Cancer as a complex adaptive system

Mathematical models of tumor ecology, evolution and development

Guim Aguadé Gorgorió

TESI DOCTORAL UPF / Year 2021

THESIS SUPERVISOR
Ricard Solé
Department Ciències Experimentals i de la Salut (UPF)
Complex Systems Lab

Agraïments

Al Jordi, per fer-me sentir a casa des del primer dia al CSL, a l'Aina per donar un punt d'humanitat en aquest món de robots, al Luíño per la passió infinita, al Raúl, que (encara) no ha acabat el món, al Víctor, per aquells dies a Santa Fe i aquest últim any de tesi. Al CSL, on he après gairebé més sobre evolució, ecologia, complexitat, biologia sintètica, neurociència o teoria del llenguatge que sobre càncer. A la June i en Quim, per la il·lusió. Al Josep Costa, per la paciència, el coneixement i la visió. A l'Stu per la confiança i per *Origins of Order*. A l'Oliver Sacks, que gairebé em fa canviar el tema de la tesi. I a en Ricard, per oferir-me un lloc al CSL on poder tenir tot l'espai per a pensar, per recordar-me sempre la importància d'una bona pregunta, per no haver oblidat el per què de fer ciència (en aquest món polític), per ensenyar-me a veure la realitat des d'un lloc nou, el dels sistemes complexos, per la curiositat i l'escepticisme, per l'estada a SFI, per tot el que has pensat i tot el que m'has fet pensar.

Al Pablo, per ser-hi sempre, tots aquests anys, al Lluís, per dur-me fins aquí i a l'Adri per espavilar-me quan m'apagava. A n'Eva, pels millors anys d'universitat i a n'Elvira, per les converses de doctorats, muntanyes i ecologia. Al Luigi, la Carolina, el Loïc, el Luís i la Lourdes, per Lisboa, i a l'Emma, també, i per retrobar-nos a temps. A la Irene per la tendresa, al Nil i la familia de Sant Pere 29. A la Jess (i a la Bruna, que va néixer ahir!), a l'Estela, per cada cop que tornes, als Davids, pel matsoni, per tants anys junts, una mica més vells. A l'Andrea, per fer-me prendre partit. Al Pallars, a Pilar i Marta, per recordar-me el camí. A la Sònia, per quedar-se aprop. Al Malet, pel Guillem i pel retrobament després de 15 anys, que no tot ha canviat. A les cases on he viscut durant la tesi: a la familia de Can Palet (la meva primera casa) per ensenyar-me a conviure, pels dijous al vespre, per tant carinyo en moments difícils, a casa Relotger, i a la familia de Casa Agnés per la il·lusió enmig de tanta incertesa. A l'Eva, que ens hem fet experts en tumors mucinosos. A l'Elisa, que vas aparèixer quan això començava, per fer-me viure d'una manera nova, a la Joana per la pau. A la Núria, en Xavi i en Roger per les parets d'aquests últims mesos, al Puxi per les muntanyes, a en Richi pels vespres pasturant el ramat a la posta de sol de Vilamur. A la Júlia, per tot el que aprenc i visc amb tu, que és tan net, bonic i fàcil.

I a casa: Jaume, Núria, Júlia. Sobretot a casa.

Gràcies per ser-hi.

Abstract

Despite decades of scientific effort, cancer remains a major cause of death worldwide. Through the accumulation of genome alterations, tumor populations evolve the capacity to circumvent the selective barriers of tissue homeostasis, eventually adapting to resist therapeutic stress. Furthermore, extensive Darwinian evolution is accompanied by an ecological engineering of the surrounding tissue micro-environment together with the alteration of cellular maturation hierarchies. To understand cancer complexity, therefore, we need a picture that spans through the domains of ecology, evolution and development. In an effort to gain understanding of the underlying patterns of treatment resistance, we introduce a mathematical approach to cancer complexity that takes into account its dynamical nature across these three axes. The resulting modeling endeavor is focused on two major fields of current research: immunotherapy and cancer epigenetics and differentiation, with the aim of providing both insight in treatment design rationale and a comprehensive perspective able to merge cancer ecological, evolutionary and developmental complexity.

Resum

Malgrat dècades d'esforços científics, el càncer continua sent una de les principals causes de mort arreu del món. Mitjançant l'acumulació d'alteracions del genoma, les poblacions tumorals evolucionen la capacitat d'eludir les barreres selectives de l'homeòstasi del teixit, fins al punt d'adaptar-se per resistir l'estrès terapèutic. A més, l'extensa evolució darwiniana s'acompanya d'una enginyeria ecològica del teixit circumdant, juntament amb l'alteració de les jerarquies de maduració cel·lular. Per entendre la complexitat del càncer, per tant, necessitem una imatge que abasti els dominis de l'ecologia, l'evolució i el desenvolupament. En un esforç per comprendre els patrons subjacents de resistència al tractament, introduïm un enfocament matemàtic de la complexitat del càncer que té en compte la seva naturalesa dinàmica en els tres eixos. L'esforç de modelització resultant se centra en dos grans camps de la investigació actual: la immunoteràpia i l'epigenètica i la diferenciació del càncer, amb l'objectiu de proporcionar una visió fonamental del disseny del tractament i una perspectiva integral capaç de combinar la complexitat ecològica, evolutiva i del desenvolupament del càncer.

Contents

List of Figures

1 Introduction

1.1 Cancer as a complex disease

Cancer is a major cause of death worldwide. By 2018, it was estimated that about 9% of females and 13% of males will die from any invasive form of the disease [Bray et al., 2018], making it one of the main causes of human death, to be only surpassed by heart disease [Jemal et al., 2011]. Furthermore, the overall number of cancer-related deaths is still increasing, mostly in relation to lifestyle changes [Jemal et al., 2011]. In this context, the most important risk factor for developing cancer is age itself [Coleman and Tsongalis, 2009], further increased by tobacco use or other habits or environmental factors. Eradication of other fatal diseases, such as smallpox [Riedel, 2005], has led to a general increase in age expectancy, with a consequent increment in the incidence of cancer around us. Has our understanding of the disease, and the consequent availability of efficient treatments, followed this trend?

Cancer has existed for all of human history and was considered, for centuries, an incurable disease [Mukherjee, 2010]. It was not until the 19th century that surgery became a first line treatment, but was later proven to be efficient only for small solid tumors with well-defined borders. For more advanced and spreading cancers the story takes a dramatic shift. Two thousand years ago, the Roman physician Celsus already observed that cancers would return even after excision [Hajdu, 2011], an evidence later followed by similar observations of tumor relapse and spreading across different treatment scenarios [Sudhakar, 2009]. This meant that, through some complex and unknown mechanism, cancers were able to survive medical attack.

Following scientific and economic efforts of the so-called *War on Cancer* [Rettig,

2005], the last fifty years have seen an improvement in both the understanding of oncogenesis and the available therapeutic protocols in place. Despite treatment for some cancers has seen significant progress, such as for some childhood leukemias [Kersey, 1997], prognosis after the diagnosis of other tumor types remains very poor after more than 40 years of intensive research. Thus, the search for a *magic bullet*, a single drug that finally eradicates the disease, is nowadays known to be, if anything, a rewardless endeavor [Strebhardt and Ullrich, 2008]. The overall trend for those cancers that are not cured remains in place: initial therapy succeeds in reaching disease regression, even to the point of no detectable tumors after initial interventions. Most advanced cancers, however, return after a given period of time, and often in the form of a more aggressive, rapidly spreading disease [Dagogo-Jack and Shaw, 2018].

So, what is cancer really, and why is it so difficult to find efficient treatment options to cure it? Cancer is a disease of abnormal growth of our own cells and tissues. Accumulated genetic alterations in otherwise-normal cells enable the acquisition of novel, *hallmark* capabilities [Hanahan and Weinberg, 2011] such as the evasion of growth-suppressing signals, activation of replicative immortality or the formation of novel blood vessels – *angiogenesis* (Fig. 1a). The activation and accumulation of these, together with other tissue-specific transformations, induces a single cell to escape multicellular control, forming a population of undifferentiated, minimal replicators [Sole et al., 2014], a *tumor*, that will invade surrounding tissues and organs, eventually killing the host if unstopped.

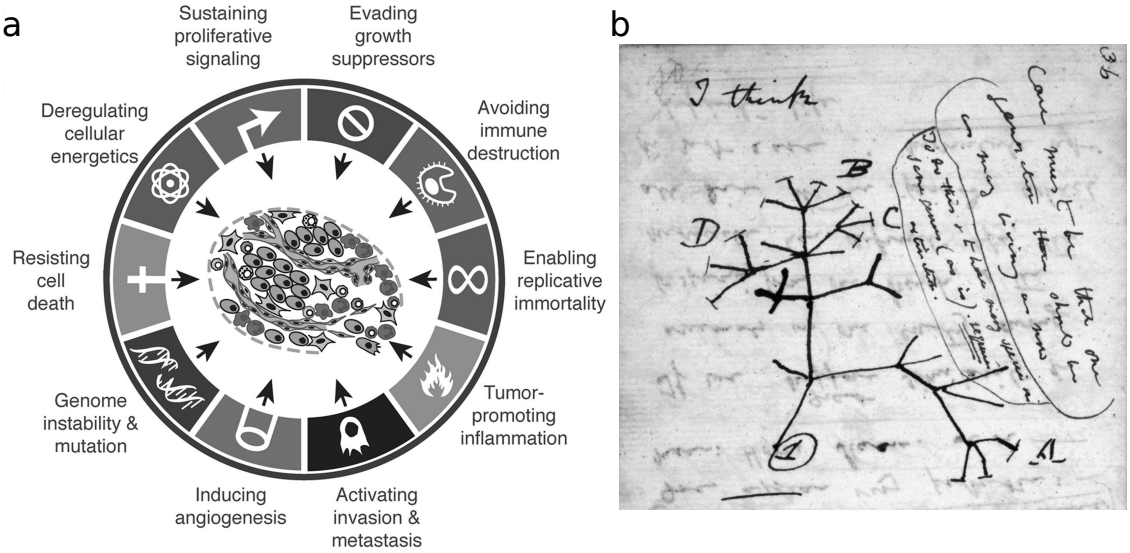

Figure 1: **The hallmarks of cancer and Darwinian speciation.** Presented in 2011 by Hanahan and Weinberg [Hanahan and Weinberg, 2011], the *Hallmarks of cancer* **(a)** comprise a set of ten universal characteristics shared across cancer types, thus providing a comprehensive picture of what tumors are. These characteristics consist of minimal replicator properties of single cells (sustaining proliferative signaling, enabling replicative immortality,...) and alterations at the tissue level (inducing angiogenesis, tumor-promoting inflammation,...) all enabled by the accumulation of DNA defects leading to unstable genomes. Evidence accumulated during the last decades indicates that cancer, and the corresponding acquisition of Hallmarks, is a disease of Darwinian evolution and speciation [Darwin, 1859, Greaves and Maley, 2012] **(b)** .

In an effort to stop tumor progression by inhibiting the hallmark capabilities, clinicians deal with a very large landscape of possible cancer therapies, that target anything from cell death [McComb et al., 2016] and DNA stability pathways [Chabner and Roberts, 2005] to inhibition of signaling cascades [Shawver et al., 2002], activation of immune surveillance [Yang et al., 2015] or the differentiation process of cancer cells [de Thé et al., 2018], among many others discussed elsewhere. Most of these therapeutic schemes do function, meaning that they effectively kill rogue cells or inhibit the expected targeted pathways of malignancy, and even more specific and precise drugs are expected to arise from future research.

It is not drug failure, then, what drives tumor relapse, or at least, not exactly. Evidence for most therapeutic schemes indicates that a first drug hit is often followed by a period of apparent lack of disease [Aguirre-Ghiso, 2007], that will eventually be disrupted by the cancer coming back, in an altered, more aggressive form. Understanding the major processes underlying tumor relapse remains a major element of cancer research. As a starting point, two main options here include the possibility that cancer cells become resistant in the presence of the drug [Misale et al., 2012], or else that, similar to resistance in pest management [Gatenby and Brown, 2018] or bacterial drug tolerance [Balaban et al., 2004], subpopulations of resistant cells were already present in the tumor prior to therapy [Dagogo-Jack and Shaw, 2018].

By the 1940s, long before the evolutionary basis of cancer started to be understood, Max Delbrück and Salvador Luria designed an experimental setting to understand the basis of Darwinian selection in bacterial resistance to stress [Luria and Delbrück, 1943]. What they were testing, in fact, was the same questions oncologists faced decades later: do single-cell agents resist drugs due to randomly accumulated preexisting mutations, or do they evolve a drug-tolerant phenotype in the presence of the toxic substance? For Luria and Delbrück, the answer indicated that resistance mutations existed previous to the drug, clearly proving the notion of Darwinian selection acting on randomly distributed, preexisting genome heterogeneity.

We know nowadays that the mechanisms for drug resistance in cancer are rooted in similar principles [Dagogo-Jack and Shaw, 2018]. However, as discussed along the present thesis, the vast and open possibilities of cancer adaptation underline a much more complex scenario, where many layers intervene, so that resistance can be the result of an adaptive process, of neutral stochasticity or an ecological or developmental cue. The present thesis focuses on gaining insight into these mechanisms with the aim of designing successful treatment approaches.

To understand the complexity of cancer drug resistance, we need to understand what is really happening during the process of oncogenesis. To accomplish that, we must consider a multi-layer picture of cancer, where the single-cell molecular perspective is broadened by considering cancer complexity at the genome and tissue levels. At the DNA level, tumorigenesis results from the following process: As tumors grow, accumulation of DNA damage and alterations result in cancers harboring a very large array of cells with different genomes [Marusyk et al., 2012]. As discussed below, the footprints of genome instability pervade the whole packaging structure of the cellular DNA content, with evidence of single-nucleotide variations, altered chromatin configurations and whole chromosome rearrangements making each cancer cell eminently unique.

This extremely diverse tumor, however, does not live in isolation. Organ multicellular-

ity, maintained by the precise fine-tuning of tissue homeostasis, implies that cancer grows in a complex ecosystem [Horning, 2017]. In it, coexistence with other cellular components such as the immune system, or multiple spatial and nutrient availability constraints, build up an array of selective pressures, different for each body site, that constantly drive cancer cells towards the limits of viability. At this point, we can descry a very heterogeneous disease, built upon multiple genomes that code for different phenotypic subpopulations, constantly facing the selective pressures induced by a complex and dynamical environment. As for the Darwinian origin of species, that results from a process of variation and selection [Darwin, 1859] (Fig. 1b), cancer is, briefly put, a disease of evolution [Greaves and Maley, 2012].

Even so, ecological interactions and Darwinian selection do not explain all processes seen across cancer types [Sell, 2004]. As a landmark example, over the last decades increasing evidence points to the existence of cellular hierarchies in tumors. In healthy tissue development, only a subset of *stem cells* differentiate into a wide array of different phenotypes [van der Kooy and Weiss, 2000]. Similar stem-driven architectures have been observed in tumors, resulting in the *Cancer Stem Cell* model of oncogenesis, with deep implications in therapy design [Clarke and Hass, 2006]. On top of that, recent evidence indicates that epigenetic changes in chromatin permissiveness enable extensive phenotypic plasticity, resulting in an additional layer of cancer evolvability [Flavahan et al., 2017].

In the light of this, cancer has to be understood as an ecological, evolutionary and developmental problem. To capture it, we need to harness previous knowledge from these areas of science. In all of them, mathematical modeling of complex systems has been a key framework to understand the fundamental laws governing their dynamics and structure. Examples of this are ecological networks to capture ecosystem architecture and stability [Sole and Montoya, 2001], the Quasispecies theory for the evolution of unstable replicators [Eigen and Schuster, 1977] or the Waddington landscape of cellular differentiation and epigenetic development [Waddington, 1957]. How does our understanding of cancer change when seen through the lens of complex systems? And, more importantly, can these mathematical models help us envisage novel therapeutic schemes that overcome cancer resistance?

1.2 Darwinian evolution in cancer

> *Everything: the minutely detailed history of the future, the archangels' autobiographies, the faithful catalogues of the Library, thousands and thousands of false catalogues, the demonstration of the fallacy of those catalogues, the demonstration of the fallacy of the true catalogue, the Gnostic gospel of Basilides [...]*
>
> *The Library of Babel*, J.L. Borges (1941)

In 1941, Jorge Luis Borges captured in a beautiful short story the conflicting similarity between the incomprehensibly vast and the infinite. In his short story *The Library of Babel* [Borges, 1941], Borges envisions a continuous, almost periodic library (Fig. 2a) that contains all possible books of 410 pages. In these precise books, each page has 3200 symbols, and there are 25 possible options for each (22 letters, the dot, the comma and the space). Simply put, the *Library*, which others call the Universe, contains

$$N = 25^{410\times 3200} = 1,956 \times 10^{1834097} \tag{1}$$

different books. How large is the space that contains them, and how vast the landscape of possible stories within their pages?

1.2.1 Complexity in the human genome

If cancer is a complex disease that takes advantage of genome modifications to explore adaptive opportunities, one may ask, as for books in the *Library of Babel*, what is the combinatorial potential of the human genome. The combinatorial landscape here entails a code of 4 symbols (the nucleotides) made of about

$$|G| \sim 3 \times 10^8 \text{ bp.} \tag{2}$$

In Borges' books, if each page has again 3200 symbols, the human genome would be 93750 pages long, and there would be

$$N \sim 4^{3\times 10^8} \tag{3}$$

different genomes, a number far beyond any library catalogue. A similar statement can be made about the genetic content that explicitly codes for proteins through an optimal genetic code [Adami, 2004] (Fig. 2b). Consider there are about

$$N_g = 2 \times 10^4 \tag{4}$$

genes in the human genome [Ezkurdia et al., 2014], which can be differentially expressed through specific chromatin configurations. Furthermore, genes interact epistatically within complex Gene Regulatory Networks [Kauffman, 1969]. In a simplifying picture, where genes are *on-off* switches of a Boolean network, the number of *attractors*, or possible stable states of protein expression, grows strikingly fast with system size [Samuelsson and Troein, 2003], *i.e.*

$$N_{att} \sim 2^{N_g} = 2^{2\times 10^4}, \tag{5}$$

making for a vast library of possible cellular phenotypes arising from each genetic configuration [Wagner and Zhang, 2011].

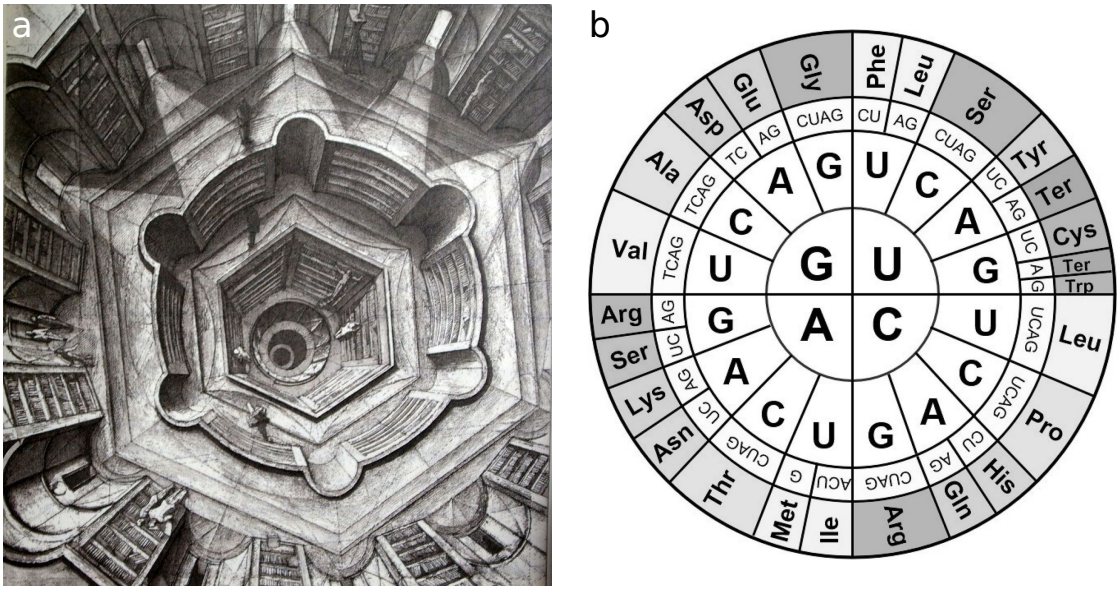

Figure 2: **The Library of Babel and genome complexity.** If all possible DNA (or corresponding messenger RNA) nucleotide genome configurations should be stored in a Borgesian library (as envisioned by Érik Desmazières [Desmazières, 1998] **(a)**), this would contain $N \sim 4^{3\times 10^8}$ different books, each of 93750 pages. In them, seemingly-infinite mRNA nucleotide configurations would encode proteins through a non-trivial dictionary **(b)**, and protein expression would in turn alter the state of other book sections, all regulated by chromatin configuration and chromosome copy numbers. The possible genetic makeups that result from genome aberrations are almost uncountable. How do cancer cells explore the confines of this complex landscape?

The architecture of the human genome, however, is build upon many layers of complexity far beyond genome sequences (Figure 3). The spatial 3-dimensional packaging of the genome within the cellular nucleus is known to play a non-trivial role in the overall coding of the phenotype, with chromosome configuration or chromatin folding altering the result of a given genetic makeup [Rowley and Corces, 2018, Zheng and Xie, 2019]. If the Library of Babel was strikingly large, how large would it become if not only symbols, but also the number of pages and their folding changed?

For human cells, this means that genome alterations could unleash an incomprehensibly vast landscape of possible expression patterns and, eventually, cellular phenotypes. Are there patterns in the seemingly random universe of carcinogenesis?

1.2.2 Universal footprints in cancer genome instability

Genes

The standard view of cancer is that it arises after the accumulation of genetic abnormalities in a single cell [Vogelstein and Kinzler, 2004]. This notion has shaped our perception of what cancer is and is not, up to the point of labeling cancer as a *disease of the genes* [Vogelstein and Kinzler, 2004]. Twenty years later, this picture has become too narrow, and many other aspects beyond genetic abnormalities are needed to understand the disease [Huang et al., 2009, Marusyk et al., 2012]. Mutations on specific genes remain, however, the enabling event of tumorigenesis [Vogelstein and Kinzler, 1993].

Three main gene families participate in tumorigenesis, namely *oncogenes*, *tumor suppressors* and *DNA stability* genes. Briefly put, mutations in these genes result in hallmark capacities of increased cellular division rate, lack of overall cell state control and loss of DNA repair machinery, respectively. Despite the apparent simplicity of this picture, and unlike diseases arising from a single genetic event such as cystic fibrosis [Zielenski and Tsui, 1995], cancer is not the result of a uniform pattern of genetic mutations. Even the most common tumor suppresor gene, *TP53*, the so-called "guardian of the genome" [Read and Strachan, 1999], is mutated in only about 39% of all cancers [Bielski et al., 2018].

In this heterogeneous landscape of possible single-nucleotide alterations, gene-specific mutations become less important when trying to look for patterns that are universal across cancer types. In turn, what becomes relevant here is the structural constraints of the mutational process itself. Evidence indicates that some cancers commonly evolve a so-called *mutator phenotype*, where defects in Mismatch Repair genes induce an incredibly high error rate that can be tens of thousands of times higher than that of healthy cells [Loeb, 2001]. A relevant question here concerns the possibility that cancer might approach evolutionary limits to mutation rates. This brought up the possibility that cancer cells could be somehow compared to viral quasispecies [Solé and Deisboeck, 2004]. In the Quasispecies Theory, evolutionary dynamics elicit the existence of a limit to the rate of genome errors per division, beyond which species fail at maintaining adaptive identity [Eigen and Schuster, 1977]. Experimental setups have proven that viruses, simple and highly mutational entities, live close to this critical region, where they are able to optimize genome variability without undergoing an identity meltdown [Solé and Elena, 2018].

Do microsatellite-unstable cancers, as viruses, live close to a critical mutational catastrophe? This is, still today, a partially unresolved question. Mathematical modeling indicates the existence of a well-defined error rate for cancer cells [Solé and Deisboeck, 2004], beyond which cell viability would be hampered due to excessive genetic mutations (Fig. 3 bottom). This has key implications on therapeutic strategies, since widely available mutagenic agents, if delivered correctly, could induce tumor arrest by overcoming this critical transition, thus pushing cells towards unviable overmutated phenotypes [Solé and Deisboeck, 2004, Fox and Loeb, 2010].

