D|H_2)} = \frac{P(H_1|D)/P(H_1)}{P(H_2|D)/P(H_2)}; BF = \frac{P(H_1|D)}{P(H_2|D)} \; when \; P(H_1) = P(H_2) \quad (2)$$

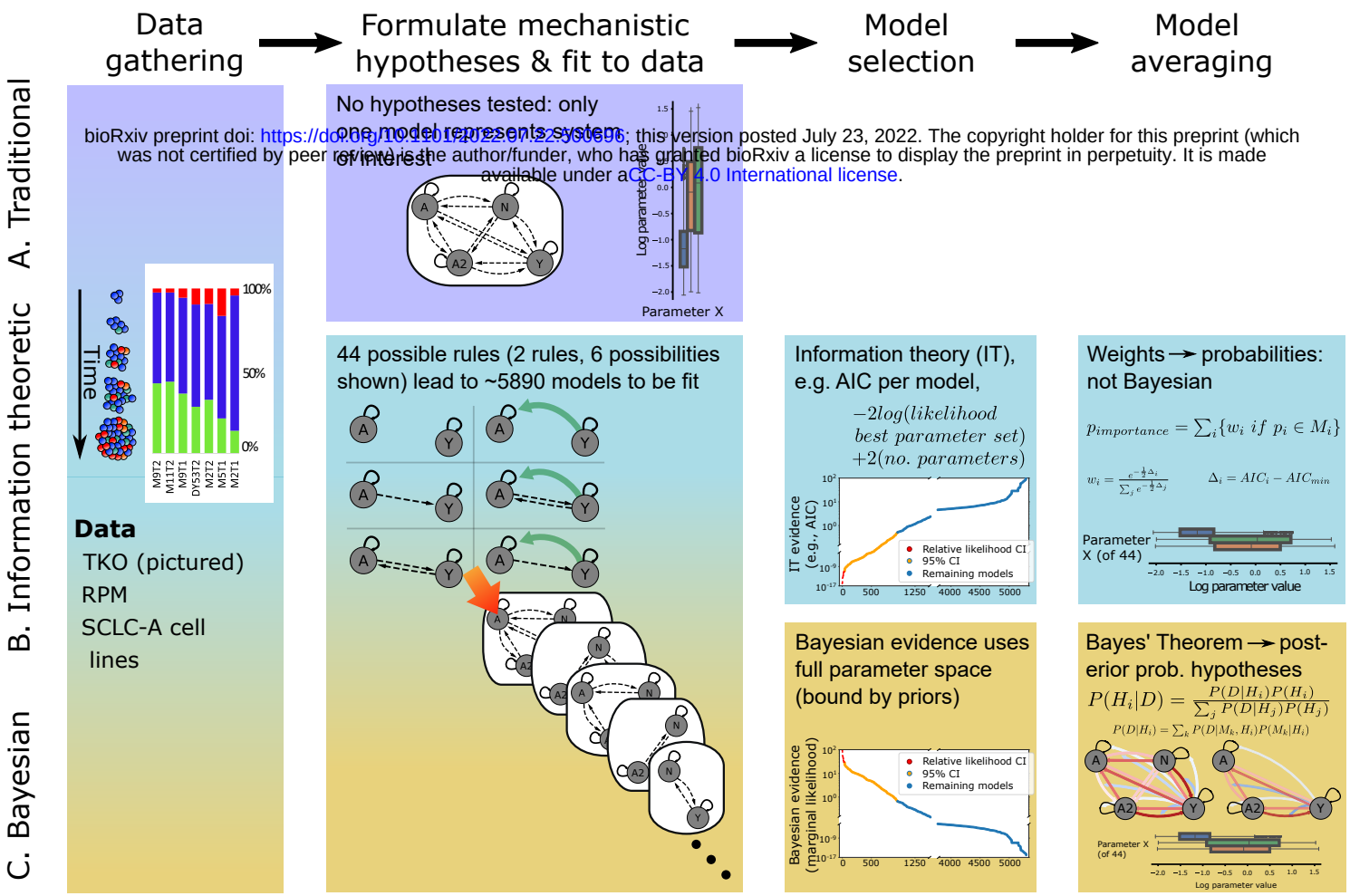
The Bayes Factor (BF) between one hypothesis and another indicates a true difference between the probability of each hypothesis when BF > $10^{1/2}$: the value of $10^{1/2}$ (~=3) is the lowest at which a difference may be determined (Kass and Raftery, 1995).