In the lack of recombination events, cancer cells accumulate DNA damage likely resulting in an increasingly unstable genome. If there are, indeed, catastrophic thresholds to genetic instability, how does cancer progress without trespassing them? A possibility here is that, as for other phenotypic characteristics such as cellular metabolism [Gillies et al., 2008] or perception of density constraints [Gerlee and Anderson, 2015], evolutionary dynamics govern genetic instability as an evolving trait itself. In this context, tumor

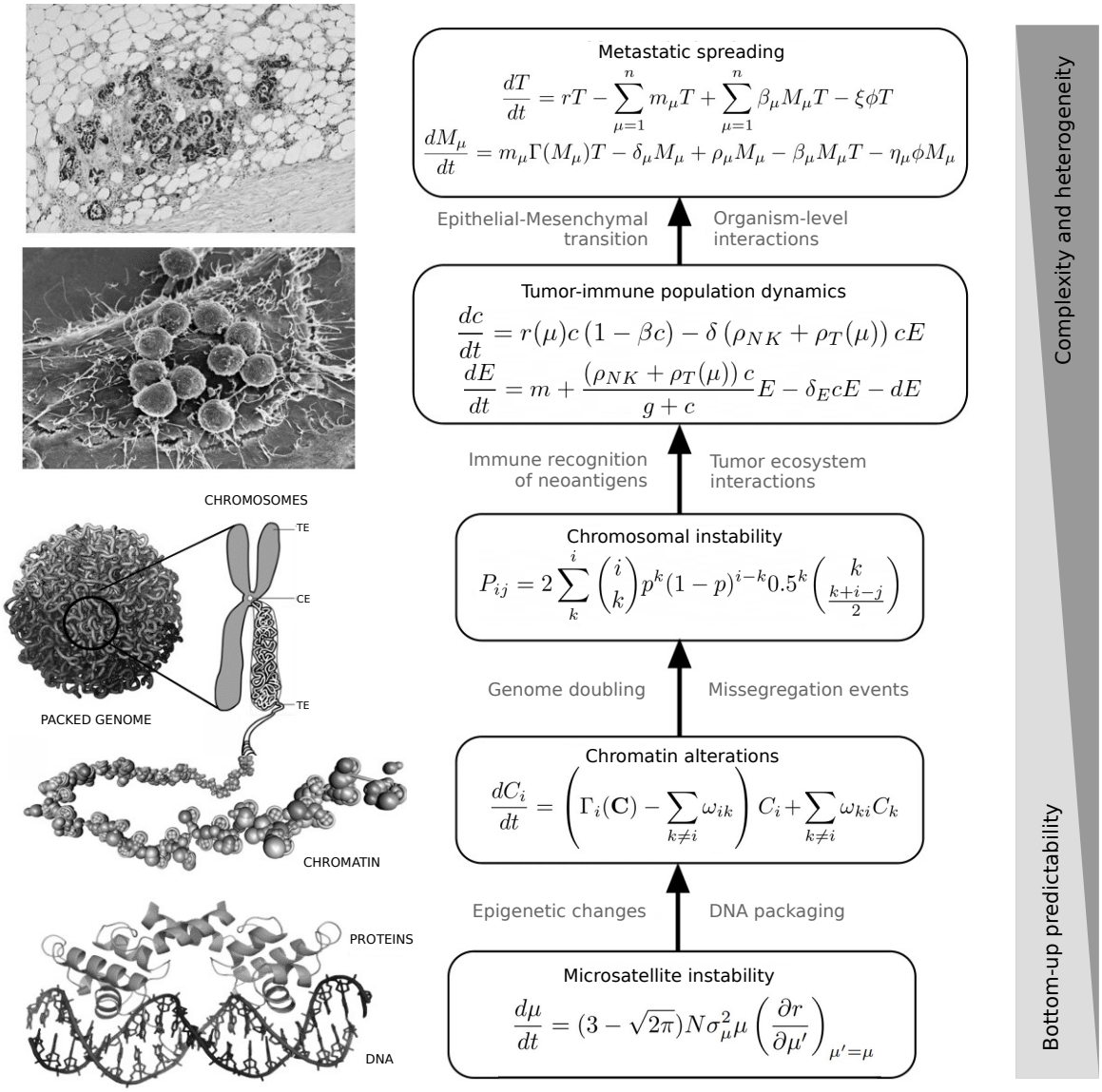

Figure 3: **Cancer instability across genome complexity layers.** Cancers explore the landscape of possible phenotypes through the accumulation of alterations across the different layers of the genome. Experimental evidence indicates the presence of instability footprints across (bottom-up) DNA sequences, chromatin configuration and chromosomes. Mathematical models can capture both the universal patterns and limits underlying these unstable dynamics. However, as in all complex systems, the dynamical motifs arising at each scale are usually not explainable by lower-scale models, so that each layer requires an approach of its own [Adami, 2002]. Single-nucleotide instability implies the evolution of so-called *mutator phenotypes* with error rates close to the error catastrophe [Solé and Deisboeck, 2004, Aguadé-Gorgorió and Solé, 2018]. Permissive chromatin states allow phenotypic switching as an alternative evolutionary mode [Aguadé-Gorgorió et al., 2020]. At the chromosome level, chromosome missegregation ensures stable copy-number distributions [Gusev et al., 2000]. The overall accumulation of genome alterations directly affects dynamics at the ecosystem level (here, neoantigen-producing mutations alter the cancer-immune interaction [Aguadé-Gorgorió and Solé, 2019]) and at the whole-organism level (chromatin plasticity facilitating epithelial-mesenchymal transitions precludes metastatic spreading [Gupta and Massagué, 2006]).

populations could optimize their mutation rate, while adapting to a given landscape, to maximize exploration while still maintaining genome identity in place. In the Results section, a possible modification to the Adaptive Dynamics framework [Diekmann, 2002] is presented, able to account for the evolution of unstable cancer populations by considering cellular mutation rates to be stochastic, evolving parameters [Aguadé-Gorgorió and Solé, 2018].

Chromosomes

At the other side of the packaging structure of the genome, opposite to nucleotides as the minimal coding element, lie chromosomes (Fig. 3 middle). More than a century ago, observation of chromosomes during cellular development and division in sea urchins led German zoologist Theodor Boveri to the notion that chromosomes where, in fact, the vectors of Mendelian heredity [Maderspacher, 2008]. Furthermore, Boveri showed that aneuploidy, the abnormal configuration of the karyotype, had a detrimental effect on organism physiology. In an astonishing scientific leap ahead of its time, Boveri also proposed that such abnormal chromosome configuration could be at the basis of oncogenesis [Holland and Cleveland, 2009]. A century later, evidence confirms that aneuploidy is, in fact, a very common feature in the aberrant genome of solid tumors [Sansregret and Swanton, 2017] (Fig. 4). However, it remains unclear whether aneuploidy is a cause or a consequence of the underlying processes of unstable tumor formation [Duesberg, 2007].

Once again, the *Library of Babel* provides a useful metaphor here. Imagine that cancer cells could not only alter nucleotide symbols and positions, but could also reshuffle or copy partial or complete chromosome sections. A *Library* with books that can also modulate page size and number would expand far beyond what is envisionable. The library, once more, would not be strictly infinite, as cells cannot harbor an unlimited load of DNA material. Together with that, most chromosomal translocations involving gene fusions would lead to novel genes unable to correctly code for proper proteins. Nevertheless, the evolutionary potential of aneuploidy as a means to both move through and expand an adaptive landscape is immense [Sansregret and Swanton, 2017]. Interestingly, as with the error catastrophe of single-nucleotide mutations, clinical and experimental evidence indicates that there might be a limit as well to chromosomal instability (CIN) that cancer cells, or the tumor as a quasispecies, cannot sustain [Sole et al., 2014].

Two landmark examples of CIN in cancer are Chronic Myelogeneous Leukemia (CML) and colorectal cancer subtypes with CIN. CML was, in fact, the first cancer to be linked to a specifically clear genetic abnormality. It is characterized by the translocations of chromosomes 9th and 22nd, resulting in a novel *fusion* gene, *BCR-ABL*, that encodes an "always-on" cell division signaling cascade [Kurzrock et al., 1988]. In such genetically-simple cancers, targeted agents that specifically inhibit gene activity –such as imatinib for *BCR-ABL*– can induce complete disease remission, resulting in a crucial step in the search for a cancer *magic bullet* [Quintas-Cardama et al., 2009]. However, the nature of aneuploidy is usually far more complex across cancer types. As an example, a large amount of colorectal cancers are known to evolve from premalignant adenomas through the so called *Chromosomal Instability* pathway. These CIN subtypes, however, encompass a large amount of different phenotypes characterized by a varying level of adaptation and resistance to therapy [Guinney et al., 2015]. Once again, pervasive inter-tumor vari-

ability recalls the need to find universal patterns that can explain the nature of aneuploidy in cancer.

These universal patterns can be found when comparing the average number of chromosome copies, *ploidy*, across different cancer types. The results are strikingly uniform. Despite variability in organ, cell of origin or patient characteristics, all cancers group around either a ploidy of 2 (no chromosomal instability) or $\sim$3.3 [Dewhurst et al., 2014, Bielski et al., 2018]. Does the pervasiveness of this ploidy of 3.3 indicate the presence of some attractor state? Mathematical models have proven that even neutral dynamics, where chromosome genetic content is not accounted for, could capture this stable ploidy value as an equilibrium between chromosome missegregation and the replicative cost of a large karyotype [Gusev et al., 2000]. However, results strongly depend on the heuristic definition of this trade-off, and the adaptive meaning of the average aneuploid attractor remains largely unknown.

Other open questions, regarding universal patterns of karyotype dynamics in cancer, remain open. Among them, a current evolutionary debate in oncology focuses in the prevalence of Whole Genome Doubling (WGD) in cancer cells [Dewhurst et al., 2014, Bielski et al., 2018]. About 30% of all cancers seem to undergo, early in the process of tumorigenesis, the duplication of their whole karyotype, making it for the second most common genome instability event after P53 mutations [Bielski et al., 2018]. Current research proposes that WGD could prevail as a buffer to the accumulation of deleterious mutations in recessive genes [López et al., 2020].

A more striking pattern arises with the prevalence of microsatellite (MSI) and chromosomal instability (CIN) across cancer subtypes. Current sequencing of cancer genomes uncovers that cancer cells either evolve through point mutations in their genes or alterations in their chromosome setup, but not both [Guinney et al., 2015, Bielski et al., 2018]. Is there some higher order constraint in evolution able to explain why apparently connected instability pathways seem to exclude each other?

The meaning of universal ploidy motifs, the pervasiveness of WGD events or the fact that MSI and CIN appear as mutually-exclusive events have been central areas of study along the present thesis. However, consistent results have not been reached by the time of the thesis presentation, so that no specific publication is presented regarding chromosome patterns in cancer. In spite of this, the overall role of CIN in cancer evolution is nevertheless discussed here, as a necessary consideration when trying to gain insight into the adaptive complexity of cancer (Fig. 3).

Chromatin and epigenetics

Only a mere 2% of the human genome is composed of genes that encode proteins, while 98% of the DNA content, previously considered junk [Doolittle, 2013], is now known to harbor elements that regulate context-specific gene activity [Van Holde, 2012]. All this DNA is wrapped around millions of nucleosomes, resulting in a massive molecular complex termed chromatin (Fig. 3, second from bottom, [Van Holde, 2012]). In between genes and chromosomes, chromatin is the genome architecture layer where transcription factors and signaling pathways alter gene activity. It is thanks to chromatin-based regulation that hundreds of different cell types in the human body can result from a single genome [Margueron and Reinberg, 2010].

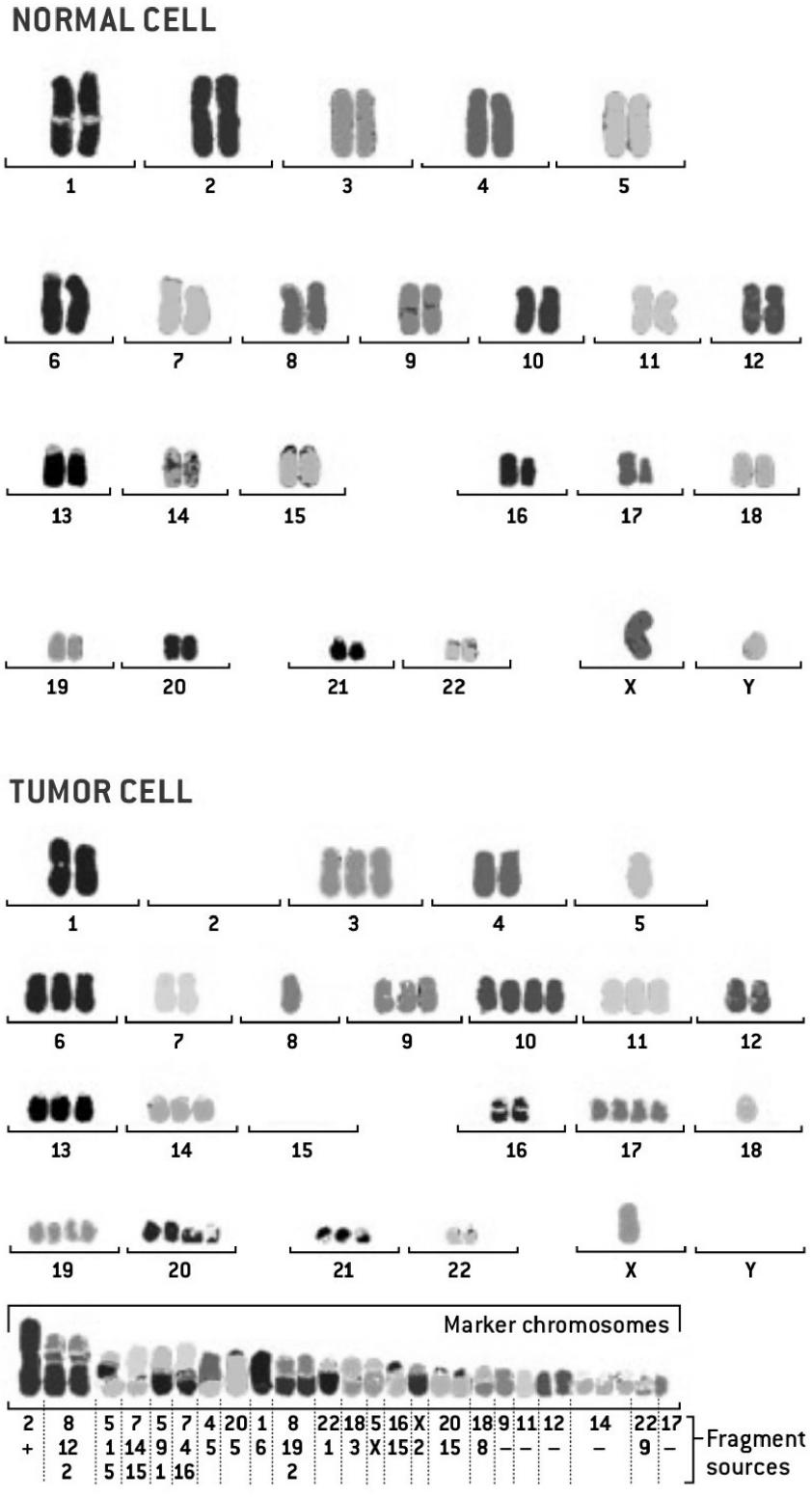

Figure 4: **Cancer aneuploidy.** More than a century ago, after observations of aberrant karyotypes in sea urchins, Theodor Boveri proposed that oncogenesis could result from chromosome mis-segregation. Still today, and despite evidence of pervasive aneuploidy in cancer genomes (In the image, karyotypes of a healthy cell and a bladder carcinoma cell, from [Duesberg, 2007]), the precise role of chromosomal instability in cancer evolution is not totally understood.

As for genes or chromosomes, the alteration of chromatin configurations is nowadays known to be a pathway for phenotypic variation in cancer [Flavahan et al., 2017], making epigenetic plasticity a novel layer of instability in tumorigenesis (Fig. 3, second from bottom). Some pediatric tumors, such as ependymoma (a tumor of the central nervous system) are a relevant example here. As discussed in the introductory section, cancer is mostly a disease related to age, as the probability of accumulating oncogenic mutations increases with number of cellular divisions and loss of DNA resilience [Bailar and Gornik, 1997]. This does not explain, however, the existence and prevalence of cancers in children of very young age, and even less for solid cancers with much more complex genomes than those of single chromosome-translocation leukemias. The stable genome of ependymomas, with no prevalence of somatic mutations or karyotypic variation, is nowadays known to be pervaded by epigenetic alterations in chromatin structure [Mack et al., 2014, Bayliss et al., 2016].

Briefly put, chromatin configuration acts on gene expression by adopting active or repressive states, dynamically helping or hiding genes from participating in the GRN in place [Margueron and Reinberg, 2010]. In cancer, genetic, environmental or metabolic signaling can dismantle the machinery regulating chromatin, resulting in overly permissive or restrictive sates [Flavahan et al., 2017]. This, in turn, can result in new oncogenic phenotypes, by either restricting healthy differentiation pathways (and thus leaving the cancer cell immature) or else allowing the overexpression of previously-silenced genes with cancerous potential [Flavahan et al., 2017].

Repressive states, most often characterized by DNA hypermethylation that blocks correct gene expression, resemble somehow the inactivation of gene function of single-gene copy errors [Baylin and Herman, 2000]. They are, however, much more plastic alterations, in that the nucleotide sequences are not changed and genes could, if reactivated, eventually recover their function [Baylin et al., 2001].

Permissive chromatin configurations represent, on the other hand, a whole new dynamical pathway to genome variation. Loss of chromatin repression not only allows for expression of previously silenced oncogenes, but may, by increasing the dimensionality and the number of attractors of the corresponding GRN [Huang et al., 2009], unleash high levels of plasticity between alternative phenotypes [Flavahan et al., 2017]. In this scenario, epigenetic alterations in cancer do not relate to the activation of specific phenotypes, as a single oncogenic mutation or chromosomal translocation does, but may directly increase the adaptive potential of cancer cells by changing the topology of the underlying phenotypic space.

What are the consequences of chromatin plasticity for tumor progression? Could a mathematical framework elucidate if there exist limits to epigenetic instability, as for genes or chromosomes? A key aspect here, discussed in *1.4 A caricature of tissue development*, is that tumors can, through epigenetic alterations, take evolutionary advantage from undifferentiated stem cell properties and tissue maturation hierarchies. As presented in the Results, mathematical models designed to understand the role of epigenetic alterations in aberrant hierarchies [Solé and Aguadé-Gorgorió, 2021] and plastic tumor structures [Aguadé-Gorgorió et al., 2020] reveal specific conditions for treatment success, different to those imposed by the accumulation of genetic or chromosomal alterations.

1.2.3 Evolutionary dynamics and therapy resistance

Heterogeneity

Cancer cells accumulate alterations across each structural layer of the genome. Similar to the experiments of Luria and Delbrück, instability footprints are mostly acquired in a random stochastic manner, often precluded by the loss of DNA machinery responsible for either correct DNA replication [Loeb, 2001], reliable chromosome segregation [Sansregret and Swanton, 2017] or homeostatic chromatin configuration [Flavahan et al., 2017] (Fig. 3). In other specific settings, mutational events can appear as a result of intense selective pressures, such as for the case of acquired adaptive resistance to immune or therapeutic attack [Sharma et al., 2017].

The first and foremost consequence of genome instability –and its multiple pathways– in cancer is the extensive heterogeneity of cellular genotypes found in a single tumor [Marusyk et al., 2012]. Furthermore, as for Darwin's speciation process, cells are grouped in *subclones*, behaving as cancer species each harboring a similar genetic makeup. Mutations shared across different subclones indicate an early appearance, thus providing a valuable backwards picture of the evolutionary process similar to that of to the Big Bang model in cosmology (Fig. 5) [Sottoriva et al., 2015].

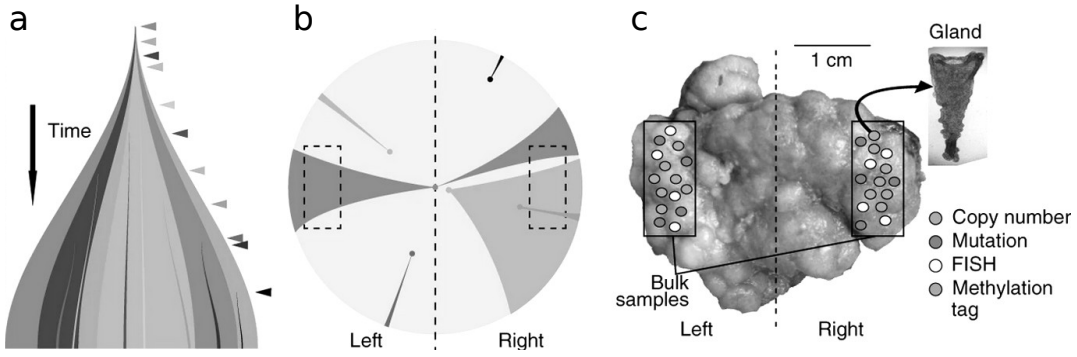

Figure 5: **The Big Bang model of tumor genome heterogeneity.** In a landmark paper by Sottoriva and colleagues [Sottoriva et al., 2015], the accumulation of mutations during cancer growth is seen as a Big Bag expansion **(a)** of shared (early) mutations and subclonal (late) mutations, from where a phylogenetic history of the tumor can be estimated. Mutation prevalence is often more related to when (and not where) the mutation occurred, owing to subclonal mixing in the primordial tumor **(b)**. In general, intra-tumor heterogeneity in cancer is likely to be subestimated in clinical settings, owing to DNA sequencing being limited to specific regions of a tumor **(c)**. In the light of this, the extent of heterogeneity can be better estimated with mathematical models of mutation and selection.

In specific settings, instability loads can be as high as to ensure that every single DNA locus can be found mutated in a given cancer [Loeb et al., 2019], precluding the notion that cancers harbor cellular subclones possibly prone to resist any single therapy [Diaz Jr et al., 2012]. This results in what is probably the most relevant aspect of a tumor: as compared to most human diseases, cancer is not a single entity, but a vast ensemble of different cellular populations that differ not only for each patient, but for each biopsy at each tumor site [Lawrence et al., 2013] (Fig. 5c). In this context, the likelihood that a

subclone will resist treatment rapidly increases as a tumor progresses, consistent with the decay in life expectancy of many advanced cancers [Dagogo-Jack and Shaw, 2018].

What is the nature and implications of the vast diversity found within a single tumor? The notion that evolution both results from and shapes tumor composition has led to a dramatic step forward in our understanding of cancer [Greaves and Maley, 2012], opening the door towards the possibility that cancer therapy and resistance could, as well, be understood from the perspective of Darwin's theory [Darwin, 1859] and its modern extensions [Huxley et al., 1942]. As for the evolution of ecosystem architecture [Valverde et al., 2018] or the fossil record [Eldredge, 2014], several (sometimes opposed) frameworks exist to describe the underlying evolutionary dynamics of intra-tumor heterogeneity (ITH), each solving specific aspects of malignant transformation. Their validity is often restricted to specific cancer types or tumor growth stages, thus asking for further research to fully understand the underlying adaptive dynamics of tumor progression.

The clonal selection model

In ecological settings, diversity precludes Darwinian selection by eliciting the survival of those individuals that are, within a given ecological setting, more prone to reproduction [Darwin, 1859]. Evidence of tumors harboring populations with different phenotypes led Peter Nowell to establish the theory of clonal evolution in cancer [Nowell, 1976]. In his paradigm-changing view, cancer was seen as the result of an evolutionary process where accumulation of genetic diversity was followed by selection and expansion of the fittest cancer clones, *i.e* ensembles of cells with the same genome or phenotype behaving as ecological species. Present day biology and genomics have proven that cancer is, indeed, a complex, Darwinian system [Merlo et al., 2006, Pepper et al., 2009].

The consequences that emerge from understanding cancer as a process of clonal evolution are many, with selection dynamics explaining some of the complex processes of oncogenesis, from the timescale of a disease that can take decades to develop to the acquisition of all the hallmark capabilities common across cancer types [Hanahan and Weinberg, 2011, Greaves and Maley, 2012] (Fig. 1). Within the evolutionary picture, our understanding of cancer therapeutic resistance has also taken a massive step forward. We now know that, as for other ecosystems, introducing a foreign substance that disrupts cellular viability also creates a selective pressure for the expansion of drug-tolerant phenotypes [Gatenby and Brown, 2018]. Cancer relapse after therapy is, therefore, a result of evolution.

In this scenario, there is an increasing interest in therapeutic approaches that take into account the evolution of resistant tumor subclones [Gatenby and Brown, 2018]. A promising example here is that of so-called *Adaptive Therapy* (AT) [Gatenby et al., 2009], a treatment rationale that aims at avoiding competitive release, the commonly observed growth of a resistant subclone after the sensitive population has been killed by therapy. By modulating drug dosing and scheduling with mathematical principles [West et al., 2020], AT aims at controlling tumor size by maintaining competition between drug-tolerant and sensitive subclones.

Similar to AT, many questions in cancer treatment resistance are grounded in the dynamical aspects of the clonal selection model. In particular, several considerations regarding the adaptive nature of the cancer-immune interaction need to be approached as a Dar-

winian process. Within this perspective, two mathematical frameworks able to account for the evolution of cancer subclones in the presence of a selective pressure (here induced by lymphocytic recognition and attack) are presented here (See Results, [Aguadé-Gorgorió and Solé, 2019, Aguadé-Gorgorió and Solé, 2020]). As for AT, the same dynamics that entail treatment resistance by selecting for drug-tolerant subclones can possibly be used to design efficient combination immunotherapy approaches [Aguadé-Gorgorió and Solé, 2020]. Besides clonal selection, the nature of these treatment approaches is also based on the underlying cell-cell ecological interactions involved [Strobl et al., 2021], discussed below in *1.3 The cancer ecosystem*.

Neutral evolution

There remain, however, several relevant aspects to cancer resistance that are not totally explained by the theory of clonal selection. As a landmark example, pervasive mutational signatures abound across apparently healthy genomes [Martincorena and Campbell, 2015]. In this context, it is unsure if selection alone can account for the abundance and diversity of mutations in sequencing efforts [Marusyk et al., 2012, Burrell et al., 2013], nor if all this genetic heterogeneity is functional to cancer adaptation. If it were not, neutral evolution [Kimura, 1983], and not stringent selection, would be the dynamical framework to understand tumor progression. Inspired by early models of population genetics [Ohta and Gillespie, 1996] and self-organized criticality [Bak et al., 1987], a simple mathematical model of neutral tumor growth has been able to consistently discover patterns of genetic drift in cancer sequencing data [Williams et al., 2016] (Fig. 6).

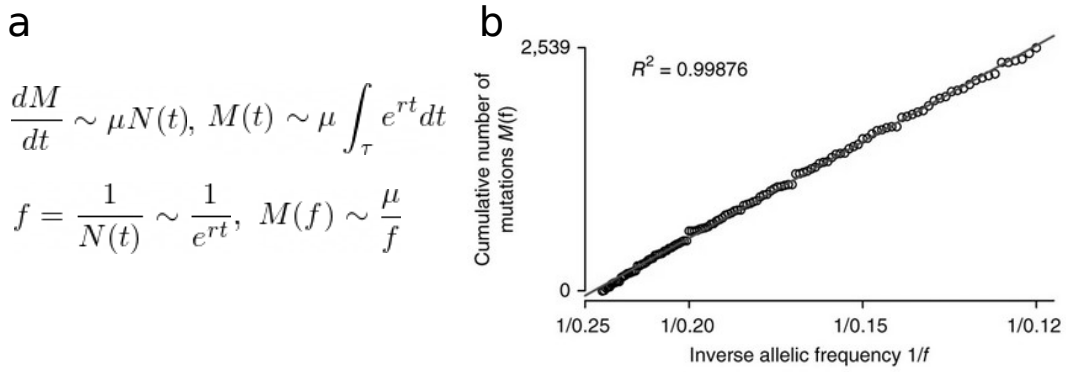

Figure 6: **Footprints of neutral evolution in cancer.** A minimal mathematical model [Williams et al., 2016] is able to elucidate that, if mutations are not selected for, their allelic frequency f must be proportional to the inverse of the population at the time they appeared, $f \sim N(t)^{-1}$. The number accumulated mutations M is, therefore, inversely proportional to f in the lack of evolutionary pressure **(a)**. Williams and colleagues found that the $1/f$ marker of neutral evolution was strikingly common across gastric cancer genomes **(b)** as well as in other cancer types.

As found in [Williams et al., 2016], many tumors accumulate mutations in a random and uniform manner, and only key driver mutations arising in founder cells are, therefore, shared across tumor lineages [Sottoriva et al., 2015] (Figs. 5,6). This is a key result when targeting tumor phenotypes with therapy: tumors that have not undergone purifying

selection harbor mutational heterogeneity to the point where most, if not all, necessary mutations for therapy resistance are probably present at the time of the first drug hit [Loeb et al., 2019]. Going back to the experiments of Luria and Delbrück, this means that, for any single therapy, at least one test tube in their experiment would always be populated by resistant bacteria [Luria and Delbrück, 1943]. In the light of this, mathematical efforts have elucidated combination treatments –the sequential use of different drugs– possibly able to overcome resistance by diminishing the probability that a single clone harbors multiple resistance mutations [Bozic et al., 2013].

Neutral evolutionary dynamics are likely to play an important role in one of the main subjects of study of the present work, namely the cancer-immune interaction. As discussed below, cancer cells often evolve the hallmark capacity of avoiding immune detection and attack [Hanahan and Weinberg, 2011, Sharma et al., 2017]. In this context of immune silence, mutations producing so-called *cancer neoantigens*[1] would no longer be deleterious [Lakatos et al., 2020]. Therefore, it is likely that neoantigen distributions evolve neutrally during tumor growth [Aguadé-Gorgorió and Solé, 2020]. Knowing that highly diverse neoantigen landscapes are likely to confound a sustained immune response [McGranahan et al., 2016], mathematical models of neutral evolution able to characterize the extent of neoantigen heterogeneity provide a powerful tool to predict immunotherapy prognosis (see Results, [Aguadé-Gorgorió and Solé, 2020]).

Once the immune system is back in place, the selective pressure to silence neoantigenic mutations is reactivated [Aguadé-Gorgorió and Solé, 2020]. Prior to that, however, mathematical models can capture how the dynamics of neutral (neoantigen) evolution are disrupted by using targeted agents affecting other cell characteristics [Aguadé-Gorgorió and Solé, 2020]. Interestingly, this puts up several opportunities regarding the use of combination immunotherapy to modulate, or even change, the dynamics of cancer evolution in favour of treatment success [Aguadé-Gorgorió and Solé, 2020].

Phenotypic plasticity

There exist several key phenomena in cancer adaptation that cannot be explained by clonal or neutral evolution theories. Two major examples here concern the existence of tissue hierarchies and phenotypic plasticity in tumors, reviewed thoroughly below (see *1.4 A caricature of tissue development*). Evidence for the first arose after cells with stem-like characteristics where found in cancers, indicating that tumor populations where separated into cells with and without tumorigenic potential among other phenotypic differences [Reya et al., 2001]. This means that heterogeneity, which we now know to preclude treatment resistance [Dagogo-Jack and Shaw, 2018], can be fueled by a (partially predictable) tissue-like maturation hierarchy [Marusyk et al., 2012], as opposed to the stochastic mechanisms of mutation-selection dynamics [Greaves and Maley, 2012]. The evolutionary consequences of the *Cancer Stem Cell* model, therefore, request a whole novel dynamical approach to treatment design [Meacham and Morrison, 2013].

Beyond cancer stem cells and hardwired maturation hierarchies, further evidence indicates that cells can activate phenotypic transdifferentiation pathways through non-mutational footprints [Marusyk et al., 2012]. Within this context, studies on drug resis-

[1] Mutational signatures in the surface of cancer cells that the immune system can recognize as *nonself* (see below, *1.3.2 Cancer and the immune system*) [Schumacher and Schreiber, 2015]

tance in cancer led to the notion that tumors were sometimes composed of plastic cellular states, able to *switch* to a drug tolerant phenotype in the presence of stress [Sharma et al., 2010]. As previously observed for bacterial populations [Balaban et al., 2004], evidence indicates that cancer could evolve phenotypic switching architectures able to maintain multiple states in place [Pisco et al., 2013, Neftel et al., 2019].

As discussed in sections *1.2.2* and *1.4*, instability at the chromatin level enables overall phenotypic plasticity, moving forward the possibility that cancer can evolve beyond the mutation-selection framework by taking advantage of aberrant cell developmental hierarchies and the vast possibilities of cell specialization and development [Huang and Ingber, 2007]. This represents a novel layer of disease complexity, where Darwinian dynamics are possibly overtaken by faster adaptive schemes. As for the previous evolutionary mechanisms, mathematical modeling provides again a relevant tool to capture the underlying roles of phenotypic plasticity in cancer resistance. In particular, we here present two alternative frameworks to elucidate how heterogeneity –and subsequent treatment resistance– can be a result of cancer stem cell hierarchies [Solé and Aguadé-Gorgorió, 2021] or phenotypic switching [Aguadé-Gorgorió et al., 2020] in given tumor types (see Results). As expected, models within each adaptive scenario indicate the presence of different threshold conditions for tumor success, highlighting how alternative evolutionary pathways entail different drug resistance processes.

Beyond developing novel drugs, overcoming therapy resistance is nowadays a main objective in cancer research. As briefly described, drug resistance is a multi-faceted mechanism, where evolution, mutational heterogeneity and phenotypic plasticity orchestrate a complex adaptive response. Two points arise from this conceptualization. First, we need to profoundly understand the complex evolutionary dynamics of cancer, their constraints and the interplay of several temporal and spatial domains in adaptation. Second, we need to design therapeutic interventions that consistently take this evolutionary framework into account in order to avoid cancer resistance. The specific shape of these therapeutic interventions, however, must follow from the selective pressures at play. These, in turn, strongly depend on the ecological interactions between cancer cells, surrounding tissue and the treatment in play. Evolutionary therapies for cancer need first to consider the complexity of the cancer ecosystem.

1.3 The cancer ecosystem

> *It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us.*
>
> *On the Origin of Species*, C. Darwin (1859)

Evolutionary dynamics govern cancer formation and progression. However, for the accumulation of heterogeneity in genome alterations to become a tumor, there has to be an ecological background of selective pressures and constraints that allow fittest phenotypes to progress [Greaves and Maley, 2012]. These barriers are, for the most part, a result of cancer cells facing the tightly organized architectures of organ homeostasis, that correctly maintains cellular growth and function in otherwise healthy tissue. In this picture, each of the the hallmark characteristics of the cancer cell [Hanahan and Weinberg, 2011] can be understood as an evolved mechanism to evade, suppress or overcome multicellular control [Sole et al., 2014].

As for most complex ecosystems, the interactions of heterogeneous cancer populations with surrounding cells, signaling molecules, soluble factors, the extracellular matrix or even mechanical cues are many and of variable sign [Balkwill et al., 2012]. Scientific progress in recent decades has helped us in understanding that there is much beyond a cancer *vs* host picture [Swartz et al., 2012]. In this context, interactions in the so-called *Tumor Microenvironment* (TME), as in Darwin's *entangled bank* [Darwin, 1859], are not restricted to cellular competition for space or resources [Tilman, 1982], but also include predator-prey dynamics [Barbosa and Castellanos, 2005], mutualism [Boucher, 1985], ecosystem engineering [Myers et al., 2020] or facilitation [Bruno et al., 2003].

The complexity of the cancer ecosystem can be visualized as a dynamical ecological network, where nodes represent cell types and oncogenic factors. The size and architecture of this network remains unknown, and is possibly built upon several heterogeneous layers [Pilosof et al., 2017] (Fig. 7). In this context, the outcomes that result from node removal or treatment that alters network topology are expected to be highly non-trivial, pointing towards the need of a methodical approach to the cancer ecosytem.

In this introductory section we present a brief outline of the specific ecological interactions of the cancer ecosystem that have been studied within the context of the present PhD thesis. This are, in general terms, the role of space and niches, the cancer-immune

interactions and the ecology of metastatic spreading, all discussed with special emphasis on the potential treatment opportunities involved.

1.3.1 Space and resources in tumor growth

Despite the complex molecular background underlying tumor growth, the observed kinetics, *i.e* the experimental shape of a tumor's growth curve, can be surprisingly simple [Benzekry et al., 2014] and where the first field of oncogenesis where mathematical modeling brought up novel insight [Steel, 1977]. Besides the predictive potential of growth equations, the fact that chemotherapy primarily acts on rapidly dividing cells implies that models capturing effective growth can be a very simple and preliminary predictor of prognosis [Swan, 1990]. In terms of differential equations, population growth in a cancer subclone c_i can be summed up to a minimal replicator model with birth b_i and death d_i rates

$$\frac{dc_i}{dt} = b_i c_i - d_i c_i, \qquad c_i(t) = c_i(0)e^{(b_i - d_i)t}. \tag{6}$$

However, a common and early finding in the dynamics of tumor size is that relative growth rates slow down with time [Collins, 1956]. In terms of differential equations, this means that there has to be an additional dynamical factor modulating birth b [Gerlee, 2013]. A basic and widely used approach here is the logistic or Levins growth model [Levins, 1969], characterized by introducing minimal modification to birth rates b that accounts for how subclone i is affected by the overall tumor metapopulation $\mathbf{c} = \sum_j c_j$

$$\frac{dc_i}{dt} = b_i \left(1 - \frac{\sum_j c_j}{K}\right) c_i - d_i c_i \tag{7}$$

This change alone accounts for an illuminating example of the role of basic ecological principles in cancer. Here K represents the carrying capacity of the system, a maximal population level beyond which the effective growth of the population becomes negative, thus setting an equilibrium total number of cells of $\mathbf{c}^* = K(1 - d/b)$. Despite more advanced models, such as the Gompertzian equation [Norton, 1988], offer an improved fit of growth curves, the fact that cancers progress under a carrying capacity-like limitation indicates that tumor populations are self-limited [Gatenby, 1991]. In ecology, this is often the result of spatial constraints or limited resource availability.

Spatial constraints participate in several ecological aspects of cancer growth (Fig. 8). As for many ecological settings [Hanski et al., 1999], three-dimensional spatial configuration of subclones has been proven to enhance diversity [González-García et al., 2002] (Fig. 8a) by deviating metapopulation dynamics from the simple *survival of the fittest* scenario [Sterelny and Turney, 2007]. As discussed along section *1.2*, this heterogeneity is a major factor of cancer resistance to treatment. Besides maintaining diversity, in such space-limited tumors cancer cells are also prone to escape density-dependent growth limitations by ignoring specific signaling pathways [Gatenby and Vincent, 2003]. Mathematical models here are able to capture the evolution of tumors with increased K-values and effective growth rates [Gerlee and Anderson, 2015] (Fig. 8b).

A further relevant aspect regarding cancer growth as an habitat-based problem concerns the existence of niches. As discussed below, evidence that cancer stem cells at the top of cancer differentiation hierarchies (see *1.4 A caricature of tissue development*) live

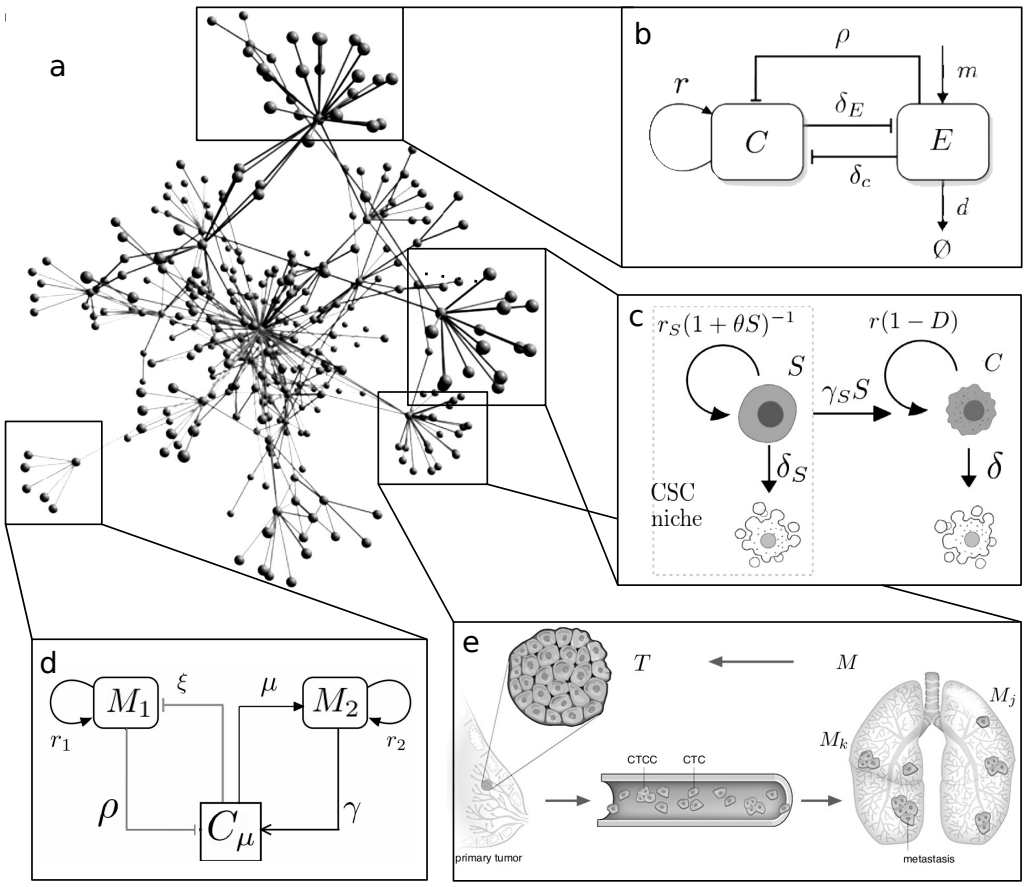

Figure 7: **Complexity in the cancer ecosystem.** As in many complex ecosystems, the dynamical interactions of cancer cells with the TME build up a complex network **(a)**. However, accurately describing the number of involved nodes (cell types) and their degree (number of interactions) and sign (interaction type) remains an elusive endeavor [Balkwill et al., 2012]. Within the ecosystem's building blocks, the cancer-immune landscape represents a landmark example of the non-trivial nature of the TME **(b,d)**. Minimal predator-prey models have been able to describe basic aspects of the dynamics between cancer and T cells [Kuznetsov et al., 1994] and the intrinsic role played by neoantigens [Aguadé-Gorgorió and Solé, 2019] **(b)**. However, evidence of macrophage polarization indicates that, beyond immune-based predation, education of immune components by cancer cells can result in mutualistic cooperation [Myers et al., 2020] **(d)**. Beyond strictly cell-cell interactions, spatial niches play important roles in cancer, such as for the Cancer Stem Cell niche in tumor hierarchies [Solé and Aguadé-Gorgorió, 2021] **(c)**. Metastatic spreading is another scenario where multiple ecological cues, such as invasion or colonization, underline a key aspect of cancer malignancy [Massagué and Obenauf, 2016] **(e)**.

in separated habitats [Borovski et al., 2011] implies that they follow a possibly alternative growth pattern to that of the principal tumor population (Figs. 7c, 8c). In the Results chapter, we propose a theoretical approach to the spatial constraints of differentiation therapy (DTH) [Solé and Aguadé-Gorgorió, 2021], built on early ecological models of habitat loss and fragmentation [Huxel and Hastings, 1999] (Fig. 8d). Interestingly, the model is able to highlight how the presence of a cancer stem cell compartment (and the underlying ecological niche) might render tumors resilient to differentiating compounds [Solé and Aguadé-Gorgorió, 2021].

Beyond space, another relevant aspect in cancer ecology regards the dynamical effects resulting from resource limitations. Cellular proliferation demands an increased import of nutrients from the immediate environment, mainly oxygen, glucose and glutamine [Pavlova and Thompson, 2016]. Within this context, tumorigenesis implies that a newly growing tumor is able to obtain enough energy blocks for uncontrolled replication, in an environment where resources are not infinite. In this context, cancer cells compete with surrounding cells, both healthy and cancerous, for nutrient uptake [Gatenby, 1991], creating an evolutionary pressure that selects for arising phenotypes that can overcome glucose limitations.

Two major cancer hallmarks arise as tumors grow and food becomes scarce [Hanahan and Weinberg, 2011]. When tumors grow beyond a size of about one cubic millimeter, the normal architecture of tissue blood vessels becomes insufficient for their supply [McDougall et al., 2006]. In this scenario, mutated cells actively secreting angiogenic growth factors such as VEGF [Carmeliet, 2005] are able attract novel capillaries, thus being selected by Darwinian dynamics against their starving competitors. The pervasiveness of this angiogenic switch across cancer types [Bergers and Benjamin, 2003] prompted the development of anti-angiogenic therapies, that have showed a marked benefit in sensitizing tumors to chemotherapy and controlling the extent of metastatic spreading [Jain, 2005].

A further oncogenic transformation in the absence of sufficient resources for uncontrolled proliferation is the reprogramming of the cellular metabolism [Hanahan and Weinberg, 2011]. Firstly observed almost a century ago by Otto Warburg [Warburg et al., 1931], it is nowadays clear that cancer cells can reprogram their glucose metabolism by limiting it solely to glycolysis, even in the presence of oxygen. This counterintuitive process (as glycolytic metabolism is an inefficient process in the presence of oxygen [Gillies et al., 2008]) could be a complex expression of facilitation dynamics, where glycolytic cells producing lactic acid maintain the energetic demand of a secondary subpopulation of acid-consuming cancer cells [Kennedy and Dewhirst, 2010].

All in all, evidence indicates that even the simplest demands of cancer cells as minimal replicators [Sole et al., 2014], *i.e* food and space for growing, are pervaded by ecological cues that influence malignant transformation. However, the previously discussed (Section 1.2.3) extent of intra-tumor heterogeneity (ITH) might not be explainable by considering intra-tumor competition alone. There are several possible –non-exclusive– explanations to ITH [Merlo et al., 2006]. Mutations could be evolutionary neutral [Williams et al., 2016], or else fitness gained from these mutations plateau as the cancer clone grows [Darch et al., 2012]. Other explanations raise that clones could evolve to occupy a specific niche [Pienta et al., 2008], either biological or geographical, thus alleviating direct competition, or that the surrounding environment, that shapes evolutionary pressures, is heterogeneous

in space or time [Marusyk et al., 2012].

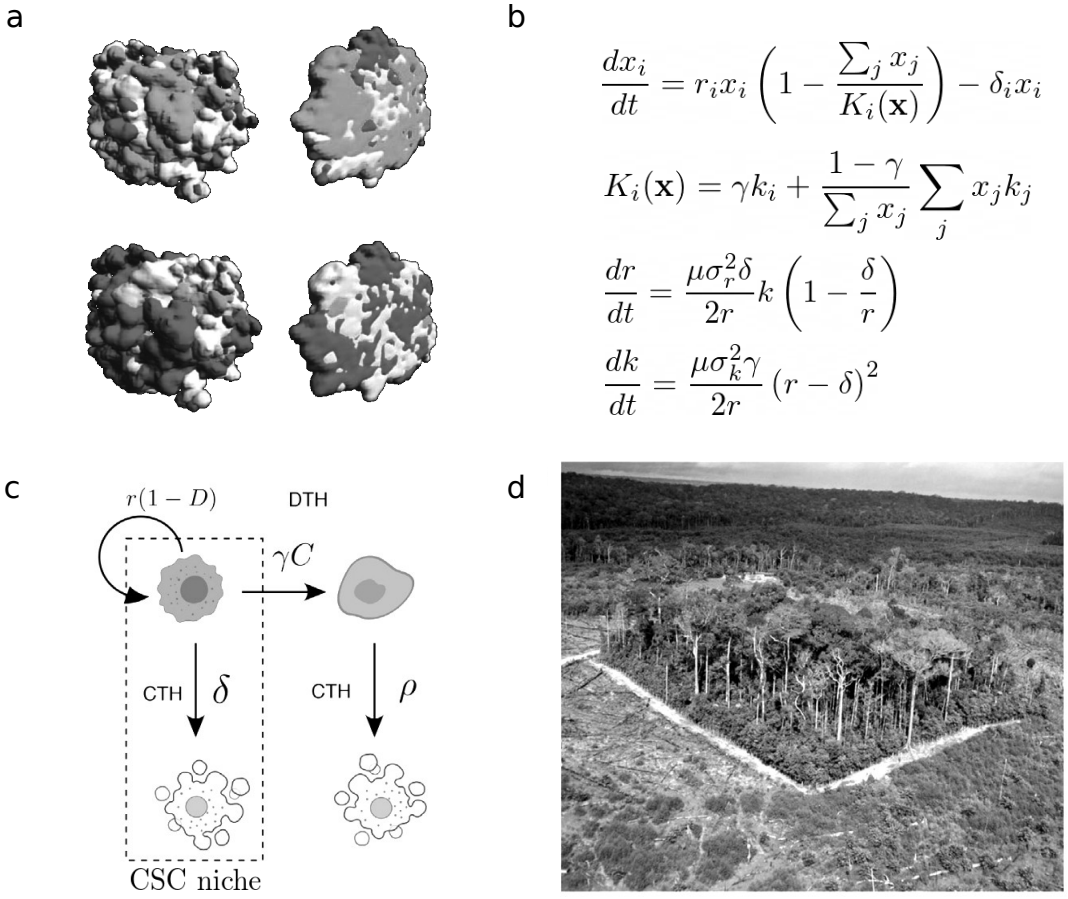

Figure 8: **Spatial constraints in cancer ecology.** Mathematical approaches have been successful in capturing several ecological cues of spatial limitations in cancer. Examples of these are three-dimensional simulations of tumor diversity in metapopulation models [González-García et al., 2002] **(a)**, adaptive dynamics modulating carrying capacities in tumor clones [Gerlee and Anderson, 2015] **(b)** and the interplay between the cancer stem cell niche **(c)** and habitat fragmentation models **(d)** [Solé and Aguadé-Gorgorió, 2021].

What other ecological dynamics exist beyond competition in the cancer metapopulation itself? As for many ecosystems, where biodiversity needs laws other than conflict to remain stable [Duffy et al., 2007], the presence of several non-cancer populations in a tumor indicates that these play non-trivial (and not always competitive) roles in tumorigenesis. Among the myriad ecological actors of the cancer environment, reviewed elsewhere [Mbeunkui and Johann, 2009], we consider here two main areas with specific –and crucial– ecological roles in cancer: the immune system – as a promising therapeutic opportunity– and metastatic spreading –as the key element of cancer aggressiveness (Fig. 3).

1.3.2 Cancer and the immune system

The immune system (IS) encompasses a giant collection of cells and chemical compounds that interact in a multi-layered complex network to provide us with a defense against external pathogens [Delves and Roitt, 2000, Subramanian et al., 2015] (Fig. 9a). As with the nervous system, the IS performs complex pattern recognition tasks and records a memory of previously encountered pathogens [Perelson and Weisbuch, 1997] (Fig. 9b), thus orchestrating the accurate responses characteristic of a *liquid brain*[2] [Solé et al., 2019, Pinero and Solé, 2019]. One of these cognition-based processes is the hallmark capacity of the IS to recognize external agents (the *non-self*) via identification of foreign antigens, and differentiate these from the myriad elements natural to the human body (the *self*) that need not be harmed [Medzhitov and Janeway, 2002] (Fig. 9c). Failure of correct discrimination is related to many immune-mediated diseases, with lack of self tolerance resulting in often serious autoimmune disorders [Sinha et al., 1990].

After correct *self-nonself* discrimination, the adaptive compartment of the immune system orchestrates a response to target and kill the external invader (Fig. 9b). But, if (human) cancer cells belong to the *self*, can a similar process take place in tumors? Back in 1893, William Coley discovered that bacterial infections could correlate with observations of remission in certain sarcomas [Coley, 1893], providing a first glance towards the possibility that the immune system, activated by an external pathogen, could target and kill cancer cells. Despite several decades-old theories establishing the basis of today's immunotherapy [Ehrlich, 1909, Burnet, 1957, Thomas and Lawrence, 1959], the underlying complexity of so-called *immune surveillance* (Fig. 9d) left immunotherapy somehow dormant for more than 100 years.

Modern evidence confirms that, indeed, immune effector cells, mostly T-cells and Natural Killer (NK) cells can recognize and kill malignant cells in our body [Miller and Sadelain, 2015] (Fig. 9d). A prominent discovery has been understanding how the mutational background of cancer cells generates a pool of so-called neoantigens, cell-surface peptides that fall in the *non-self* category and are eventually recognizable by immune cells [Schumacher and Schreiber, 2015]. Upon presentation of a given neoantigen by Antigen Presenting Cells (APCs) (Fig. 9c), helper T cells with the matching T Cell Receptor (TCR) release cytokines that generate a cytotoxic clone of effector T cells able to recognize and attack a tumor [Starr et al., 2003]. The possibility that the IS can be harnessed to target cancer has prompted an enormous endeavor towards disentangling the complexity of immune surveillance of tumors.