For model selection, the part of MMI addressing the likelihood that a candidate model represents the underlying process that generated the data, the hypotheses H in (Eq. 1) represent candidate models M within the set of all models $\mathcal{M}$ (see **Figure 2**), and data D represents each dataset to which each model is fitted (see **Figure 1**). Nested sampling calculates the marginal likelihood that each model represents the data, thus we can use Bayes Theorem to calculate the posterior probability that a model M_k best represents the data (**Figure 3**) (see Methods).

Given that our candidate models incorporate the range of possibilities of SCLC behavior, (**Figure 2**) we can use the marginal likelihood of each candidate model to calculate the posterior probability of each hypothesis (see **Table 1** for biological hypotheses). This can be done via Bayesian model averaging, (BMA) the part of MMI addressing the likelihood that a model variable is supported by the data. In this case, the likelihood $P(D|H)$ for a hypothesis H_i is represented by:

$$P(D|H_i) = \sum_k P(D|M_k, H_i)P(M_k|H_i) \quad (3)$$

where $P(D|M_k, H_i)$ is the marginal likelihood of a model M_k if M_k falls under the hypothesis H_i, and $P(M_k|H_i)$ is the prior probability of M_k across all models falling under H_i (Burnham and Anderson, 2002; Fragoso et al., 2017). Incorporating $P(D|H_i)$ in (Eq. 2) back into (Eq. 1) enables us to calculate posterior probabilities (**Figure 5**) for each hypothesis and Bayes Factors comparing competing hypotheses (**Table 2**).



In a traditional approach (A), a single model is selected to represent a system based on an investigator's best reasoning. Parameter fitting is performed (see Box 1). Resulting fitted parameters for the sole model in this approach may be biased: if a mechanistic process (model reaction) is wrongfully included or excluded, a fitted rate might be skewed to accommodate the inclusion/exclusion by driving simulations toward a better fit to the data.

In an information theoretic (IT) multimodel inference approach (B), known features of a system are used to generate candidate models with differing reaction rules. All candidates are fit to the same data. An information criterion (IC) such as Akaike Information Criterion (AIC) (the IT example in this Box; see (Dziak et al., 2020) for a discussion of IC in model selection), is calculated per model based on best-fitting parameters and parameter number. A confidence set of models can be chosen based on differences between AICs, or "weights" (Burnham and Anderson, 2002). Model averaging may be performed, including calculation of parameter importance.

Bias may play a role in several areas of the IT process. Firstly, IC weights are commonly determined after optimization via sampling algorithm like Markov Chain Monte Carlo, which only partially explores parameter space (van Ravenzwaaij et al., 2018). Thus, a confidence set based on IC weights misses information, not accounting for the parameter space's full complexity. Next, parameter importance is commonly estimated by summing model's weights where a parameter appears, then dividing by the sum of all model weights, approximating a probability (Burnham and Anderson, 2002; Galipaud et al., 2014). Galipaud et al. caution against interpreting sums of weights (importance values) as probabilities, as they do not accurately represent variable effect on simulation outcome (Galipaud et al., 2014, 2017). Finally, importance values are difficult to interpret if there are more candidate models with a given feature than without, (Burnham and Anderson, 2002) resulting in inflated feature importance. Attempts to resolve this bias have included assigning prior weights based on candidate model similarity (Garthwaite and Mubwandarikwa, 2010). We posit that we have solved the above issues in part, by using marginal likelihood values from a nested sampling algorithm rather than approximations like AIC. Marginal likelihood replaces weights and the "importance" calculation results in a posterior probability by definition of Bayes' Theorem. Parameter importance inflation can be solved by implementing the appropriate prior probability in Bayes' Theorem.

Thus, in our approach (C), nested sampling fits candidate models to data, saving all parameter set scores to later use for estimating marginal likelihood, or Bayesian evidence (see Box 2; Methods). Model averaging then incorporates uncertainty into posterior rate parameter distributions via Bayes' Theorem (Fragoso et al., 2017). Parameter posterior probability (rather than importance) is calculated, resulting in the posterior probability of each model feature given prior probability and data. In this way we use the principles of Bayesian inference to incorporate uncertainty into our conclusions.