Early experimental and mathematical models described cancer-immune interactions as a predator-prey ecological system, with lymphocyte clones (the predator, T) grow by feeding on cancer cells with the corresponding antigens (the prey, c). Here the simple mean-field approach adds an equation for T cells:

$$\frac{dc}{dt} = b\left(1 - \frac{c}{K}\right)c - \delta cT - dc \tag{8}$$

$$\frac{dT}{dt} = m + f(\rho, c)T - dT \tag{9}$$

[2] As opposed to standard *solid brains*, where a static set of networks interact in persistent architecture, *liquid brains* comprise a variety of cognitive systems with a dynamical network of interacting moving agents. Ant and termite colonies, the immune system or microbiome communities are examples of such liquid cognitive networks [Solé et al., 2019].

with δ introducing immune-mediated killing, m the steady immune migration to the tumor site and $f(\rho, c)$ the rate at which T cells recognize cancer cells and start a growth cascade [Kuznetsov et al., 1994]. This system captured the dynamical background of cancer dormancy and remission (Fig. 3, 7b), confirming the hypothesis that these where immune-related phenomena [Hellström and Hellström, 1969, Kuznetsov et al., 1994]. In the light of cancer pervasiveness, it appears clear that the IS does not always maintain tumors at bay. In many cases, cancer cells seem to escape immune surveillance and inevitably progress into carcinomas. Immune escape is a hallmark example of how evolution, here selecting for preys able to elude their predators by altering either δ or $f(\rho, c)$, drives oncogenic transformation [Hanahan and Weinberg, 2011, Sharma et al., 2017]. In one of the possible mechanisms to immune escape, cancer cells activate immune checkpoints (such as CTLA-4 or PD-L1, Fig. 9d), molecular regulators of immune attack originally evolved to avoid collateral tissue damage in pathogenic infections [Topalian et al., 2015], thus dismantling the predator-prey system by setting $\delta \sim 0$. The discovery of drugs inhibiting the action of these pathways led to a dramatic revolution in the field of cancer immunotherapy [Hodi et al., 2010].

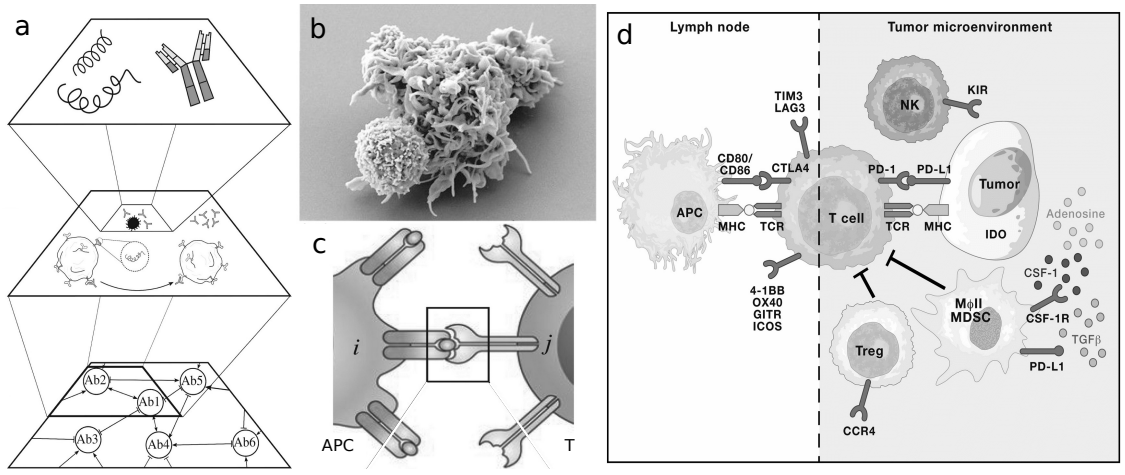

Figure 9: **The immune system and cancer surveillance.** The many layers of immune system complexity, involving the idiotypic network, multiple cellular populations or the specific epitope and surface receptor agents, provide the IS with *liquid brain*-type cognitive capacities [Pinero and Solé, 2019] **(a)**. Among these, a major task is the elimination of foreign pathogens **(b)** after recognition of *nonself* antigens **(c)**. In cancer, immune surveillance is built upon a complex network of molecular and cellular components **(d)**, where not only neoantigens, but immune checkpoints or protumorigenic immune cell types participate in the cancer-immune ecology [Sharma et al., 2017].

However, not all tumors remit after T cells are summoned by anti-PD-L1 or anti-CTLA-4 checkpoint inhibitors [Sharma and Allison, 2015]. Because T cells recognize mutated antigens, cancer types characterized by low somatic mutation burden are likely to remain silent to immune surveillance [Sharma et al., 2017]. This includes a wide variety of malignancies where adaptation follows from chromosomal or epigenetic instability (Fig. 3), posing a keystone limitation to immunotherapy efficiency. One of the primary endeavors of the present thesis has been the study of how do predator-prey models change

under an increment in neoantigen load resulting from induced mutagenesis. The crucible of the problem here is that genetic instability both activates cancer adaptation and immune recognition of neoantigens, thus establishing an evolutionary trade-off modulating the outcomes of system (8)-(9). In the Results chapter we propose a mathematical framework able to capture this trade-off [Aguadé-Gorgorió and Solé, 2019]. Interestingly, the model indicates specific therapy conditions for tumor clearence, where mutagenic or immunotherapy approaches alone are unlikely to success. Instead, results entail a possible a combination therapy approach able to success in complete tumor remission [Aguadé-Gorgorió and Solé, 2019].

In given settings, clinical evidence indicates that even cancers with high mutational burden might not elicit proper immune responses [McGranahan et al., 2016]. The underlying mechanism here might result from the complex laws of the TCR repertoire [Perelson and Weisbuch, 1997], implying that heterogeneity of tumor neoantigens is likely to exceed the capacity of the T cell *clonal selection* process to generate a correspondingly diverse pool of TCRs [Perelson and Oster, 1979]. Can we capture the nature of this heterogeneity limit? Continuing with research on the instability-surveillance interplay [Aguadé-Gorgorió and Solé, 2019] and inspired in previous models of HIV diversity [Nowak and May, 1991], we present below an additional model that captures the critical nature of neoantigen *diversity limits* (see Results, [Aguadé-Gorgorió and Solé, 2020]). The model is grounded on a multiscale approach to the heterogeneity-surveillance tradeoff, where cancer cell death results from the immunogenic capacity of presented antigens and their intrinsic accumulation and distribution across the tumor, implying a background evolutionary layer. The model provides both a diversity biomarker able to predict immunotherapy outcome in melanoma patients, as well as an open possibility for reducing neoantigen heterogeneity prior to immune attack through a combination treatment approach [Aguadé-Gorgorió and Solé, 2020].

Deciphering the interplay between cancer and effector cells has been a major landmark of modern cancer research –the fact that T cells can target cancer cells makes our immune system one of the most advanced (if not the most) treatments available. However, evidence indicates that the cancer-immune interactions (Fig. 3) embody a much more complex network. One of the first observations of this was the fact that inflammation correlated with increased cancer incidence [Coussens and Werb, 2002]. In the scenario of a wounding injury, immune inflammation orchestrates a regenerative environment promoting cell proliferation and angiogenesis. This growth factor-rich ecosystem becomes a tasty site for malignant cell progression, making tumors act as *wounds that do not to heal* [Dvorak, 1986]. Innate immunity can be, therefore, a protumorigenic agent as a result of indirect facilitation dynamics [Bruno et al., 2003, Lin et al., 2007].

Cancer as a result of previously existing inflammation is only at the tip of the protumorigenic roles of the immune system. An example of further complexity can be found in the confusing role of macrophages in tumor progression. Macrophages, mostly characterized by their phagocytosis activity, can be roughly categorized into two subtypes, namely M1 and M2 [Mosser and Edwards, 2008]. M1s are mostly killer, proinflammatory macrophages, responsible for pathogen phagocytosis. M2s, in turn, are termed *repair* macrophages, and often act after M1s in complex anti-inflammatory and tissue-repairing cytokine networks inhibiting harmful immune attack [Mosser and Edwards, 2008]. Briefly put, M1s are anti-cancerous, while M2s favor a protumorigenic

environment. Current experimental observations indicate that, even in the absence of previous inflammation, cancer cells can, by secreting the necessary factors, attract and polarize macrophages into an M2 phenotype [Sica et al., 2008]. Since both cancer cells facilitate M2 reproduction, and M2s engineer a protumorigenic response, the tumor here behaves as an indirect mutualistic system, creating an ecosystem coengineering network that might be targeted by therapy (Fig. 7d, [Myers et al., 2020]).

All in all, the complex nature of cancer and immune populations underline a constant eco-evolutionary process, where cancer cells can evolve the capacities to escape immune predation, while clonal selection of matching TCRs constantly pursues mutated cancer antigens. On top of that, evolution of protumor immune populations indicates that the cancer-immune network is, itself, pervaded by multiple and diverse ecological interactions. To efficiently explore immunotherapy opportunities, oncology needs the theoretical background able to capture the ecological complexity at play and the resulting evolutionary directions that tumors are likely to take after treatment. Despite overwhelmingly complex, the dynamical and cognitive nature of the immune system as a *liquid brain* provides us with the hope of having found one of the few agents that could match cancer in terms of adaptive potential.

1.3.3 The ecology of metastatic spreading

About 90% of cancer-related deaths do not result from the growth of the *primary* tumor, but from the dispersal of cancer cells that eventually colonize distant host organs [Gupta and Massagué, 2006]. The so-called *metastatic cascade* involves a complex multi-step process where a single cancer cell looses cellular adhesion, increases motility and invasiveness, penetrates and survives in the circulation to eventually exit and colonize a distant site [Fidler, 2003] (Fig. 3 top, 10a). Several, if not all, of these processes are very unlikely in terms of the probability of both the genetic events and cellular survival involved. However, the ecological constraints of the primary site, where cancer outgrowth reduces resource availability, are likely to select for invasive phenotypes, similar to the evolution of foraging strategies in moving animals [Schoener, 1987], in cooperation with angiogenic phenotypes that create the necessary blood vessels for departure (Fig. 7e, [Weidner et al., 1991]).

Besides the genetic and molecular alterations characteristic of the metastatic cascade, the *seed and soil* hypothesis proposed by Stephen Paget in 1889 remains the benchmark of metastatic ecology thinking [Paget, 1889]. Paget's proposal was that the non-random dispersal of metastases[3] was comparable to the affinity of certain seeds (here, cancer cells) to the specific milieu of given soil types (here, organs or tissues). This ecological idea, deemed inconcievable for decades, started to be accepted in experimental terms by the 1970s [Weiss, 2000].

Once a single tumor has spread to distant organs with diverse ecological pressures, a main barrier to the treatment of metastatic disease is the underlying heterogeneity of each metastasis [Fidler, 2003]. In this scenario, surgery is not a feasible option, and systemic treatment usually fails due to the diversity of cancer cells at alternative sites, further

[3]Metastatic distribution across organs does not happen at random nor uniformly, but specific cancers seem to have preferred metastatic sites. A well-known example here is the common metastases of breast cancer towards the bones and lungs [Nguyen et al., 2009].

enhanced by the possible microenvironmental differences between tissues [Fidler, 2003] (Fig. 3). The context here is comparable to that of invasion ecology of islands [Lockwood et al., 2009, Lloyd et al., 2017], that establishes a framework to answer questions regarding how original biodiversity (ITH) is transfered to invaded ecosystems (organs). Here again, mathematical models have obtained key results that unravel a consecutive seeding scenario, where not a single, but many primary tumor cells eventually reach and transfer genetic information to distant metastases [Heyde et al., 2019]. Further evidence also indicates that seeding between the primary tumor and distant metastatic sites is actually bidirectional [Comen et al., 2011] (Fig. 10a). In this context, the *seed and soil* hypothesis expands to a much more complex scenario where tumors and their metastases coevolve by exchanging genetic information in a recombination-like fashion.

Another aspect of metastatic spreading governed by ecological interactions is *concomitant resistance* [Prehn, 1993]. Concomitant resistance (CR) is the paradoxical observation, first discussed by Paul Ehrlich [Ehrlich, 1909], that an inoculated secondary tumor will fail at colonizing the host if a large tumor is already present within the same organism [Chiarella et al., 2012]. This apparently distant–competition process governs the more important phenomenom of sudden metastatic expansion after primary tumor surgery [Coffey et al., 2003]. A relevant example here is that of advanced breast cancers, where late but successful surgery of the primary tumor is often followed by the sudden growth of aggressive metastatic disease [Retsky et al., 2010], indicating that the presence of the primary tumor was somehow maintaining secondary tumors under dormancy. Despite mathematical efforts, explicit derivations of the type of ecological interactions that could result in CR remain elusive, and direct or indirect competition [Benzekry et al., 2017], concomitant immunity [Gorelik et al., 1981] or angiogenic inhibition [O'Reilly et al., 1999] do not totally explain systemic CR dynamics. Can we envisage a metapopulation model (Fig. 10b) that explains which molecular and ecological cues induce concomitant resistance?

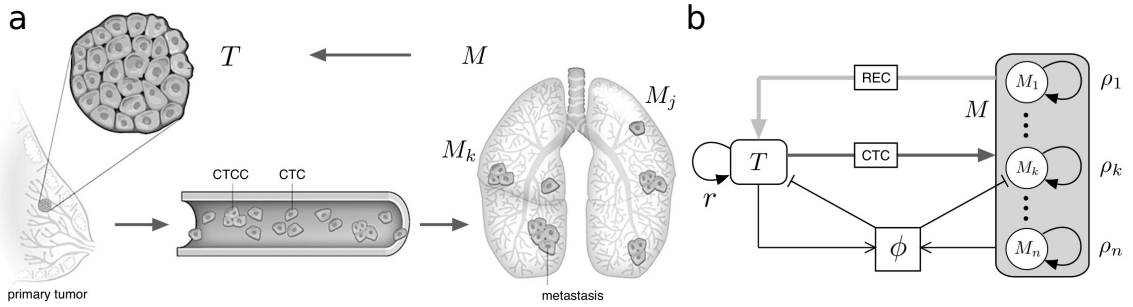

Figure 10: **Ecological cues in metastatic spreading.** Metastatic spreading involves a complex process of cellular transformation built upon several steps, from the EMT transition to an invasive mesenchymal phenotype, the navigation across populated blood vessels to the colonization of a distant organ **(a)**. Furthermore, evidence indicates the presence of non-trivial dynamics, such as metastases-tumor reseeding (implying the sharing of heterogeneity) or distant tumor inhibition (concomitant resistance). Ecological multi-species models **(b)** are needed in order to capture the possible critical thresholds limiting metastatic invasion.

1.4 A caricature of tissue development

Back in 1829, before most of the biological and molecular knowledge of cancer was established, one of the first observations of tumors under the microscope was their resemblance to embryonic tissue [Recamier, 1829]. Decades later, these observations developed into the theory of *embryonal rest*, by which adult tissues contain embryonic remnants that could be activated from dormant states into cancer, in part due to their inherent capacity for proliferation [Cohnheim, 1914]. By 1941, studies on teratomas (germ cell tumors made up of several well-differentiated tissues, such as hair, muscle or teeth) led to the evidence that tumors could in fact contain both differentiated and undifferentiated cells [Jackson and Brues, 1941].

During the same period, evolution of our knowledge on tissue development and regeneration led to the discovery of *Stem Cells*, undifferentiated cells with unlimited capacity to either proliferate or differentiate into various cell types [van der Kooy and Weiss, 2000]. These were first defined as such by Theodor Boveri in the late 19th century [Ramalho-Santos and Willenbring, 2007], at the same years he developed a theory for the chromosomal origin of cancer (Section 1.2.1). By the 1960s, discovery of blood-forming (hematopoietic) stem cells (HSC) moved to the front page the relevance and key properties of stem cells [Becker et al., 1963]. Furthermore, the existence of stem cells uncovered the underlying *cellular lineages*, fine-tuned hierarchies of cells of diverse differentiation degree that maintain stable tissue architecture [Visvader and Clevers, 2016].

Could tumors arise from these embryonic cells? Would heterogeneity then be a result of an aberrant hierarchy, instead of a process of the somatic accumulation of mutations? The discovery that some cancers can arise from specific cells with tumorigenic potential, as if tumors followed from a process of aberrant development, led to the establishment of the *Cancer Stem Cell* (CSC) model [Reya et al., 2001]. If tumors are indeed *caricatures of normal tissue development* [Pierce and Speers, 1988], and not the simple Big Bang expansion of a given cell of origin [Sottoriva et al., 2015], our understanding of oncogenesis, cellular heterogeneity and consequent therapeutic resistance needs to be revisited.

1.4.1 Cell fate and the Waddington landscape

The human body is made of about 200 different cell types [Alberts, 2008], organized in complex multicellular patterns to maintain body function. All these different cells, however, share the same DNA sequence and originate from a single embryonic stem cell [Carroll et al., 2013], meaning that a single genome can encode many different phenotypes. The underlying processes governing how cells develop into alternative phenotypes with specific functions is the key area of study of developmental biology. During embryogenesis, undifferentiated *stem cells* replicate and specialize into different cell types by receiving molecular signals from their surroundings [Carroll et al., 2013]. A similar process can be observed later on in adult organs, where tissue stem cells maintain and regenerate specialized cellular populations when needed [Wagers and Weissman, 2004].

A relevant example here is that of hematopoiesis, the formation of all blood cells from HSCs and subsequent myeloid and lymphoid progenitors (see [Morrison et al., 1995] for an extensive description).

The findings of developmental biology have been of utmost importance to our understanding of classical Darwinian evolution, as genes appear no longer to be the only agents of phenotypic expression nor heritability [Goldberg et al., 2007]. This was first proposed by Conrad Waddington in 1942 [Waddington, 1942], who coined the term *epigenetics* to include all non-genetic mechanisms of heritable information [Goldberg et al., 2007]. To understand how single DNA sequences develop into multiple (and stable) cellular types, he proposed a visual landscape, now known as the *Waddington landscape* [Waddington, 1957], where cell fate is visualized as a downhill descent from less differentiated states into canals representing possible cell fates (Fig. 11a).

At the top of the landscape, an initial totipotent stem cell[4], precludes tissue development, with differentiation (here downhill pathways) governed by external signaling cascades affecting gene expression patterns [Huang, 2012]. Once the embryo becomes an adult individual, hierarchies of cells at different landscape heights choosing between self-replication or differentiation maintains a fine-tuned equilibrium between phenotypic populations that ensures correct organ function [Lehrer et al., 1998, Derényi and Szöllősi, 2017] (Fig. 11b).

What are the *epigenetic* programs determining the topology of such landscape? A relevant framework here is provided by Gene Regulatory Networks (GRNs) [Davidson and Levin, 2005]. As discussed along Section 1.2.1, genes interact between one another, so that the product or expression level of a gene is a function of the expression of N other genes, and so on, creating a network of epistatic interactions [Wolf et al., 2000]. In an enlightening work, Stuart Kauffamn proposed GRNs could be modeled as Boolean networks [Kauffman, 1969], where each gene state is a logic function of other N states. Interestingly, such network has a set of S stable states, or attractors, towards which the system evolves [Kauffman, 1969], that map the possible cell types that arise from the original DNA sequence depending on gene expression patterns. Surprisingly, even very large random networks entail a much smaller and limited number of attractor states, consistent with the divergence between the number of DNA components and resulting cell types [Solé and Goodwin, 2000].

As a corollary and more realistic version of random Kauffman networks, network models where node states are continuous and stochastic provide a first mathematical tool that maps genome expression to the topology of the Waddington landscape [Huang, 2012]. Chromatin, previously studied in *1.2*, is the key epigenetic element regulating gene expression in the GRN [Flavahan et al., 2017]. Alterations in gene expression or chromatin configuration, therefore, could induce changes in the topology of walls separating attractor states, thus allowing the developmental process to access previously forbidden phenotypes [Huang, 2012]. The Waddington landscape is, therefore, a source for cellular heterogeneity at two characteristic levels. On the one hand, the existence of developmental hierarchies provides a stable mechanism for tissue plasticity. On the other, epigenetic changes alter the topology of accessible phenotypes, eliciting the creation of novel developmental paths. Can cancer take advantage of this two adaptive opportunities? And, if

[4]Not all stem cells have the same capacity –or *potency*– to differentiate into alternative cell types. Totipotent stem cells, at the top of the hierarchy, are the initial cells in the embryo such as zygotes [Gardner, 2002].

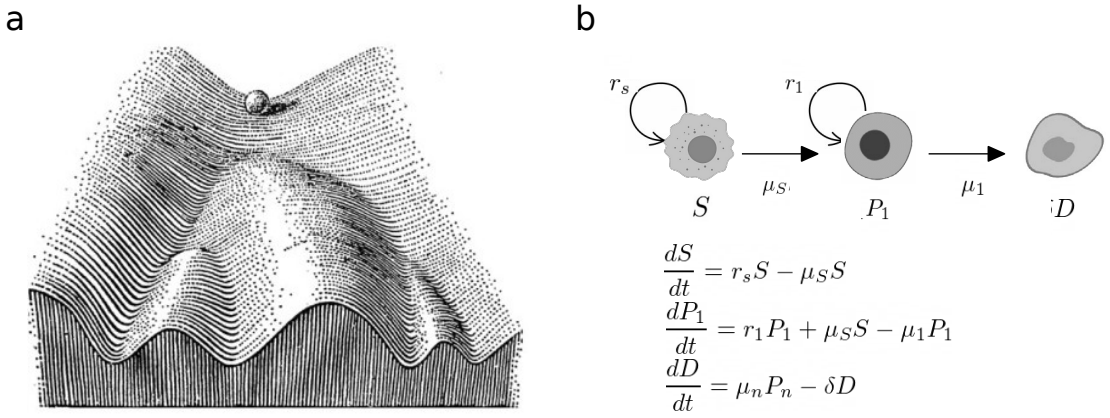

Figure 11: **Developmental complexity and tissue hierarchies.** Cell development as captured by the Waddington landscape **(a)**, where cell fate is visualized as a downhill roll from less differentiated states (high regions) towards one or another *attractor state*, representing alternative cellular phenotypes. In adult organs, the Waddington landscape sustains cellular hierarchies maintaining tissue homeostasis **(b)**. In some cancers, aberrant cells often take advantage of phenotypic heterogeneity by exploring the phenotypic heterogeneity of these stable hierarchies, with *Cancer Stem Cells* at the top of the process being the agent responsible for tumor proliferation and resilience.

so, what are the necessary tools to understand tumor progression under both scenarios?

1.4.2 Cancer Stem Cells

The proliferative and adaptive potential of healthy stem cells should already remind us of cancer cells themselves. The possibility that cancer was a disease related to stem cell transformation gained attention in the 1990s, when Acute Myeloid Leukemia (AML) was discovered to follow a hierarchical structure as that of blood lineages, thus originating from a single HSC compartment with unique tumorigenic capacity [Bonnet and Dick, 1997] that maintained a heterogeneous population of more differentiated leukemia cells. Evidence for AML, breast cancer and many other tumor types indicated that only a small set of particular cell types within each tumor where potential candidates for tumorigenesis after transplantation, making these the stem-like architects of oncogenic growth [Al-Hajj et al., 2003]. The existence of so-called *cancer stem cells*, consistent with observations of the embryonic appearance of tumors, was followed by the notion that loss of multicellular traits could be accompanied by a backwards differentiation process [Sole et al., 2014], or else that tumors could online arise from mutations in healthy stem cells [Sell, 2004]. In any case, this precluded the possibility that cancers progress imitating healthy tissue development [Pierce and Speers, 1988].

Discovery of hierarchies with different tumorigenic potential across tumor types (see [Nassar and Blanpain, 2016] for a comprehensive review) has led to the establishment of the so-called *cancer stem cell model* [Batlle and Clevers, 2017], that can be summarized into four premises. First, extensive ITH in tumors follows from its hierarchical organization (Fig. 11b), eventually resembling that of the tissue of origin. Second, these hierarchies contain rare self-renewing CSCs that maintain a larger tumor population with only

transient capacity for proliferation. Third, such tumor hierarchical structure is considered stable, following from the overall lack of tumorigenic potential of non-CSCs. And fourth, common relapse after treatment can be explained by CSC tolerance to standard chemotherapy and radiation therapies, mostly as a result from their typically quiescent state [Batlle and Clevers, 2017]. The CSC model is able to accurately explain clinical observations, such as tumor dormancy, the epithelial-mesenchymal transition precluding metastases or overall therapy resistance [Batlle and Clevers, 2017]. As so, it provides a conceptual framework to understand cancer progression without the need to explicitly account for Darwinian dynamics.

At first view, the possibility that cancer is organized as a hardwired hierarchy apparently contradicts the clonal evolution model (Section 1.2.3, [Greaves and Maley, 2012]). Here, cellular properties are no longer randomly distributed across clones, as expected from the Luria-Delbrück approach to neutral mutation accumulation [Luria and Delbrück, 1943, Williams et al., 2016] (Figs. 5,6), meaning that the adaptive capacity of cancers should be understood as a remnant of stem cell plasticity instead of a result of the Darwinian selection of ecologically beneficial phenotypes [Shackleton et al., 2009]. However, current state of the art acknowledges that not all cancers follow a well-defined CSC model, and it is necessary to mention that CSCs themselves are sometimes ill-defined and context-dependent [Shackleton et al., 2009]. Furthermore, the clonal and CSC models of tumor progression have been found to be non-exclusive, and well established tissue hierarchies (Fig. 11b) can in turn accumulate mutations and undergo Darwinian selection at each level of the hierarchy [Fulawka et al., 2014].

For those tumors that do progress under a CSC model, treatment design must take into account several potential barriers. CSCs are characterized by the capacity to adopt quiescent states with slow proliferative cycling [Li and Bhatia, 2011]. Given that chemotherapy and radiation –the most widely used cancer therapeutics– both target DNA damage in rapidly dividing cells, CSC quiescence ensures that the stem cell compartment persists after therapy, steadily repopulating the tumor after drug release [Batlle and Clevers, 2017]. As a proof of concept, mathematical models of multi-layered cellular compartments (Fig. 11b) predict that therapy targeting a very large fraction of the tumor is likely to fail if a CSC population persists, and only therapy targeting this compartment can effectively achieve tumor remission [Dingli and Michor, 2006]. Treatment design, therefore, should focus on therapies directly targeting CSCs.

A prominent example here is that of differentiation therapy (DTH). As a promising alternative to convential cytotoxic approaches, DTH involves the use of molecular agents able to induce cellular maturation, thus targeting loss of differentiation and moving cells away from the proliferative compartment and downhill into Waddington canals (Fig. 11, [Sell, 2004, de Thé et al., 2018]). DTH attention was fueled by its use in Acute Promyelocytic Leukemia (APL), a subtype of AML characterized by the proliferation of immature granulocytes. APL is a cancer driven by a particularly simple genomic event, a chromosomal translocation resulting in the fusion of the retinoic acid receptor-alpha gene on chromosome 17 (RARA) with the promyelocytic leukemia gene (PML) on chromosome 15 [de Thé et al., 2018]. All-trans retinoic acid was discovered to induce differentiation of the immature cells in APL without directly killing them. In turn, a combination of standard chemotherapy and all-trans retinoic acid increased the likelihood of disease remission up to 90% [Meng-Er et al., 1988], an astonishing improvement for a disease

previously considered a hyperacute fatal illness, with a median survival time of less than a week in the 1950s [Coombs et al., 2015].

In many other cancers, however, the search for a DTH approach has not granted successful results [de Thé et al., 2018]. The more complex genome alterations of solid tumors, that also generate non-trivial differentiation hierarchies [Nassar and Blanpain, 2016], have hampered the extension of DTH to solid cancers in particular. pointing towards several potential spatial limitations to effective differentiation. In particular, the fact that CSCs could live in alternative niches with specific differentiation signaling might confer an additional layer of resistance to DTH (Fig. 8c). In the present thesis we present a mathematical approach to the role of space in DTH design, by considering how the differentiated, unproliferative compartment alters cancer growth (Eq. 7) through occupying invadable habitat [Solé and Aguadé-Gorgorió, 2021]. This modifies the original Levins model to account for a minimal hierarchy

$$\frac{dc}{dt} = b\left(1 - \frac{c+D}{K}\right)c - \delta c - \gamma f(c,D) \tag{10}$$

$$\frac{dD}{dt} = \gamma f(c,D) - \rho D \tag{11}$$

where $\gamma f(c,D)$ captures the signaling cascade by which the amount of neighbouring progenitor (c) and differentiated (D) cells affect the basal differentiation rate of progenitor cells γ [Solé and Aguadé-Gorgorió, 2021]. The key element here is observing how D produces an indirect spatial competition by introducing an habitat-loss term in the cancer birth rate b. Interestingly, this minimal system is able to capture the failure of differentiating agents alone, consistent with observations on APL. For tumors with more complex architectures, the same modeling approach is able to capture how does the minimal system change when a CSC population –living in an independent compartment– survives habitat loss, pointing towards the *CSC niche* [Plaks et al., 2015] as a buffer against maturation-inducing compounds (See Results, [Solé and Aguadé-Gorgorió, 2021].

1.4.3 Phenotypic plasticity in cancer

Our understanding of cellular differentiation programs changed around 2006, with the discovery that only a few transcription factors, the so-called *Yamanaka Factors*, could induce dedifferentiation of otherwise mature cells [Takahashi and Yamanaka, 2006]. By 2011, evidence of spontaneous dedifferentiation of mammary basal-like human cells [Chaffer et al., 2011] ignited the possibility that cellular hierarchies are not as hardwired as previously expected [Poulsom et al., 2002]. Nowadays, a large body of evidence confirms that committed cells are capable of moving up and down differentiation hierarchies under a wide array of molecular or environmental signals [Batlle and Clevers, 2017]. On top of that, transcriptional noise might put a further layer of stochasticity in the (already complex) genotype-to-phenotype map [Marusyk et al., 2012], further deforming the structure of tissue hierarchies.

In cancer, *phenotypic plasticity* –the ability of cells to move around the Waddington landscape (Fig. 11a)– opens a totally new scenario in cancer adaptation [Flavahan et al., 2017]. A prominent example here was the discovery that breast cancer cell lines could consistently display well-defined stem-, basal- and luminal-like phenotypes on top of a

much more noisy heterogeneous genetic background [Gupta et al., 2011]. Interestingly, a tumor grown out of any single cell in the hierarchy recapitulated the original distribution, indicating that heterogeneity between cellular states could not result from random accumulation of mutations. With a simple computational model of a Markov process, Gupta and colleagues where able to show that these cellular states coexisted under a phenotypic switching (PHS) strategy [Gupta et al., 2011].

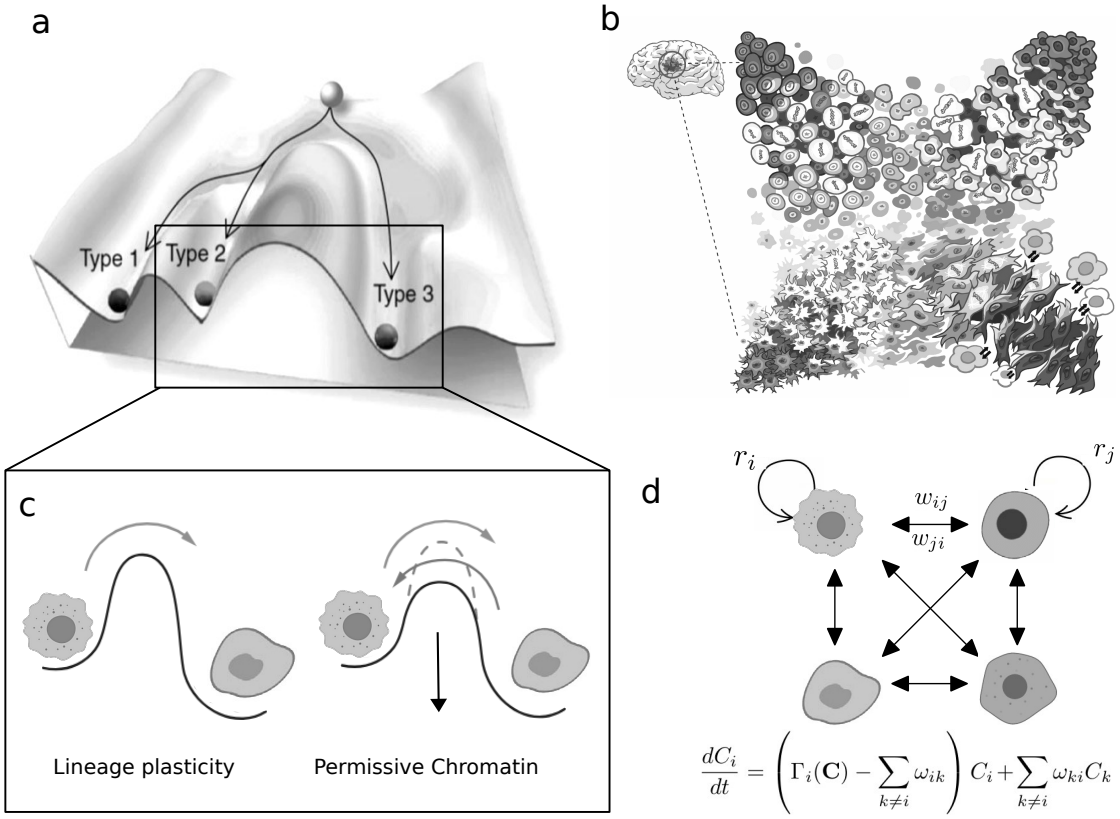

Figure 12: **Phenotypic plasticity in cancer.** Genetic and epigenetic cues build up the underlying topology of the Waddington landscape **(a)** (redrawn from [Ollé-Vila et al., 2016]). Evidence for PHS strategies in cancer indicates that tumors could take advantage from modifications on this landscape (in **(b)**, a schematic representation of four well-defined PHS phenotypes in Glioblastoma, redrawn from [Neftel et al., 2019]), namely related to changes in chromatin configuration facilitating the accessibility of alternative phenotypic states [Flavahan et al., 2017] **(c)**. Mathematical models of PHS architectures allow for the dynamical description of possible thresholds to treatment efficiency **(d)** [Aguadé-Gorgorió et al., 2020].

Phenotypic plasticity between cellular phenotyes has been shown to govern cancer adaptation in a wide array of settings [Flavahan et al., 2017], such as in epithelial–mesenchymal heterogeneity[5], the appearance of transient drug-tolerant states during therapy [Sharma et al., 2010] or lineage plasticity[6] as response to therapeutic intervention

[5]The epithelial-mesenchymal transition (EMT) encodes the phenotypic change between epithelial cells that loose adhesion properties and embark into a mesenchymal stem cell phenotype with migratory potential [Kalluri et al., 2009]. The EMT is, therefore, a key event in the formation of metastatic disease.

[6]Lineage plasticity (LP) refers to the recent observation that, upon targeted treatment, cell lines in adeno-

[Quintanal-Villalonga et al., 2020]. More complex PHS architectures, with more than two phenotypes involved, have been shown to boost cancer heterogeneity and adaptation. A recent illuminating discovery is that of Glioblastoma (GBM), an aggressive form of brain cancer, that has been shown to develop PHS between four well-defined genetic meta-modules, resembling originary cell lines in brain development (Fig. 12b, [Neftel et al., 2019]). This finding has major implications for GBM treatment design.

The Waddington landscape is again a useful tool to visualize the epigenetic laws behing PHS [Waddington, 1957] (Fig. 12a). Configuration of the valleys and walls of the developmental landscape follows from the architectural topology of a GRN [Huang, 2012]. On the one hand, gene expression patterns define the possible attractor states [Marusyk et al., 2012]. Extensive genome alterations in malignant cells can therefore induce the creation of novel valleys of variable proliferative potential, the so-called *cancer attractors* [Huang et al., 2009]. On the other hand, the effect of chromatin configuration in DNA folding alters the height of walls separating these valleys [Marusyk et al., 2012] (Fig. 12c). Restrictive chromatin states ensure that wells separating the canals are high enough so that transcriptional noise does not induce excessive cellular transformation [Flavahan et al., 2017] (Fig. 12c left). In cancer, epigenetic alterations of chromatin configuration often induce *permissive* chromatin states, that translate into the Waddington landscape as a decrease of the landscape ruggedness (Fig. 12c right) allowing for more probable stochastic transitions between cellular phenotypes [Flavahan et al., 2017].

Mathematical modeling coherent with experimental observations has been able to demonstrate that phenotypic switching dynamics can be captured by simple stochastic models inspired in *bacterial persistence*, the ability of bacterial communities to maintain a small drug-tolerant population in place [Balaban et al., 2004]. The crucible of the problem here is solved by introducing positive switching rates w_{ij} between phenotypes [Aguadé-Gorgorió et al., 2020] (Fig. 12d). The equation for each subclone under PHS here writes

$$\frac{dc_i}{dt} = \left(r_i(\mathbf{c}) - \sum_{k \neq i} \omega_{ik} \right) c_i + \sum_{k \neq i} \omega_{ki} c_k \tag{12}$$

so that population growth now includes those cells that switch into the c_i phenotype minus those that leave into alternative ones. Interestingly, these minimal set of replicator equations completely transforms the landscape of system attractors [Aguadé-Gorgorió et al., 2020], by allowing softer conditions for multicellular coexistence than those characteristic of replicators in the clonal model (Eq. (7)). In this scenario, one particular difficulty of therapy targeting tumors with PHS strategies is the resilience of heterogeneity, since targeted phenotypes can be held in place by the rest of the system [Aguadé-Gorgorió et al., 2020] (see Results). Can we design alternative therapies that take into account the dynamical resilience of PHS?

carcinomas loose dependency upon initial oncogenic drivers and transform into aggressive neuroendocrine derivatives. Lung and prostate cancers are so far the most studied adenocarcinomas showing LP [Quintanal-Villalonga et al., 2020].

2 Objectives

The complexity of cancer – and the search to find successful therapeutic approaches – spans across the fields of ecology, evolution and development. For each of these, a large body of different questions arise, many of which are possibly answered by using theoretical biology and mathematical models. The initial objective of the present thesis, therefore, was to identify relevant questions within oncology that should be approached through the lenses of complexity theory, and a large part of the research time has been involved in properly defining a concise question for each specific aspect in place. Some of the questions that have been targeted, with different rate of success, during the time of my PhD have been the following:

Evolution

- If there are viability limits to genetic instability [Solé and Deisboeck, 2004], how can cancer cells survive them? In this context, what is the dynamical nature underlying mutation rate evolution? Do cancer cells live at the critical mutation rate, as viruses do [Solé and Elena, 2018]? If so, can lethal mutagenesis be applied as a therapeutic approach to unstable tumors?
- How does chromosomal instability (CIN) modulate the adaptive landscape of cancer cells? Why do many cancers organize around an average ploidy of $\sim$3.3? Is this an optimal evolutionary value, or a physiological constraint? Are there error thresholds to chromosomal instability?
- What is the relation of Whole Genome Doubling (WGD) with the previous question? Why is WGD such a common event in cancer, despite being apparently detrimental to fast cellular proliferation?
- Why do CIN and micro-satellite instability pathways appear as mutually-exclusive evolutionary mechanisms in cancer [Guinney et al., 2015]?

Ecology

- What is the role of genetic instability in the cancer-immune (predator-prey) interaction? If T cells recognize cancer neoantigens, and these result from accumulated mutations, is there a trade-off balancing cancer evolution and immune recognition? Can mutagenic therapies be combined to checkpoint blockade inhibitors to enhance immune surveillance? (an **Eco-Evo** question)

- How does heterogeneity in neoantigen landscapes affect the previous question? Can a mathematical model capture why (and how) increased neoantigen diversity correlates with immunotherapy failure? Are there specific limits to neoantigen heterogeneity, and can therapy take advantage of them? (another **Eco-Evo** question)

- Is the Warburg effect (the use of inefficient glycolytic metabolism even in the presence of oxygen) an example of a critical transition with hysteresis?

- Can mathematical modeling help us understand the ecosystem coengineering nature of cancer and type-2 macrophages [Myers et al., 2020]? Provided that M1-M2 equilibria can be restored by therapy, can we predict the conditions under which the immune system will transit from a pro-tumor to an anti-tumor response?

- What is the ecological nature of Concomitant Resistance? Why do tumors apparently secrete molecular compounds that do not allow secondary tumors to grow? Can a mathematical model capture this apparently indirect competition process? If surgery is successful for small tumors, but awakens large metastatic burdens in large tumor excisions, can we predict the presence of a threshold tumor size with potential therapeutic value?

Development

- What is the dynamical nature of Differentiation Therapy (DTH) in solid tumors, and why does it usually fail? What role does the CSC niche play here? Can we use spatial ecological models to understand why DTH affects differently each cellular compartment in a hierarchy? (an **Eco-Devo** question)

- If tumors adapt through stochastic phenotypic switching (PHS) instead of somatic mutation accumulation, what is the nature of therapeutic resistance? Can a simple analytical framework inform about the adaptive potential of PHS versus the clonal evolution model? Would it provide clues for the treatment of multi-phenotype structures such as those of Glioblastomas? (another **Eco-Devo** question)

- Is it possible to develop a mathematical framework that unites Waddington's landscape and Kauffman's cancer attractors to capture the dynamical limits of epigenetic plasticity? Can it account for why (and how) PHS strategies evolve in the first place? Can we understand if PHS architectures have evolved *de novo* or from an aberrant –but already existing– tissue hierarchy?

A range of mathematical tools have been considered to approach these questions, most often those of population ecology, such as ordinary differential equations to study population dynamics, diversity models or habitat loss and fragmentation in ecology [Bascompte and Solé, 1996]. We have also used adaptive dynamics as a framework to study stochastic trait evolution [Diekmann, 2002]. When needed, we have taken advantage of computational simulations to reproduce *in silico* experiments.

The main overall goal of the research process has been to obtain, for each question, a minimal model both able to provide a basic understanding of the system and a concise result with specific implications on cancer treatment. Experimental and clinical data from other research teams and publications has been used, when possible, to assess the validity of the models' results.

Interestingly enough, few of the questions for each specific field have yielded answers constrained to that field, meaning that ecological questions have been target by using evolutionary or developmental models, and viceversa. By the end of the PhD thesis, the interaction between ecology, evolution and development in cancer has yielded another open question:

- Is it possible to build a conceptual framework to understand how ecology, evolution and development are intertwined in cancer? Can we map how different cancers explore complexity along each of these axes? Does this provide a cartography of alternative tumorigenic pathways?

3 Results

The present chapter contains the following research articles, published within the scope of the thesis:

- Aguadé-Gorgorió, G., & Solé, R. (2018). Adaptive dynamics of unstable cancer populations: The canonical equation. *Evolutionary applications*, 11(8), 1283-1292.
- Aguadé-Gorgorió, G., & Solé, R. (2019). Genetic instability as a driver for immune surveillance. *Journal for immunotherapy of cancer*, 7(1), 1-13.
- Aguadé-Gorgorió, G., & Solé, R. (2020). Tumour neoantigen heterogeneity thresholds provide a time window for combination immunotherapy. *Journal of the Royal Society Interface*, 17(171), 20200736.
- Solé, R., & Aguadé-Gorgorió, G. (2021). The ecology of cancer differentiation therapy. *Journal of Theoretical Biology*, 511, 110552.

Another article presented below was under revision for publication at the time of the thesis deposit (May 2021):

- Aguadé-Gorgorió, G., Kauffman, S., & Solé, R. (2021). Transition therapy: tackling the ecology of tumour phenotypic plasticity.

  (Preprints: https://www.preprints.org/manuscript/202007.0547/v1)

Received: 12 November 2017 | Accepted: 15 February 2018

DOI: 10.1111/eva.12625

ORIGINAL ARTICLE

WILEY Evolutionary Applications Open Access

Adaptive dynamics of unstable cancer populations: The canonical equation

Guim Aguadé-Gorgorió[1,2] | Ricard Solé[1,2,3]

[1]ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Barcelona, Spain

[2]Institut de Biologia Evolutiva (CSIC-UPF), Barcelona, Spain

[3]Santa Fe Institute, Santa Fe, NM, USA

Correspondence
Ricard Solé, ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Barcelona, Spain.
Email: ricard.sole@upf.edu

Funding information
Fundacion Botin-Banco de Santander; Botín Foundation by Banco Santander through its Santander Universities Global Division; MINECO, Grant/Award Number: FIS2015-67616; Universities and Research Secretariat of the Ministry of Business and Knowledge of the Generalitat de Catalunya; European Social Fund; Santa Fe Institute

Abstract

In most instances of tumour development, genetic instability plays a role in allowing cancer cell populations to respond to selection barriers, such as physical constraints or immune responses, and rapidly adapt to an always changing environment. Modelling instability is a nontrivial task, since by definition evolving instability leads to changes in the underlying landscape. In this article, we explore mathematically a simple version of unstable tumour progression using the formalism of adaptive dynamics (AD) where selection and mutation are explicitly coupled. Using a set of basic fitness landscapes, the so-called canonical equation for the evolution of genetic instability on a minimal scenario associated with a population of unstable cells is derived. We obtain explicit expressions for the evolution of mutation probabilities, and the implications of the model on further experimental studies and potential mutagenic therapies are discussed.

KEYWORDS

cancer adaptation, critical points, genome instability, Moran process, unstable dynamics

1 | INTRODUCTION

Cancer can be understood as the failure of those regulatory mechanisms that guarantee the maintenance of tissue and organ homoeostasis. Cooperative interactions along with extensive feedback signalling loops and replication checkpoints provide multiple paths to avoid the emergence of undesirable mutations or chromosomal abnormalities that can allow rogue cells to start proliferative growth. In dynamical terms, what has to be avoided within multicellular organisms is any kind of individual cell Darwinian evolution (Gatenby & Brown, 2017; Greaves & Maley, 2012; Nowell, 1976).

It is generally acknowledged that genetic instability plays a key role in tumour progression and carcinogenesis (Hanahan & Weinberg, 2011). Unstable genomes result from the failure of molecular mechanisms responsible for the maintenance of genome integrity (Negrini, Gorgoulis, & Halazonetis, 2010). That cancer cells are unstable is fairly well illustrated by the observation of their karyotypes: in sharp contrast with healthy cells, cancer chromosomal arrangements reveal a wide degree of aneuploidy (Lengauer, Kinzler, & Vogelstein, 1998). Such high levels of mutational load deploy the potential to overcome selection barriers, as well as involve a rather uncommon process from multicellularity to reduced cellular complexity (Solé et al., 2014), giving place to a highly adaptive and heterogeneous population. Genetic instability acts as a driver as well as the search engine for disease progression. An important (and not always appreciated) consequence of instability is that, once unleashed, it can easily grow as the lack of proper repair can damage other components of the check-and-repair cellular network.

Despite increasing knowledge of the molecular basis of unstable tumorigenesis, there is still the need for understanding the role of instability on cancer evolution, namely discerning if it is a cause or a consequence of carcinogenesis, how does it evolve along tumour development, and what are the treatment strategies that arise from answering such questions. Many mathematical models

have provided interesting points into this topic, with the introduction of relevant ideas such as the mutator phenotype (Loeb, 2001) and several multi-step models of mutation acquisition (see e.g., Komarova et al., 2002; Nowak et al., 2002) that have investigated the possible scenarios of correlation between instability and cancer progression.

The fact that genetic instability itself changes over cancer evolution makes it difficult to properly model its behaviour. Particular efforts, such as the computational models of Komarova, Sadovsky, & Wan, 2008 and Datta et al., 2013; have given interesting insight into understanding how a changing instability level affects, by means of modifying the probability of mutations, all kinds of replication and control mechanisms within the vast pathways towards cancer malignancy. Within this picture, instability cannot be taken as a parameter, but rather as an evolving phenotypic trait affected by the selective pressures of the tumour microenvironment. In this scope, the recent work by Asatryan and Komarova represents a further step for its proposal of an analytical approach where both instability and heterogeneity of cancer populations can be traced along time (Asatryan & Komarova, 2016). As a complementary point of view, we consider the need to include stochasticity in the process of acquiring either advantageous or deleterious mutations, together with considering instability as a trait evolving through changes within each single cell, compared to the idea of measuring it by following the competition dynamics between subpopulations with fixed mutation probabilities.

Here, we propose that the mean evolutionary paths of such stochastic process followed by unstable populations are describable by means of the framework of adaptive dynamics (AD) (Champagnat, Ferriere, & Ben Arous, 2001; Dieckmann & Law, 1996), which has been used in the study of cancer when focusing on niche construction (Gerlee & Anderson, 2015). AD models provide a powerful alternative to previous formal approaches by explicitly including replication, mutation and selection in a consistent way, allowing an exploration of the evolutionary dynamics of adaptive traits, while at the same time keeping a minimal, treatable model able to produce explicit expressions for trait evolution depending on a few parameters.

A central object in the AD framework is the so-called canonical equation. For a given quantitative phenotypic trait s, this equation describes the evolutionary trajectory for the mean trait value as

$$\frac{d\langle s\rangle}{dt} = \frac{1}{2}\mu(\langle s\rangle)\sigma^2(\langle s\rangle)n(\langle s\rangle)\left(\frac{\partial r(s,s')}{\partial s'}\right)_{s'=\langle s\rangle} \quad (1)$$

where $\mu(\langle s\rangle)$ is the probability under which mutant individuals are generated, σ^2 is the variance of the mutant distribution s' derived from an individual with trait s, n the stationary population size and the last term in the right-hand side stands for the fitness gradient associated with the specific landscape at work. The standard formulation involves some strong assumptions on the mutation-selection process, and we will therefore review the mathematical process in order to understand up to which point the framework is suitable for our problem.

In the AD models, and in the work presented here, evolution takes place within a constant population context, where mutants appear and invade in a stepwise process, leading to a formalism for evaluating the trajectories of evolving trait values. This picture of cancer dynamics stems from classic work on ecological competition (Gatenby, 1995) where tumour cells act as invaders that cause the disruption of the local (tissue) ecology. These simplified models reveal how a proper formulation of competition can yield useful predictions (Gatenby, 1991). In this context, although the constant population falls short to describe the behaviour of some tumour growth processes, it is a much needed first approximation. Moreover, it can also be appropriate when dealing with some in vitro experiments involving long-term evolution of unstable cancer cell populations. We will go back to this at the end of the paper.

Understanding how instability becomes a driver of evolvability can give further insight about its role as a cancer hallmark, and might as well produce relevant steps towards contemplating genetic instability as potential target for treatment. Is it possible to formulate a canonical equation describing the time evolution of instability? The answer is affirmative and here we show how it can be obtained.

2 | POPULATION DYNAMICS

With the aim of obtaining a clear understanding of the questions proposed above, we look for a minimal model to implement the unstable evolutionary dynamics. Our goal is to consider the process of cancer progression, which involves a heterogeneous population of cells (Figure 1a). In this population, cells are only characterized by their particular replication rate r_i and mutation probability μ_i. However, and in the eyes of the AD approach, this preliminary model uses a constant population approximation where mutation probabilities are small enough so that the dynamics remain in equilibrium in-between invasions. This approach, whose limitations will be later thoroughly discussed, is best described by means of a so-called Moran process (Moran, 1958).

A particularity of the Moran process—here coarse-grained into a continuous process, keeping in mind the long-term evolution of tumour progression—is that cells of type c_i give birth by means of occupying other, randomly chosen cell sites at rate r_i, so that the birth–death process is coupled into a single event (Figure 1c) that will eventually lead to selection towards cells with higher r_i, thus producing a minimal environment where selection can take place. Furthermore, mutation is introduced by considering that cells can give birth to mutant offspring at probability μ_i.

Mutations, however, do not occur as in quasispecies or replicator-mutator models, where genomes mutate from one to another. In our model, a newly born mutant cell will have a modified mutation probability $\mu' = \mu_i + \Delta\mu$, where $\Delta\mu$ is taken from a continuous distribution

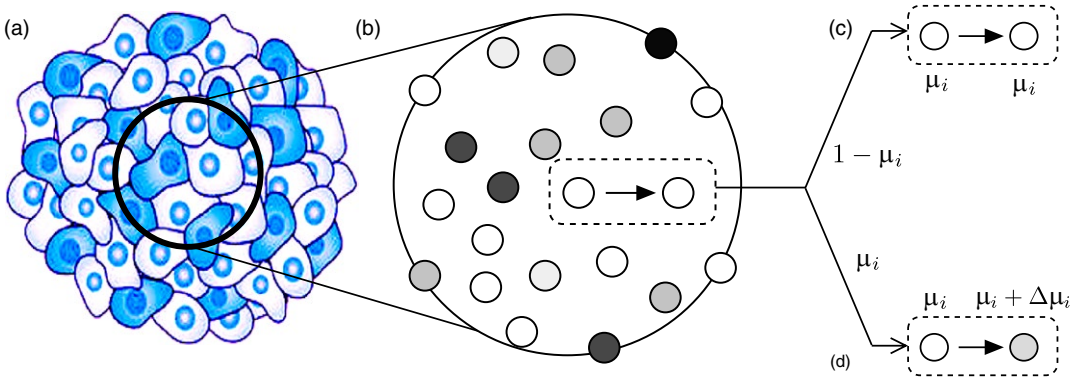

FIGURE 1 The Moran process rules associated with the model of a population of unstable cells competing for resources. We consider an idealized model of a heterogeneous cancer cell population (a) described by a well-mixed (mean field) model (b). Here cells occupy a given domain that is not explicit and each cell has a distinct phenotype described in particular by its intrinsic instability μ_i. In the Moran process, when a cell replicates it occupies another cell's niche and produces an identical daughter (c) or a slightly different one due to a mutation event proportional to μ_i, which can lead to an increase $\Delta\mu_i$ of the instability levels (d)

that we discuss later on. With this, we emphasize the wide levels of heterogeneity and genomic configurations found within tumours by means of giving a different phenotype to each cell rather than grouping populations into a countable, finite set of possible genome configurations.

Within this minimal model, we aim at understanding how selection and mutation are coupled when instability, and thus the individual mutation probability μ_i, can itself change and affect the rate of cell replication r_i, and what the evolutionary consequences of this coupling are.

3 | SELECTION ON INSTABILITY

As discussed above, a most common event during the process of tumorigenesis are mutations in oncogenes that usually result in increased levels of replication (Vogelstein & Kinzler, 2002), thus giving to instability a role in activating the paths towards higher replicative capacity. On the other hand, the same elevated levels of instability can trigger deleterious mutations in genes that are vital for correct cellular metabolism and functioning, eventually leading to reduced cell viability or death. This apparent trade-off supposes the existence of a clear coupling between replicative capacity, cell viability and mutation probability that sits at the basis of tumour replication, evolvability and adaptation. We hereby propose a minimal adaptive landscape that translates such coupling into replication rates being a function of instability, $r(\mu)$.

3.1 | Adaptive landscape

Within our scope of producing a minimal model we expect to describe evolutionary dynamics on an adaptive landscape containing a reduced, treatable set of components. Taking into account the previously mentioned trade-off, these follow from considering the effect of mutations enhancing malignant cell replication, provided that such mutations have not damaged any of the necessary machinery for cell viability. We start by considering that mutations on oncogenes can translate into a linear increase in replication rate, such that $r(\mu) = r_0 + N_R \delta_R \mu$, with r_0 being the basal replication rate of normal cells, N_R the number of oncogenes responsible for increased replication and δ_R the mean effect on replication rate when mutating one of such genes. Following a linear approximation, we do not include a saturation term for the number of nonmutated oncogenes. This is actually consistent with early stages of tumour evolution, where only a small fraction of oncogenes has been affected and so N_R can be kept as a constant. In this picture, we need as well to take into account the minimal genetic material needed for a cell to keep its basic functions. If we group such material into a number of house-keeping genes, N_{HK}, the probability that none of them has been mutated is $(1-\mu)^{N_{HK}}$. Grouping both considerations together we obtain an analytical description of the coupling between replication and instability

$$r(\mu) = (r_0 + N_R \delta_R \mu)(1-\mu)^{N_{HK}} \quad (2)$$

(Solé et al., 2014). This adaptive landscape is of course of qualitative nature, and realistic fitness landscapes for unstable tumour environments are still far from our knowledge. However, certain points can be made if we give values within realistic parameter ranges to our function. The number of both oncogenes and house-keeping genes have been widely assessed, and we take them to be about $N_R \approx 140$ (Vogelstein et al., 2013) and $N_{HK} \approx 3804$ (Eisenberg & Levanon, 2013), respectively. Interestingly enough, considering small replication effects for δ_R, such experimental values produce an adaptive landscape (Figure 2) that has a positive gradient within the region of

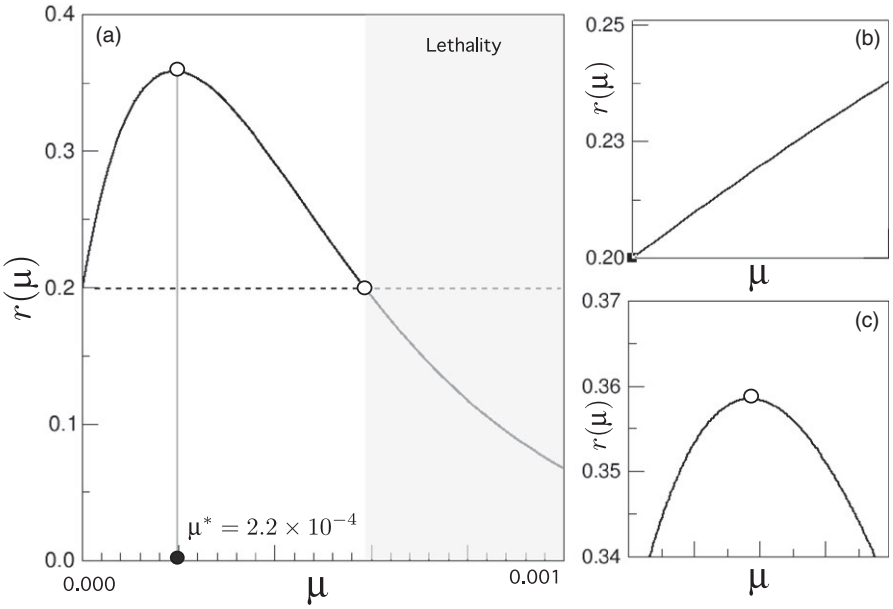

FIGURE 2 Fitness landscape function associated with the evolutionary dynamics of unstable tumour cells. In (a) the full landscape, given by a replication rate $r(\mu) = (r_0 + N_R\delta_R\mu)(1-\mu)^{N_{HK}}$, is plotted against the instability probability μ. Further discussion is focused on two limit cases representing initial linear progression of instability (b) and optimal mutation (c) domains

$\mu \in [10^{-9}, 10^{-4}]$, so that our evolutionary trajectories will be bounded within a region of instability levels in accordance with those experimentally measured for tumour cells (Tomlinson, Novelli, & Bodmer, 1996).

3.2 | Distribution of new mutations

We have assessed so far what is the effect of instability in proliferation, thus coupling mutation and selection for mutation level. Up next, we need to evaluate how does instability change during reproduction, so that we can finally compute the effects on replicative capacity of a mutated cell. As previously discussed, a broad range of mechanisms relates to variations in DNA replication fidelity. Such variations, however, are hardly in the direction of increasing DNA stability, and in general account for an increase in the mutation probability of cancer cells due to accumulation of further tumour-suppressor or care-taker gene mutations (Vogelstein & Kinzler, 2002).

This trend of generating more unstable offspring is translated into a positively skewed distribution of mutants $M(\mu,\Delta\mu)$. To keep the mathematical background of our model treatable, a Rayleigh distribution peaked at $\Delta\mu = 0$ has been chosen[1] . Under this scheme, instability of a daughter cell is likely to be similar or slightly higher from its parent, controlled by a scale parameter σ_μ^2 depicting the general size of mutational increases.

[1]The Rayleigh distribution is an asymmetric probability distribution defined for positive random variables (Forbes et al., 2010). We displaced so that its mode is zero, with shape

$$M(\Delta\mu;\sigma_\mu^2) = \frac{\Delta\mu+\sigma_\mu}{\sigma_\mu^2}\exp\left(\frac{-(\Delta\mu+\sigma_\mu)^2}{2\sigma_\mu^2}\right),$$

accounting for asymmetric probabilities of possible mutation levels, with small forward mutations being the most common.

4 | ADAPTIVE DYNAMICS

Adaptive dynamics is a set of techniques or a mathematical framework that models long-term phenotypic evolution of populations. Several works by different authors cover a broad scope of possible applications, and we hereby focus on the work of Dieckmann and Law and others (Champagnat et al., 2001; Dieckmann & Law, 1996) and adapt it to our particular system. The main biological background behind the maths sits in considering the evolutionary step as a mutant appearing and invading in a population in ecological or dynamical equilibrium (Dieckmann & Law, 1996). Under this picture, the ecological and evolutionary time scales are considered to be uncoupled, so that the process of the mutant competing against the resident population, and eventually fixating in it, is considered instantaneous in the evolutionary process.

General AD literature (see e.g., Champagnat et al., 2001; Dieckmann & Law, 1996; Geritz et al., 1998) follows the evolution of a quantitative phenotypic trait or set of traits, s, that can change through mutations. In these studies, the probability μ at which mutations appear is considered a possible function of the trait s, but afterwards and further on in the AD literature is usually left as a constant of each model. In the light of what we have discussed in the previous section, however, instability itself is a quantitative trait if computed as a mutation probability, and so the coupling of mutation and selection results in $s = \mu$ being the studied trait value.

The starting point of the AD modelling is to consider the evolutionary process, where the population's mutation probability changes as mutants appear and fixate, as a Markov chain for the probability of finding the population at time t having trait value μ

$$\frac{dP(\mu,t)}{dt} = \int \left(w(\mu|\mu')P(\mu',t) - w(\mu'|\mu)P(\mu,t)\right) d\mu'. \quad (3)$$

The transition probabilities $w(\mu'|\mu)$ describe the evolutionary step and contain the probability of the mutant with trait μ' appearing (A) and fixating (ρ) in the population, so that $w(\mu'|\mu) = A(\mu,\mu')\rho(\mu,\mu')$. The probability that a mutant appears is $A(\mu,\mu') = Nr(\mu)\mu M(\mu,\mu')$, the size of the population at equilibrium N, the probability of birth and mutation $r(\mu)\mu$ and the probability that the mutant has mutation probability μ' provided the parent cell had probability μ.

The probability $\rho(\mu,\mu')$ that a mutant with fitness advantage $r(\mu')/r(\mu)$ fixates in a population of N individuals has an analytical expression for the Moran model (Ewens, 2004)

$$\rho(\mu,\mu') = \frac{1-(r(\mu)/r(\mu'))}{1-(r(\mu)/r(\mu'))^N}. \tag{4}$$

A common procedure of the AD framework is to expand ρ for small variations of the trait value under the assumption of large populations, assumptions that are not a restriction for our problem. Under this view, the probability that the μ' mutant fixates is zero for $r(\mu') \leq r(\mu)$ and

$$\rho(\mu,\mu') = \frac{\mu'-\mu}{r(\mu)}\left(\frac{\partial r}{\partial \mu'}\right)_{\mu'=\mu} + O(\Delta\mu^2), \tag{5}$$

for $r(\mu') > r(\mu)$. Once a complete expression for the transition probabilities is build, we only need to recall how the evolution of the mean mutation probability can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\mu\rangle(t) = \int \mu \frac{\mathrm{d}}{\mathrm{d}t}P(\mu,t)\mathrm{d}\mu, \tag{6}$$

so that, using the original master equation, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\mu\rangle(t) = \int\int (\mu'-\mu)w(\mu'|\mu)P(\mu,t)d\mu'\mathrm{d}\mu = \langle a_1(\mu)\rangle, \tag{7}$$

where one recalls that $a_k(\mu) = \int (\mu'-\mu)^k w(\mu'|\mu)d\mu'$ is the k-th jump moment. If the first jump moment were a linear function of μ, then $\langle a_1(\mu)\rangle = a_1(\langle\mu\rangle)$ and the previous expression becomes directly treatable.[2]

The evolutionary trajectory for the mean path (we cease denoting it by angle brackets) will therefore follow

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu(t) = N\mu\frac{\partial r}{\partial \mu}\int (\mu'-\mu)^2 M(\mu,\mu')\mathrm{d}\mu' \tag{8}$$

[2]It is interesting to understand the implications of such condition and how do they relate to the assumptions of the AD method and the limitations of our model. Between many approaches (see e.g., van Kampen, 1981; Kubo, Matsuo, & Kitahara, 1973), and without pretending to expose here a deep discussion on this aspect, one might consider extracting a Fokker–Planck equation from the Markov chain (6). By means of computing the evolution of the mean value $\langle\dot{\mu}\rangle$ for that equation, it can be seen that the dynamics of $\langle\mu\rangle$ are only equivalent in both frameworks if either the selection gradient $\partial r/\partial\mu$ does not depend on μ (a linear adaptive landscape), or the population is strictly monomorphic on the trait μ. A detailed explanation of this considerations can be found in the Appendix section. These conditions will impose a strong constraint when considering a realistic tumour environment, so that it will become necessary along our computational experiments and the discussion to assess the regions of validity of our model.

At this point, we recall that only fitter mutants can invade and so may eventually contribute to the exploration of the adaptive landscape. This translates onto the domain of the integration being restricted to $\mu' > \mu$, and so we integrate the positive part of our σ^2_μ skewed mutant distribution.

These considerations of selection on instability and nonsymmetrical mutations result in our first-order approximation of the evolution of instability for a minimal cancer cell population:

$$\frac{\mathrm{d}\mu}{\mathrm{d}t} = \gamma N\sigma^2_\mu\mu\left(\frac{\partial r}{\partial\mu'}\right)_{\mu'=\mu} \tag{9}$$

where

$$\gamma = \frac{3}{\sqrt{e}} - \sqrt{\frac{\pi}{2}}\mathrm{erfc}\left(\frac{1}{\sqrt{2}}\right) \tag{10}$$

is simply a positive constant that results from integrating the asymmetric Rayleigh mutations distribution. Equation 9 defines the canonical equation for unstable cancer dynamics, describing the evolution of the mean mutation probability depending on population size, probability and effect of mutations and the steepness of the adaptive landscape defined by the effects of instability on cellular replication and viability.

5 | EVOLUTION OF INSTABILITY

The canonical Equation 9 describes the evolution of instability in our model population depending on the population size N, the distribution of mutation jumps σ^2_μ and the product of the mutation probability and the gradient of the adaptive landscape $\mu\,\partial_\mu r$. For our model landscape (Equation 5), this turns out to be

$$\begin{aligned}\frac{\mathrm{d}\mu}{\mathrm{d}t} = \gamma N\sigma^2_\mu\mu\big(&N_R\delta_R(1-\mu)^{N_{HK}} \\ &-N_{HK}(r_0+N_R\delta_R\mu)(1-\mu)^{N_{HK}-1}\big)\end{aligned} \tag{11}$$

Complicated analytical solutions for this equation might not give best insight of the underlying dynamics. However, as a first test of our model we compare its numerical solution to averaged Moran Process simulations (Figure 3). It is both relevant and useful to understand the factors that cause deviations between computer experiments and our analytical approach, in order to further comprehend the approximations on which AD is build.

In terms of parameter range, these are mostly translated into the population being large enough, and mutation probabilities being proportionally small. The second is easily met for both healthy and cancerous human cells, but simulating full-size clinically detectable tumours (more than 10^8 cells (Bozic et al., 2013)) is of large computational cost, and keeping our model and

exercise minimal, we have used smaller populations, modelling smaller subclones or spatially segregated populations where drift comes into play. Such drift produces a nonmonomorphic population where evolution deviates from the gradient trajectory and so proceeds slightly slower than our estimate. As previously stated and discussed along with the Appendix, the high nonlinearity of our landscape ensures that $\langle a_1(\mu)\rangle = a_1(\langle\mu\rangle)$ will be only valid up to a certain degree of approximation. It can be seen from Figure 3 that, still within this restricted range of validity, the canonical equation can capture the dynamics of instability up to a reasonable point.

A better understanding of the underlying dynamics can result from dividing the exploration of the landscape in well-behaved regions where simpler equations will arise.

On the one hand, in an initial phase of malignancy exploration for small values of μ, the shape of the adaptive landscape is dominated by the linear increase of mutated oncogenes, $r(\mu) = r_0 + N_R\delta_R\mu$. Within this region, dynamics of instability follow

$$\frac{d\mu}{dt} = \gamma N\sigma_\mu^2\mu N_R\delta_R \tag{12}$$

and the mean evolutionary trajectory is

$$\mu(t) = \mu_0 e^{\gamma\sigma_\mu^2 N N_R\delta_R t}. \tag{13}$$

It is remarkable to understand how, even in a linear adaptive landscape, the coupling between mutation and selection on unstable cells introduces a further nonlinearity that will account for exponential exploration of the space of instability and the consequent exponential increases in replication capacity. Such results can be again compared to computer simulations of mutator-replicator cells (Figure 4). The smaller nonlinearity also ensures that AD remains a good approximation despite stochastic deviation.

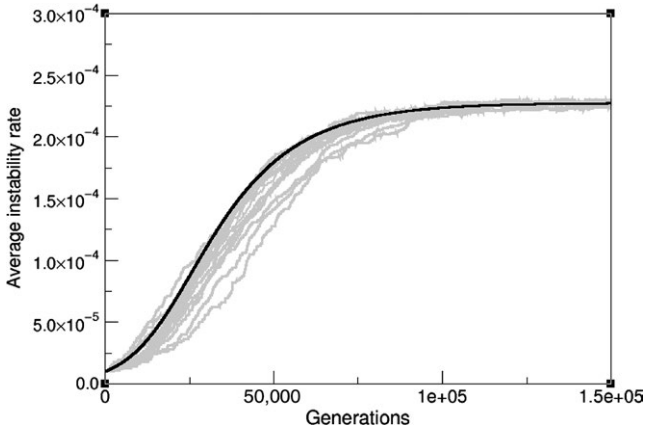

FIGURE 3 Evolutionary trajectories of the simulated Moran process (grey lines) and numerical solution of the predicted AD result (black curve), (population [N] = 2,000, distribution scale parameter [effect of mutations, σ] = 0.01, initial instability (μ_0) = 10^{-5})

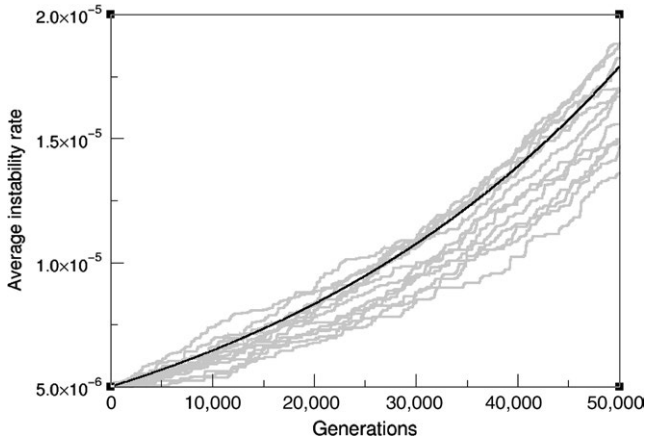

FIGURE 4 Exponential evolution of the mean mutation probability on a linear landscape: Moran process simulations (grey lines) of populations of 2,000 cells and the AD approximation (black curve), (σ = 0.01, $\mu_0 = 5 \times 10^{-6}$)

Another interesting point is to understand the behaviour of the mean instability levels as the population approaches the landscape peak. This kind of behaviour is easily studied if one considers a simple landscape containing a peak, such as $r(\mu) = r_0 + \delta_R N_R\mu - \delta_{HK}N_{HK}\mu^2$, where the role of house-keeping genes is not considered totally deleterious but just reducing fitness quadratically with the mutation probability. This landscape has an optimal value at $\mu^* = \delta_R N_R/2\delta_{HK}N_{HK}$, and this peak is explored through

$$\frac{d\mu}{dt} = \gamma N\sigma_\mu^2\mu(N_R\delta_R - 2\delta_{HK}N_{HK}\mu). \tag{14}$$

By means of rewriting this trajectory as $d\mu/dt = A\mu(B - C\mu)$, with $A = \gamma N\sigma_\mu^2$, $B = N_R\delta_R$ and $C = 2\delta_{HK}N_{HK}$, its solution simplifies to

$$\mu(t) = \frac{Be^{B(At+c_1)}}{Ce^{B(At+c_1)} + 1}, \tag{15}$$

where c_1 ensures that $\mu(0)=\mu_0$, the normal mutational probability of healthy cells. This trajectory saturates for long times at the expected result $A/B=\mu^*$, and can be again compared to computational experiments of replicating cells (Figure 5).

The same deviation between simulations and the numerical fit is found in this case, with evolution proceeding slower than our estimate. However, this minimal landscape approximation is able to capture the dynamical behaviour of our gene-related landscape model, mainly with an initial exponential growth followed by saturation around the peak, which can be proven to be an evolutionary stable strategy (Geritz et al., 1998).

Provided that the canonical equation has a nontrivial, singular point, as we found for $\mu^* = \delta_R N_R/2\delta_{HK}N_{HK}$, one can study the evolutionary stability of a quantitative trait. We can easily compute if this

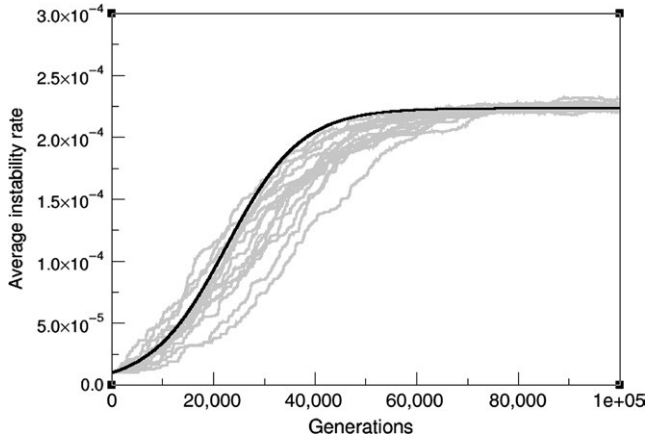

FIGURE 5 Mean mutation probability saturation at the fitness peak: Moran process simulations (grey lines) of populations of 2,000 cells and the AD approximation (black curve), ($\sigma = 0.01$, $\mu_0 = 10^{-5}$)

singular mutation probability will be an evolutionary trap, *that is*, a strategy that no further mutants can invade, if

$$\frac{\partial^2 r}{\partial \mu^2}_{\mu=\mu*} < 0 \tag{16}$$

which holds for our strategy: $\partial_{\mu\mu} r(\mu) = -2\delta_{HK} N_{HK} < 0$.

6 | DISCUSSION

In the present work, we have discussed the implications of the coupling between selection and instability for a minimal model of a population of mutating cells. We have shown how to determine the evolutionary trajectory for the mean instability levels in a basic landscape of cancer-related genes. Our AD model, as defined by our canonical equation (and consistently with simulated trajectories), describes the tempo and mode at which mutation probabilities increase and saturate around fitness peaks. For a simple but sensible fitness landscape, a general canonical equation has been derived from the Moran process scenario. Several approximations have also been considered.

A first relevant result of our model arises from evaluating the canonical equation for unstable cells in a linear landscape, to be associated with a premalignant stage. The nonlinearity resulting from the coupling of mutation and selection predicts an exponential increase of instability levels, whereas a trait different from instability would only increase linearly within such landscape. This result is presented as a mathematical description of genomic instability being an enabling characteristic of cancer, by means of generating fast exploration of the space of possible mutations towards malignancy. Similarly, we obtained consistent matchings between simulated and average predicted instability values for the near-optimum state. In this scenario, our model predicts an exponential increase followed by saturation around a critical mutational load, where, at least for this initial model of a nongrowing population, tumour cells are robust to further mutations. Considering that the distribution of mutational effects of cancer cells is hard to describe, it is important to understand that these results are qualitatively independent of the Rayleigh distribution, which we have only chosen in search for an asymmetric and analytically treatable function to work with. Other distributions would account for the same dynamics of exploration and saturation at different evolutionary paces. All in all, the possible applications of such minimal evolutionary descriptions of tumour instability follow from our set of examples and computer simulations.

Mounting evidence indicates that a successful approach to cancer therapy requires an explicit evolutionary perspective (Gatenby et al., 2009). One possible instance of this is provided by mutagenic therapies that have produced key results in the field of virology (Loeb et al., 1999). Would they be effective for cancer? Given some key analogies between RNA virus populations and unstable tumours (Solé & Deisboeck, 2004), this is an appealing possibility, although drug design or resistance mechanisms have yet to be assessed (Fox & Loeb, 2010). Prior to that, conceptual questions arise, such as do cancer cells live near critical instability levels, beyond which viability is no longer possible? is there a sharp error threshold for the mutation probability? what evolutionary outcomes should we expect when inducing variations on the mutational load of cancer cells, and how can these shed new light on mutagenic therapy?

Regarding the later, our model allows to bring instability as the evolving trait, while providing potential insights, particularly before and beyond the optimal instability levels. The exponentially fast increase of small mutational loads indicates that reducing instability levels in hope for progression delay might result in rapid re-exploration of the mutator phenotype. On the other hand, pushing instability beyond optimal levels, even if a critical point is not trespassed (Solé & Deisboeck, 2004), might render tumour cells too unstable, and there exist relevant efforts towards using DNA repair inhibitors to produce critical instability levels (Helleday et al., 2008).

Our model differs from previous work in its simple analytical formulation, which do not depend on chosen parameter ranges, such as those of Datta et al., 2013 and Asatryan & Komarova, 2016; and they are thus qualitatively robust. On the one hand, this means that we are able to obtain analytical expressions for the exponential evolution and saturation of the mutation probability, which could eventually be used when studying *in vitro* long-term evolutionary experiments with cancer populations, using serial transfer methods similar to those performed on viruses (see e.g., Drake, 1993; Sanjuán et al., 2010; Solé et al., 1999) or bacterial populations (see e.g., Moxon et al., 1994; Sniegowski, Gerrish, & Lenski, 1997; Barrick et al., 2009 for experiments and Taddei et al., 1997 for an early model for mutator alleles). Given the remarkable similarities found between microbial communities found both in the ecological and

evolutionary time scales (Lambert et al., 2011), it would be worth exploring the evolution of instability of cancer cell cultures over many transfer generations (Langdon, 2004).

It remains an open question to analyse if any sort of saturating dynamics occur for both the fitness or the mutation probability when these experiments are performed on malignant cells. Furthermore, interesting theoretical approaches have been performed to infer the underlying adaptive landscape from the observable evolution of traits (Kryazhimskiy, Tkacik, & Plotkin, 2009). This seems a plausible point regarding how our model directly relates dynamics and landscape gradient and could therefore shed light onto understanding the evolutionary pressures underlying genetic instability at each stage of tumour progression.

On the other hand, while trying to produce a model that can be treated without the use of complex mathematical tools, we have been constrained to leaving aside many relevant considerations, the one we are most concerned with is the lack of growing population dynamics. This leaves our model interesting for the previously discussed specific confined experiments, while not yet complete when trying to study three-dimensional growing tumours. While studying modifications to our formalism, following the work of evolutionary game theory on growing populations (see e.g., Li et al., 2015; Melbinger, Cremer, & Frey, 2010), we have decided to present this basic model as it remains a first step into a comprehensible and qualitative insight for the dynamics of populations able to evolve their mutation probability.

ACKNOWLEDGEMENTS

The authors thank P. Ruiz and E. Beltran, as well as the CSL laboratory members for useful discussions. Special thanks to Oliver Law for his inspiring ideas. This work has been supported by the Botín Foundation by Banco Santander through its Santander Universities Global Division, a MINECO grant FIS2015-67616 fellowship, by the Universities and Research Secretariat of the Ministry of Business and Knowledge of the Generalitat de Catalunya and the European Social Fund and by the Santa Fe Institute.

CONFLICT OF INTEREST

None declared.

ORCID

Ricard Solé http://orcid.org/0000-0001-6974-1008

REFERENCES

Asatryan, A. D., & Komarova, N. L. (2016). Evolution of genetic instability in heterogeneous tumors. *Journal of Theoretical Biology*, *396*(Suppl. C), 1–12. https://doi.org/10.1016/j.jtbi.2015.11.028

Barrick, J. E., Yu, D. S., Yoon, S. H., Jeong, H., Oh, T. K., Schneider, D., … Kim, J. F. (2009). Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, *461*, 1243–1247. https://doi.org/10.1038/nature08480

Bozic, I., Reiter, J. G., Allen, B., Antal, T., Chatterjee, K., Shah, P., … Nowak, M. A. (2013). Evolutionary dynamics of cancer in response to targeted combination therapy. *eLife 2*, e00747.

Champagnat, N., Ferriere, R., & Ben Arous, G. (2001). The canonical equation of Adaptive Dynamics: a mathematical view. *Selection*, *2*, 73–83.

Datta, R. S., Gutteridge, A., Swanton, C., Maley, C. C., & Graham, T. A. (2013). Modelling the evolution of genetic instability during tumour progression. *Evolutionary Applications*, *6*, 20–33. https://doi.org/10.1111/eva.12024

Dieckmann, U., & Law, R. (1996). The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, *34*, 579–612. https://doi.org/10.1007/BF02409751

Drake, J. W. (1993). Rates of spontaneous mutation among RNA viruses. *Proceedings of the National Academy of Sciences*, *90*, 4171–4175. https://doi.org/10.1073/pnas.90.9.4171

Eisenberg, E., & Levanon, E. Y. (2013). Human housekeeping genes, revisited. *Trends in Genetics*, *29*, 569–574. https://doi.org/10.1016/j.tig.2013.05.010

Ewens, W. J. (2004). *Mathematical population genetics*. Interdisciplinary Applied Mathematics. New York, NY: Springer.

Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2010). Rayleigh distribution. In: *Statistical distributions* (pp. 173–175). John Wiley Sons, Inc. https://doi.org/10.1002/9780470627242.ch39

Fox, E. J., & Loeb, L. A.. (2010). Lethal mutagenesis: Targeting the mutator phenotype in cancer. In: *Seminars in cancer biology 20.5*. Mutator Phenotype in Cancer: Origins and Consequences, 353–359.

Gatenby, R. A. (1991). Population Ecology Issues in Tumor Growth. *Cancer Research*, *51*, 2542–2547.

Gatenby, R. A. (1995). Models of Tumor-Host Interaction as Competing Populations: Implications for Tumor Biology and Treatment. *Journal of Theoretical Biology*, *176*, 447–455. https://doi.org/10.1006/jtbi.1995.0212

Gatenby, R. A., & Brown, J. (2017). Mutations, evolution and the central role of a self-defined fitness function in the initiation and progression of cancer. *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer 1867*, 162 –166. https://doi.org/10.1016/j.bbcan.2017.03.005

Gatenby, R. A., Silva, A. S., Gillies, R. J., Frieden, B. R. (2009). Adaptive Therapy. *Cancer Research*, *69*, 4894–4903. https://doi.org/10.1158/0008-5472.CAN-08-3658

Geritz, S. A. H., Mesze, G., Metz, J. A. (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology*, *12*, 35–57. https://doi.org/10.1023/A:1006554906681

Gerlee, P., & Anderson, A. R. A. (2015). The evolution of carrying capacity in constrained and expanding tumour cell populations. *Physical Biology*, *12*, 056001. https://doi.org/10.1088/1478-3975/12/5/056001

Greaves, M., & Maley, C. C. (2012). Clonal evolution in cancer. *Nature*, *481*, 306–313. https://doi.org/10.1038/nature10762

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, *144*, 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Helleday, T., Petermann, E., Lundin, C., Hodgson, B., & Sharma, R. A. (2008). DNA repair pathways as targets for cancer therapy. *Nature Reviews Cancer*, *8*, 193–204. https://doi.org/10.1038/nrc2342

van Kampen, N. G. (1981). *Stochastic processes in physics and chemistry*. Amsterdam: North Holland.

Komarova, N. L., Sadovsky, A. V., & Wan, F. Y. M. (2008). Selective pressures for and against genetic instability in cancer: a time-dependent problem. *Journal of The Royal Society Interface*, *5*, 105–121. https://doi.org/10.1098/rsif.2007.1054

Komarova, N. L., Lengauer, C., Vogelstein, B., & Nowak, M. A. (2002). Dynamics of Genetic Instability in Sporadic and Familial Colorectal Cancer. *Cancer Biology & Therapy*, *1*, 685–692. https://doi.org/10.4161/cbt.321

Kryazhimskiy, S., Tkacik, G., & Plotkin, J. B. (2009). The dynamics of adaptation on correlated fitness landscapes. *Proceedings of the National Academy of Science of the United States of America*, *106*, 18638–18643. https://doi.org/10.1073/pnas.0905497106

Kubo, R., Matsuo, K., & Kitahara, K. (1973). Fluctuation and relaxation of macrovariables. *Journal of Statistical Physics*, *9*, 51–96. https://doi.org/10.1007/BF01016797

Lambert, G., Estévez-Salmeron, L., Oh, S., Liao, D., Emerson, B. M., Tlsty, T. D., & Austin, R. H. (2011). An analogy between the evolution of drug resistance in bacterial communities and malignant tissues. *Nature Reviews Cancer*, *11*, 375–382. https://doi.org/10.1038/nrc3039

Langdon, S. P. (ed.) (2004). Basic principles of cancer cell culture. In *Cancer cell culture: Methods and protocols* (pp. 3–15). Totowa, NJ: Humana Press.

Lengauer, C., Kinzler, K. W., & Vogelstein, B. (1998). Genetic instabilities in human cancers. *Nature*, *396*, 643–649. https://doi.org/10.1038/25292

Li, X.-Y., Pietschke, C., Fraune, S., Altrock, P. M., Bosch, T. C., & Traulsen, A. (2015). Which games are growing bacterial populations playing? *Journal of The Royal Society Interface 12*, 20150121.

Loeb, L. A. (2001). A Mutator Phenotype in Cancer. *Cancer Research*, *61*, 3230–3239.

Loeb, L. A., Essigmann, J. M., Kazazi, F., Zhang, J., Rose, K. D., & Mullins, J. I. (1999). Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proceedings of the National Academy of Science of the United States of America 96*. Mutator Phenotype in Cancer: Origins and Consequences, 1492–1497.

Melbinger, A., Cremer, J., & Frey, E. (2010). Evolutionary Game Theory in Growing Populations. *Physical Review Letters*, *105*(17), 178101. https://doi.org/10.1103/PhysRevLett.105.178101

Moran, P. A. P. (1958). Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, *54*, 60–71. https://doi.org/10.1017/S0305004100033193

Moxon, E. R.Lenski, R. E., & Rainey, P. B. (1994). Adaptive evolution of highly mutable loci in pathogenic bacteria. *Current Biology*, *4*, 24–33. https://doi.org/10.1016/S0960-9822(00)00005-1

Negrini, S., Gorgoulis, V. G., & Halazonetis, T. D. (2010). Genomic instability – an evolving hallmark of cancer. *Nature Reviews Molecular Cell Biology*, *11*, 220–228. https://doi.org/10.1038/nrm2858

Nowak, M. A.Komarova, N. L., Sengupta, A., Jallepalli, P. V., Shih, Ie. M., & Lengauer, C. (2002). The role of chromosomal instability in tumor initiation. *Proceedings of the National Academy of Science of the United States of America*, *99*, 16226–16231. https://doi.org/10.1073/pnas.202617399

Nowell, P. (1976). The clonal evolution of tumor cell populations. *Science*, *194*, 23–28. https://doi.org/10.1126/science.959840

Sanjuán, R.Nebot, M. R., Chirico, N., Mansky, L. M., & Belshaw, R. (2010). Viral Mutation Rates. *Journal of Virology*, *84*, 9733–9748. https://doi.org/10.1128/JVI.00694-10

Sniegowski, P. D., Gerrish, P. J., & Lenski, R. E. (1997). Evolution of high mutation rates in experimental populations of E. coli. *Nature*, *387*, 703–705. https://doi.org/10.1038/42701

Solé, R. V., & Deisboeck, T. S. (2004). An error catastrophe in cancer? *Journal of Theoretical Biology*, *228*, 47–54. https://doi.org/10.1016/j.jtbi.2003.08.018

Solé, R. V., Ferrer, R., González-García, I., Quer, J., & Domingo, E. (1999). Red Queen Dynamics, Competition and Critical Points in a Model of RNA Virus Quasispecies. *Journal of Theoretical Biology*, *198*, 47–59. https://doi.org/10.1006/jtbi.1999.0901

Solé, R. V., Valverde, S., Rodriguez-Caso, C., & Sardanyés, J. (2014). Can a minimal replicating construct be identified as the embodiment of cancer? *BioEssays 36*, 503–512. https://doi.org/10.1002/bies.201300098

Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P. H., Godelle, B. (1997). Role of mutator alleles in adaptive evolution. *Nature*, *387*, 700–702. https://doi.org/10.1038/42696

Tomlinson, I. P. M., Novelli, M. R., & Bodmer, W. F. (1996). The mutation rate and cancer. *Proceedings of the National Academy of Science of the United States of America*, *93*, 14800–14803. https://doi.org/10.1073/pnas.93.25.14800

Vogelstein, B., & Kinzler, K. W. (2002). *The genetic basis of human cancer*. New York: McGraw-Hill.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A. Jr, & Kinzler, K. W. (2013). Cancer Genome Landscapes. *Science*, *339*, 1546–1558. https://doi.org/10.1126/science.1235122

How to cite this article: Aguadé-Gorgorió G, Solé R. Adaptive dynamics of unstable cancer populations: The canonical equation. *Evol Appl*. 2018;11:1283–1292. https://doi.org/10.1111/eva.12625

APPENDIX

LINEARITY OF $\langle a_1(\mu) \rangle$

As discussed when introducing the AD approach, we have considered that the first jump moment is linear in μ to obtain the canonical equation. Understanding the mathematical basis behind such condition can give further insight into the assumptions our model sits on. Let us start again from a Master equation that describes the evolutionary biased random walk: the trait substitution sequence that we use to model evolution. This is

$$\frac{dP(\mu,t)}{dt} = \int \left(w(\mu|\mu')P(\mu',t) - w(\mu'|\mu)P(\mu,t) \right) d\mu' \tag{17}$$

We knew that $w(\mu, \mu')$, the transition probability, was the probability of a mutant appearing times the probability of a mutant surviving. If we consider the mentioned Moran process as a good evolutionary constant population model, the transition probabilities are those of our previous AD model

$$w(\mu,\Delta\mu) = N(\mu)r(\mu)\mu M(\mu,\Delta\mu)\frac{\Delta\mu}{r(\mu)}\frac{\partial r}{\partial \mu} \tag{18}$$

and, supposing for simplicity that M is a symmetric distribution, we can obtain a Fokker–Planck transport equation of the form

$$\frac{\partial p(\mu,t)}{\partial t} = -\frac{1}{2}N\sigma_\mu^2 \frac{\partial}{\partial \mu}\left(\mu \frac{\partial r}{\partial \mu} p(\mu,t)\right) \tag{19}$$

The question is now: Under which assumptions the equation of the first moment of this distribution gives rise to the Canonical equation? Understanding these assumptions can give better insight into what are we actually doing with the canonical equation.

We can compute the dynamics of the first moment of the Fokker–Planck equation by means of multiplying by μ and integrating over the phase space, obtaining

$$\frac{d\langle\mu\rangle}{dt} = \frac{1}{2}N\sigma_{\mu}^{2}\langle\mu\frac{\partial r}{\partial\mu}\rangle, \tag{20}$$

which will only produce the Canonical equation if either the landscape is linear (and so $\langle\mu r'\rangle \propto k\langle\mu\rangle$) or the population is Dirac-distributed (monomorphic), and so

$$\langle\mu\frac{\partial r}{\partial\mu}\rangle = \int \mu\frac{\partial r}{\partial\mu}\delta(\mu-\langle\mu\rangle)d\mu = \mu\frac{\partial r}{\partial\mu}\Big|_{\mu=\langle\mu\rangle} \tag{21}$$

This can serve as a basic mathematical description of why do populations have to be monomorphic in the AD framework for the canonical equation to arise, and why simulations with highly nonlinear landscapes deviate from our model approximation.

Aguadé-Gorgorió and Solé *Journal for ImmunoTherapy of Cancer* (2019) 7:345
https://doi.org/10.1186/s40425-019-0795-6

Journal for ImmunoTherapy of Cancer

RESEARCH ARTICLE Open Access

Genetic instability as a driver for immune surveillance

Guim Aguadé-Gorgorió[1,2*] and Ricard Solé[1,2,3*]

Abstract

Background Genetic instability is known to relate with carcinogenesis by providing tumors with a mechanism for fast adaptation. However, mounting evidence also indicates causal relation between genetic instability and improved cancer prognosis resulting from efficient immune response. Highly unstable tumors seem to accumulate mutational burdens that result in dynamical landscapes of neoantigen production, eventually inducing acute immune recognition. How are tumor instability and enhanced immune response related? An important step towards future developments involving combined therapies would benefit from unraveling this connection.

Methods In this paper we present a minimal mathematical model to describe the ecological interactions that couple tumor adaptation and immune recognition while making use of available experimental estimates of relevant parameters. The possible evolutionary trade-offs associated to both cancer replication and T cell response are analysed, and the roles of mutational load and immune activation in governing prognosis are studied.

Results Modeling and available data indicate that cancer-clearance states become attainable when both mutational load and immune migration are enhanced. Furthermore, the model predicts the presence of well-defined transitions towards tumor control and eradication after increases in genetic instability numerically consistent with recent experiments of tumor control after Mismatch Repair knockout in mice.

Conclusions These two main results indicate a potential role of genetic instability as a driver of transitions towards immune control of tumors, as well as the effectiveness of increasing mutational loads prior to adoptive cell therapies. This mathematical framework is therefore a quantitative step towards predicting the outcomes of combined therapies where genetic instability might play a key role.

Keywords: Genetic instability, Neoantigen load, Mismatch repair, Immune surveillance, Combination therapies

Background

Cancer is a disease resulting from Darwinian evolution in cellular tissues[1]. Following depletion of a vast set of genetic insults altering normal multicellularity phenotypes, rogue cells are able to adapt and evade selection barriers leading to uncontrolled proliferation. In this context, genomic instability plays a key role as a driver of the genetic novelties required for tumor progression and rapidly adapting phenotypes [2, 3]. High levels of evolving instability sustain a very diverse population [4], and intratumor heterogeneity lies at the very core of why cancer is still difficult to define, characterize and cure [5].

In this paper we aim at understanding an important relationship between the effectiveness of cancer immunotherapy and genetic instability. The relevance of such link needs to be found in the challenges faced by immunotherapies based on immune checkpoint inhibition or adoptive cell transfer [6], where mutational burden seems to play a key role. Due to the underlying complexity of cancer immunology, interdisciplinary efforts towards novel immunotherapies are much required [7–9]. As discussed below, the crucible of the problem might be to the nonlinear dynamics associated to cancer neoantigen production and the consequent enhancement of immune surveillance.

*Correspondence: guimaguade@gmail.com; ricard.sole@upf.edu
[1]ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain
[2]Institut de Biologia Evolutiva (CSIC-UPF), Psg Maritim Barceloneta, 37, 08003 Barcelona, Spain
Full list of author information is available at the end of the article

A key point in cancer immunotherapy lies on the mechanisms by which T cells actually recognize cancerous from healthy tissue [10] and eventually attack tumor cells expressing tumor-specific antigens [11]. On a general basis, such antigens can be common proteins for which T cell acceptance is incomplete, or more importantly, novel peptides [10, 12]. Except for specific tumor types of viral etiology, these so-called *neoantigens* arise after DNA damage resulting in the production of novel proteins. Recent advances highlight the importance of understanding neoantigen generation as a consequence of the tumor mutational load and dissecting specific neoantigen immunogeneicity [10, 11, 13]. Furthermore, direct correlations have been suggested between neoantigen production at high microsatellite instability, eventual immune surveillance and clinical response to immunotherapies [14–16].

Several experimental and clinical sources are pointing towards a causal relation, including tumor growth impairment after inactivation of MLH1 [17], or the positive response to PD-1 blockade across different mismatch repair (MMR) defficient cancer types [18]. The inactivation of MMR results in increased mutational burden of cancer cells, promoting the generation of neoantigens which improve immune surveillance and eventual tumor arrest. These obxservations suggest a novel view on immunotherapy, where targeting mutagenic pathways can result in an alternative mechanism to unleash immune responses [9, 19].

All in all, genetic instability seems to play a conflictive role in cancer evolution and proliferation. It appears that the same genome alterations that activate cancer progression can trigger T cell recognition and immune attack. The extent of such trade-off and its application to therapy, however, is not clear. On the one hand, mutagenic therapies coexist with an intrisic risk, as increased genetic instability on heterogeneous populations might activate oncogenic outgrowth in previously stable cells. Moreover, a reactive immune system might pose a selective pressure for immune editing, leading to selection for T cell evading tumor subclones. How do these two components -instability and immune response- interact and what are the consequences? Is it possible to provide useful insight from mathematical models without a detailed picture of the immune landscape of cancer?.

Nonlinear responses associated to cancer-immune system interactions have been known from the early days of cancer modelling, from more classical approaches [20] to recent perspectives based on neoantigen recognition fitness [21]. These studies have revealed a number of interesting properties exhibited by toy models, including in particular the existence of shifts and breakpoints separating cancer progression from its extinction (see [22] and references therein). Such shifts are of exceptional importance in our context: they indicate the existence of well defined conditions (and perhaps therapeutic strategies) allowing an all-or-none response. However, a mathematical description of the specific role of genetic instability in cancer immunology has not yet been developed. Below we provide a first approach to such goal, based on considering both cancer adaptation and immune surveillance as influenced by mutational burden, and we analyze how genetic instability can account for transitions towards states of cancer control and elimination. The implications of these transitions on combination therapies are discussed, pointing towards possible cross-therapies activating neoantigen production and immune stimulation.

Methods

Population dynamics of the tumor-immune interaction

The ecology of the cancer-immune system interaction pervades several complexity levels, from a vast antigenome [23] to multilayer cellular competition dynamics [24], and a first step towards modeling such ecology lies in dissecting which specific ingredients are key drivers in the phenomena we aim to understand.

Recent research points out that there might be up to 28 immune cell types with both antitumor and immunosupressive roles infiltrated within a tumor [25]. Focusing on the immuno-surveillance mechanism of tumor growth inhibition following immune system recognition (early introduced in [26]), a minimal modelling approach recalls at least considering a population of tumor cells growing in competition with immune cells. It is commonly accepted that the immune response to cancer is mostly driven by an adaptive cohort of *cytotoxic immune cells*, such as $CD8^+$ T cells, together with a cellular compartment of the *innate* immune system such as NK cells [27, 28]. Despite this work focuses on the adaptive response to neoantigen presentation, including an innate effector response will allow for understanding relevant non-antigenic immune effects.

Even if further models have been useful at depicting very advanced properties of the immune system [29], we have chosen to keep a minimal scenario able to describe the competition dynamics at play. We apply a well characterised model (see e.g. [30]) that has been used to account for experimental results in cancer immunology such as tumor-immune equilibrium [31]. This model has been studied using parameter ranges measured from experimental setups consistent with several tumor types (Table 1, see [20, 32]).

The cellular interactions considered here involve a commonly used well-mixed (mean-field) model [20, 22] where the population of cancer cells c follows a logistic growth (at effective replication rate $r = b - d$ and carrying capacity K) and immune-cell mediated death (at rate δ_c). This saturating growth model captures several

Table 1 Parameter values for the cancer-immune ecology model, estimated from experimental data of BCL_1 lymphoma in the spleen of chimeric mice (see [20])

Parameter	Meaning	Kuznetsov et al. (1994) estimate
r	Cancer cell replication rate	$0.18 day^{-1}$
K	Tumor carrying capacity	$2 \times 10^9 cells$
δ_c	Immune-mediated cancer cell death rate	$1.101 \times 10^{-7} day^{-1} cells^{-1}$
δ_E	Cancer-mediated immune cell death rate	$3.422 \times 10^{-10} day^{-1} cells^{-1}$
m	Rate of T cell migration towards tumor site	$1.3 \times 10^4 cells\ day^{-1}$
g	Tumor size limit for effective T cell infiltration	$2.019 \times 10^7 cells$
ρ	Rate of cancer cell recognition by the immune system	$0.1245 day^{-1}$
d	Intrinsic T cell death rate	$0.0412 day^{-1}$

tumor microenviroment effects of malignant cell competition and death, such as spatial constraints or nutrient availability [33].

$$\frac{dc}{dt} = rc\left(1 - \frac{c}{K}\right) - \delta_c cE. \tag{1}$$

The effector immune population includes both NK and T cell compartments. Despite further modeling has been able to capture specific dynamics of T cell activation by cancer-NK cell encounter [27], activation of both cell types by malignancy can be described in a similar form [22], here described by

$$\frac{dE}{dt} = m + \rho\left(\frac{c}{g + c}\right)E - \delta_E cE - dE, \tag{2}$$

In this framework, the innate and adaptive immune populations are encapsulated into a single Effector compartment that grows due to a constant migration of cells and a predation term ρ that is commonly acknowledged to obey a Michaelis-Menten-like saturation due to limitations in immune cell circulation through the tissue [20] and penetration within the solid tumor [32, 34]. The peculiarity of the model lies in considering this predation term different for both NK and T cells. As discussed below, ρ is split into a constant rate refering to innate NK predation (see [27] and references therein) together with a variable part that will relate to antigen recognition by T cells, so that $\rho = \rho_{NK} + \rho_T$. Effector cells also have a natural decay rate, d, and die when competing with tumor cells at a rate $-\delta_E c$. The complete set of interactions described by (1) and (2) is schematically shown in Fig. 1.

Ecological trade-offs in genetic instability

As discussed above, genetic instability plays a key role in tumor evolution, acting as the driving mechanism towards phenotypic variation and adaptation. Within our model, this can be translated as the replication rate being a function of its level of genetic instability μ. On the other hand ρ_T, the rate of cancer cell recognition by T cells, is also μ-dependent because of neoantigen production. Below we propose a minimal characterization of r and ρ able to describe how genetic instability modulates such trade-off.

Cancer adaptation as a function of genetic instability

Cancer adaptation, here summarized to modulations in its replication rate, stems from the phenotypic plasticity resulting from mutations and copy-number alterations. On a general basis, enhanced tumor replication follows from mutations affecting oncogenic pathways, which poses a trade-off on genetic instability as it can, as well, damage any of the necessary machinery for cell viability.

Following previous research [35, 36], an adaptive landscape is build on several assumptions based on the probabilities of mutating oncogenic and house-keeping genes.

Genetic instability has a twofold impact on cell fitness. Specifically, replication rate r will be considered a function of mutation probability μ. A landscape $r(\mu)$ is now in place [35, 37], and follows from considering that mutations on oncogenes can translate into a linear increase in replication rate. This follows from assuming that reproductive effects of oncogenes, as for advantageous mutations on many systems, are exponentially distributed [38], so that their sum is gamma distributed with average increasing with the number of mutated oncogenes. This will be expressed as $R_1(\mu) = r_0 + N_R \delta_R \mu$ with r_0 being the basal replication rate of normal cells, N_R the number of oncogenes responsible for increased replication and δ_R the mean effect on replication rate when mutating one of such genes.

To account for cell viability, the number of house-keeping genes N_{HK} is taken into account so that mutations affecting them result in null replication [39]. This introduces the constraint of not having any of them mutated, $R_2(\mu) = (1 - \mu)^{N_{HK}}$. Grouping both considerations together we obtain an analytical description of the coupling between replication rate and mutation probability $r(\mu) = R_1(\mu)R_2(\mu)$ which reads:

$$r(\mu) = (r_0 + N_R \delta_R \mu)(1 - \mu)^{N_{HK}} \tag{3}$$

This adaptive landscape is of course of qualitative nature, and realistic fitness landscapes for unstable tumor environments are still far from our knowledge. However, certain points can be made if we give values within realistic parameter ranges to our function. The number of both

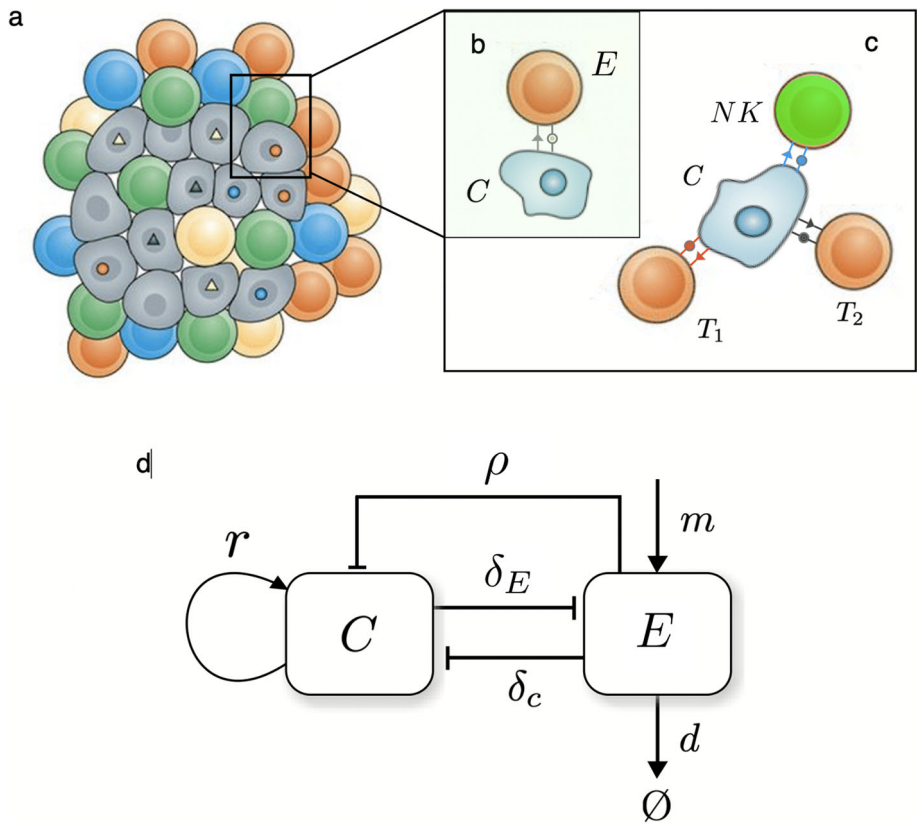

Fig. 1 A schematic summary of the basic cancer-immune cell-cell interactions. The two key components are (**a**) a tumor population driven by genetic instability and (**b**–**c**) interactions associated to tumor cell recognition and attack by T and NK cells. The strength T cell attack depends on the number of surface neoantigens (**c**), while NK killing is constant [27]. In (**d**) the population-level interaction diagram is displayed based on the model in [20]. Here *c* and *E* indicate the number of cancer and T and NK cells, respectively. Cancer cells grow at a rate *r* (and have a limited carrying capacity) while immune cells enter the system at a constant production rate *m* and react at malignant cells at a rate ρ that will be different for NK cells and instability-dependent T cell recognition. A constant average death rate *d* is associated with their removal. Two constant cross-interactions rates are also indicated as δ_T and δ_c associated to the removal efficiency of cancer cells and the death of immune cells resulting from the same process, respectively

oncogenes and house-keeping genes have been widely assessed, and we take them to be about $N_R \approx 140$ [40] and $N_{HK} \approx 3804$ [39] respectively. Interestingly, considering small replication effects for δ_R, such experimental values produce an adaptive landscape that has an optimal region for tumor replication at about $\mu \approx 10^{-5} - 10^{-4}$, which is in accordance with the point-mutation probability levels experimentally measured for unstable tumor cells [41].

Immune recognition of malignancy as a function of genetic instability

Building a mathematical description of how the immune system reacts at the mutational burden of cancer cells is not straightforward. This stems from the fact that such behavior is yet starting to be understood at the molecular level and it probably builds upon many layers of complexity [10]. In our minimal mathematical approach, the first step is describing immune reactivity as proportional to the adaptive compartment of cancer cell recognition ρ_T, a rate that itself depends on the dynamics of neoantigen expression. Under our assumptions, since adaptive immune response follows from neoantigen detection we expect ρ_T being a function of the overall mutational landscape of a tumor, μt, which is eventually responsible for such neoantigen dynamics. Following recognition probability distributions from [21], we expect the average dominance to initially increase with mutations as more and more neoantigens are generated and eventually saturate as very dominant neoantigens are rare.

The mathematical shape of this dependency $\rho_T(\mu t)$ could stem from purely stochastic dynamics, but recent research gives better insight into the shape of this correlation. Rooney and colleagues provided an enlighting perspective in this direction by comparing a measure of immune response from the transcript levels of two key

cytolytic effectors with the total mutation count for eight tumor types [42].

Cytolytic response strengths in [42] seem to indicate a dependency on tissue and tumor microenviroment, which we have not included in our study since our model is not tumor type-specific. For each tumor type, a least-squares linear regression is used (Melanoma in Fig. 2). When comparing across tumor types the shape of the immune response seems to obey a common pattern across many cancers, once cytolytic response values are normalized (Table 2). A linear relation can be found for which normalized cytolytic activity scales with mutational load as $\mathrm{CYT} \sim 4.35 \times 10^{-4} \mu t$ when averaged across the range of tumor types explored here. However, we expect a function depending only on mutation probability. The variable t in this expression refers to the evolutionary life history of mutations accumulation of the tumor. This time scale is much larger than the faster ecological dynamics that govern the cancer-immune system interactions, so that we can consider it an average measure of tumor age at the time of detection, and consider it constant when introducing ρ in the ecological dynamics. From these facts, the only variable governing immune recognition at the cancer-immune competition level is the point mutation probability μ.

A very rough estimate for t could be either inferred from average cell replication data or from the fact that values for the mean mutation rate and the absolute mutational load are known for many tumors [43]. For example, we can use the notion that mutator tumors have mutation rates of about 10^{-5} mutations per gene per cell division [44], which account for the accumulation of about 10^3 somatic mutations per tumor life [42], so that average tumor divisions lies at about $t \sim 10^7$. Using this approximation we obtain our preliminar expression for how the immune reactivity rate depends on the mutation levels, $\rho_T(\mu) = 4.35 \times 10^3 \mu$.

In this first correlation measure from [42], however, immune recognition grows constantly with mutational load. This growth should not be indefinite, and many factors counteract the cytolytic effect of antigen-producing mutations. As an example, increases in genetic instability can also account for antigen silencing and immune editing, which itself would reduce cytolytic activity [45]. All in all, it seems plausible to consider that antigenic and immune-suppressing mutations could balance beyond certain mutational threshold. Following data from [42] it seems that the tumor-immune cytolytic interaction is far from saturation, with an estimated saturation behavior to happen beyond $\mu \sim 10^{-4}$, a mutational level higher than those of most tumors measured by recent methodologies (see e.g. [42]). This saturating function follows the same trend of the data-based linear relationship and reads

$$\rho_T(\mu) = 4.35 \times 10^3 \mu \sim \frac{2}{1.4 + e^{14000\mu}} - \frac{5}{6}, \tag{4}$$

and can be compared with tumor adaptability $r(\mu)$ (Fig. 3) to obtain a full mutational landscape for tumor progres-

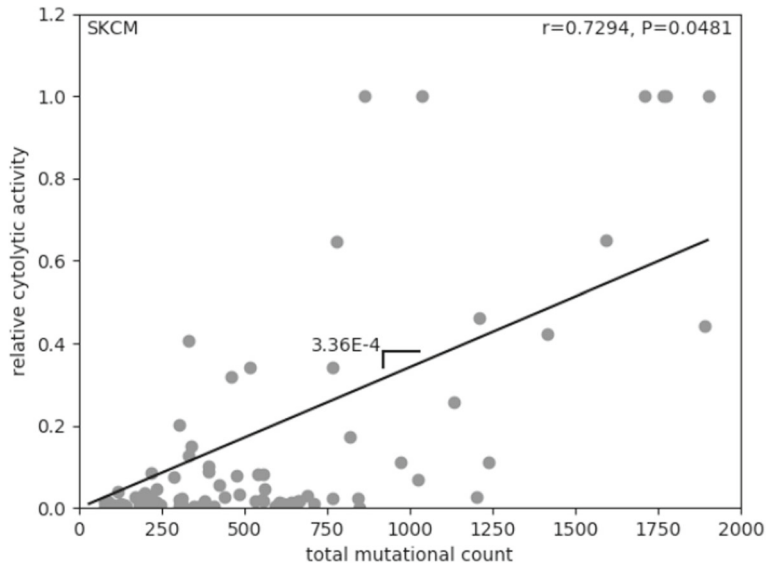

Fig. 2 Measuring immune reactivity as a function of the mutational load. Melanoma is plotted as an example, where a linear regression (black line, scale=3.36E-4) between total mutation count and relative cytolytic activity is evaluated. Results for 12 cancer types in Table 2. Data is obtained from [42]. As in the original work, the correlation spans the 5^{th} to 95^{th} percentile of the mutation count variable

Table 2 Linear regressions for $\rho(\mu t)$ across 12 cancer types, resulting in $\rho(\mu t) = 4.35 \times 10^{-4} \mu t$

Cancer Type	Gradient of $\rho(\mu t)$
GBM	1.38×10^{-4}
LUAD	4.73×10^{-4}
LUSC	6.98×10^{-4}
BRCA	2.17×10^{-3}
UCEC	2.30×10^{-4}
CRC	2.88×10^{-4}
STAD	3.29×10^{-4}
HNSC	8.37×10^{-4}
SKCM	3.36×10^{-4}
CESC	9.25×10^{-4}
BLCA	3.98×10^{-4}
LGG	2.98×10^{-3}

Data is obtained from [42] with linear regressions performed as in Fig. 2

sion in the presence of T cells. Assumptions on immune response saturation at high genetic instabilities do not affect the outcome of the model. Finally, the death rate of cancer cells increases as they become immunogenic and detectable by T cells [10, 46]. This is translated in the model as cancer cells dying at rate $\delta_c = (\rho_{NK} + \rho_T(\mu))\delta$, the rate of immune detection (ρ) times the rate of T cell killing (δ). Since saturating dynamics are already present in the mathematical shape of ρ_T, this last rate δ is considered constant, which is consistent with other recent modeling efforts [46].

Cancer-Immune system attractor states

Once the proper role of genetic instability on cancer adaptation and immune response is defined, the original model is reinterpreted as a pair of coupled populations with instability-dependent rates, i.e.

$$\frac{dc}{dt} = r(\mu)c\left(1 - \frac{c}{K}\right) - \delta(\rho_{NK} + \rho_T(\mu))cE \tag{5}$$

$$\frac{dE}{dt} = \frac{(\rho_{NK} + \rho_T(\mu))}{g + c}cE + m - \delta_E cE - dE \tag{6}$$

A global picture for the behavior of the system is obtained by studying its possible attractor states taking into account the variability of the mutational load. Together with the cancer free attractor $(c^*, E^*) = (0, m/d)$, other attractors can be inferred from the intersections between nullclines

$$\begin{aligned} E_1(c) &= \frac{r(\mu)}{\delta(\rho_{NK} + \rho_T(\mu))}\left(1 - \frac{c}{K}\right) \\ E_2(c) &= \frac{m}{\left(\delta_E c + d - \frac{(\rho_{NK} + \rho_T(\mu))c}{g+c}\right)} \end{aligned} \tag{7}$$

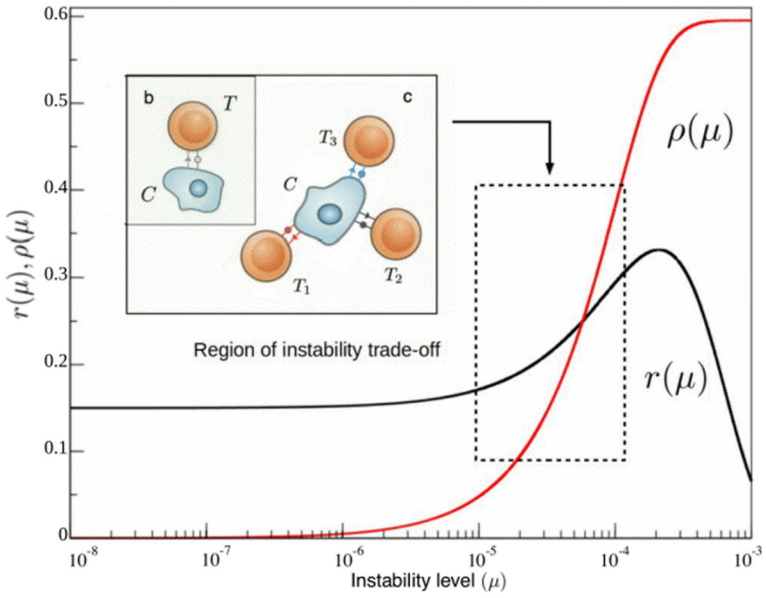

Fig. 3 Functional forms for cancer replication $r(\mu)$ and the adaptive compartment of immune recognition $\rho_T(\mu)$ related to neoantigen presentation. The first (black curve) provides a representation of the cancer instability landscape, as predicted from our theoretical approach (see Methods section) and calibrated by available data. It reveals a very slow increase (in this log-linear diagram) at low instability levels followed by an increase associated to favourable mutations allowing for faster replication and a marked decay at high instabilities due to mutations on viability genes. The immune reactivity to genetic instability function $\rho(\mu)$ (in red, obtained from [42]) rises from zero to saturation beyond $\mu \sim 10^{-4}$. The relevant domain of common cancer instability levels is highlighted. The innate response, ρ_{NK}, is not depicted as is not a function of genetic instability and lies in a smaller order of magnitude of around $\rho_{NK} = 2.5 \times 10^{-}2$ [27]

Nullcline 1 is a simple line with a negative slope controlled by the inverse of the carrying capacity of cancer cells. On the other hand, nullcline 2 is a peaked curve, with a height controlled by immune cell migration and a denominator that might eventually produce divergences. Through their crossings we will find which steady states coexist under which parameter domains (See Results section and Fig. 4).

Along with genetic instability, another parameter is key to the dynamics of the system. Regarding the second nullcline, we can see its size is linearly affected by the influx m of immune cells arriving at the tumor site. It is therefore interesting to understand how μ and m are related to cancer-immune scenarios, since this will open the door to further discussion on mutagenic and immune activation therapies.

By solving $E_1(c) = E_2(c)$, we can understand how the values of m and μ affect the nature and number of possible solutions of the system. We here write $(\rho_{NK} + \rho_T(\mu)) = \rho$ for simplicity. The previous identity leads to a cubic expression of the form $Ac^3 + Bc^2 + Cc + D = 0$, with

$$\begin{aligned} A &= -\frac{\delta_E r(\mu) b}{\delta\rho} \\ B &= -\frac{r(\mu)}{\delta\rho}\Big(b(d-\rho) + \delta_E(bg-1)\Big) \\ C &= \frac{r(\mu)}{\delta\rho}\Big(d + g\delta_E - \rho - bgd\Big) - m \\ D &= \frac{r(\mu)gd}{\delta\rho} - mg. \end{aligned} \tag{8}$$

The sign of the discrimant $\Delta = 18ABCD - 4B^3D + B^2C^2 - 4AC^3 - 27A^2D^2$ will define of which combinations of m and μ belong to which scenarios of Fig. 4. Knowing that three real roots exists for $\Delta > 0$ and only one for $\Delta < 0$, the transitions between attractor scenarios happen to occur at $\Delta = 0$. This condition can be used to easily describe the whole bifurcation space as seen in the results

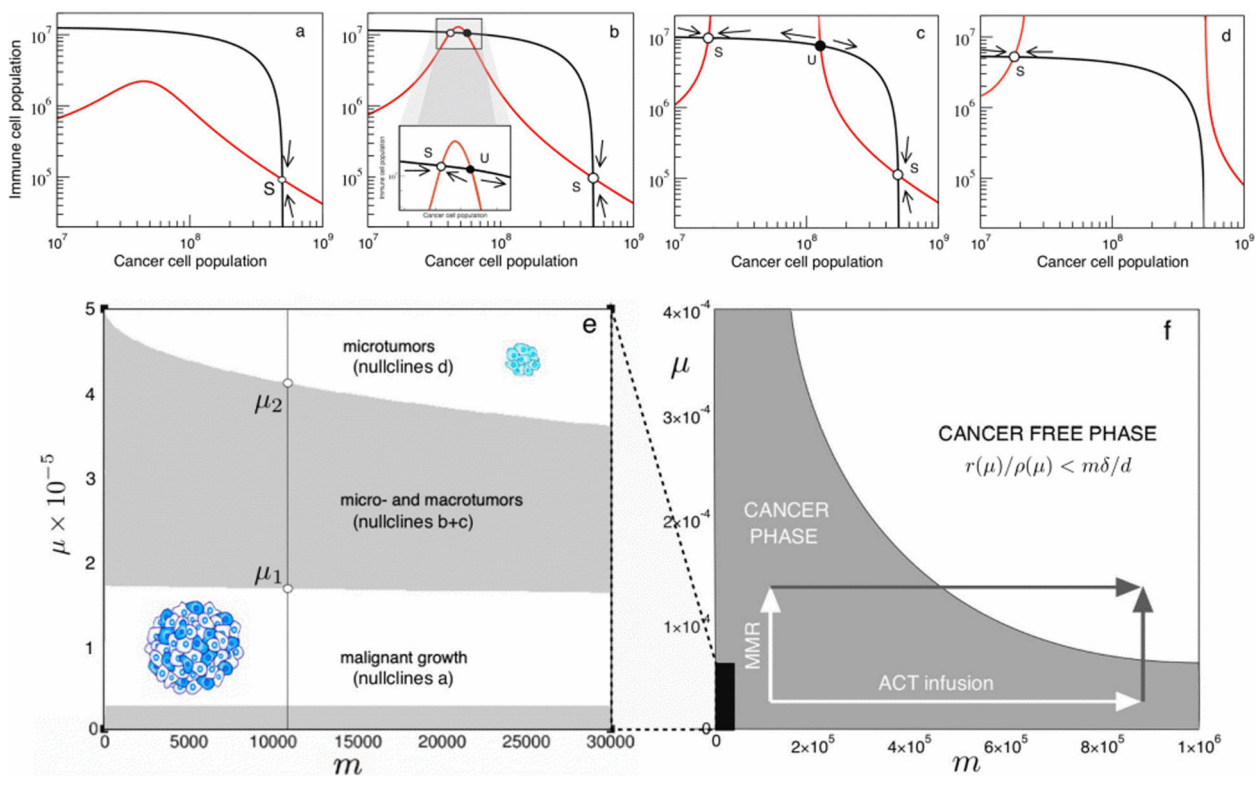

Fig. 4 Cancer-Immune response attractors driven by instability. In (**a**–**d**) we display the nullclines as we increase mutation probability values. Arrows indicate the system flow towards the small and large tumor attractors. Two transitions can be seen. **a** At low genetic instability levels of 10^{-5} mutations per gene per division, such as those common in mutator tumors, only a large cancer attractor coexists with the unstable tumor-free equilibrium left from the graph at $c = 0$. **b** Beyond $\mu^* \sim 1.6 \times 10^{-5}$, two new attractors are created, which correspond to a stable microtumor attractor and an unstable twin [30]. **c** At $\mu^* = 2.0 \times 10^{-5}$, the microtumor attractor becomes smaller; until eventually the attractor of uncontrolled tumor growth is eliminated (**d**) at mutational levels similar to those attained after Mismatch-Repair knockout [40]. In (**e**) and (**f**) we summarise the bifurcation diagrams for the possible scenarios as a function of μ and m. For standard immue migration rates (**e**, black region in **f**), mutational increases drive the system across the two transitions observed in (**a**–**d**) and towards the controlled tumor state. However, by increasing both μ and m through combining Mismatch Repair knockout with adoptive cell therapy, the total cancer clearance state can be accessed

and Fig. 4e and f, showing how mutation frequencies and immune stimulation affect the possible outcomes of the system.

Results

Minimal mutation rate for an efficient immune response

Before engaging into a full analysis of the complete model, we can study the behavior of the system for initial phases of progression. This corresponds to a small tumor of size $c << K = 2 \times 10^9$ cells. Under this assumption, the population dynamics of $c(t)$ simplifies to

$$\frac{dc}{dt} = c\Big(r(\mu) - \delta(\rho_{NK} + \rho_T(\mu))E - d_c\Big) \tag{9}$$

where we have now included a natural death rate $-d_c$ that accounts for growth barriers of initial malignant cells if away from the microenviroment carrying capacity [33]. From (9) we can isolate a condition for tumor control, i.e.:

$$\frac{dc}{dt} < 0 \tag{10}$$

which leads to a crude estimate of the amount of effector immune cells required to counterbalance tumor growth, namely

$$E(\mu) > \frac{r(\mu) - d_c}{\delta_C(\rho_{NK} + \rho_T(\mu))}. \tag{11}$$

The inequality consistently shows that $E(\mu)$ will be proportional to the instability landscape of cancer growth rate divided by both NK and immune-mediated death. This acknowledges that both NK or T cells can play crucial roles in cancer surveillance. To understand the role of the adaptive compartment and genetic instability in controling a growing cancer population, we use validated data from [20] (Table 1) and consider a healthy adaptive immune population of $T \sim 10^7$ cells ([29] and following sections), to obtain that the immune control condition is fulfilled for $\mu > 5.75 \times 10^{-5}$ mutations per gene and replication. This can be understood as the minimal mutation rate required to generate a critical neoantigen load for T-cell immune attack, not considering here NK or other innate components away from the scope of the work. The estimated value is consistend within the range of genetic instability levels associated to MMR knockout [47], indicating a connection between mutagenic therapies enhancing genetic instability and a threshold level to activate the immune response.

Transitions to tumor control and eradication at genetic instabilities within the mMR-knockout range

For well-formed tumors, no similar approach can be performed, but we can study the effects of changes in genetic instability in the sytem defined by equations (4) and (5) by picturing the intersections between nullclines described in the Methods section. As we are interested in the specific role of genetic instability and neoantigen presentation, we will focus here on the adaptive part of immune recognition, $\rho(\mu)$. It is straightforward to see how several transitions regarding creation and anihilation of steady states are governed by mutational probability μ (Fig. 4a-d).

As expected from [30] and previous discussions, we know that the cancer-free attractor will always be present, but local stability will be ensured if $r(\mu)/(\rho_{NK} + \rho_T(\mu)) < m\delta/d$ (depicted in Fig. 4f). Without an innate component, the condition is only fulfilled at very high instability levels above 10^{-4} mutations per gene per division. This implies that no complete tumor clearance solely by neoantigen recognition seems possible at realistic mutation rates for fixed m, meaning that an innate response might also play a role in complete respondant patients, as many therapies do elicit total tumor eradication [45]. Additionally, we can see that a large-tumor solution c_L is also present at low instabilities (Fig. 4a), and it is globally asymptotically stable. Interestingly, a transition seems to occur as the value for μ becomes larger: before $E_2(c)$ diverges, a smaller stable attractor c_S is created together with its unstable *twin* (Fig. 4b), which is often described as a microtumor controlled by the immune system. Furthermore, nullcline 2 diverges at $\mu \sim 1.75 \times 10^{-5}$ (Fig. 4c), and, as the two values for divergence of $E_2(c)$ grow further appart, the large cancer attractor disappears and only the controlled microtumor coexists with the cancer free attractor and is globally asymptotically stable (Fig. 4d). These results are consistent with those of [30], where such solution is considered a microtumor controlled by the immune system. However, both transitions of microtumor creation and large tumor elimination being a function of the mutational levels of the tumor population are new to the present work.

At this point it is clear that understanding at what instability levels these transitions happen is key to the possible outcomes of the tumor-immune interaction. For the given parameter region and in the absence of a strong innate response, a basic computational approach lets us see that the first transition happens around $\mu \sim 1.65 \times 10^{-5}$ (Fig. 4b), whereas another transition where the large tumor attractor disappears happens at higher μ values of about $\mu \sim 4 \times 10^{-5}$ (Fig. 4d).

Following extensive data, unleashed genetic instability after Mlh1 knockout in mice accounts for increasing mutation frequencies ranging from $10^{-6} \sim 10^{-5}$ up to 10^{-4} mutations per gene per division (values assessed for transgenic mice containing *supFG1* or *cII* from [47]). Interestingly, instability levels before MMR knockout put our system within a region where the large cancer attractor is stable and no controlled microtumor exists. However, the increase after Mlh1 knockout might be pushing cancer cells to a region beyond μ_1^*, where the stable

microtumor attractor appears, or even μ_2^*, where the stable large cancer attractor has disappeared (Fig. 4e).

The resemblance between the model and experiments linking genetic instability to adaptive immune surveillance seems intuitive enough. Following [17], we think that there is a connection between the observed phenomenon of immune reactivity and tumor collapse after Mismatch Repair knockout and the qualitative behavior of our model, which depicts a transition of this kind at high μ values. Furthermore, we have taken advantage of recent research in order to use quantitative data to build our model. The fact that our model predicts the range for which immune surveillance reacts at increased cancer instability levels emphasizes the possible existence of transitions like the ones studied here.

Assessing if these two transitions are in fact well defined in vitro or if genetic instability can modulate tumor evolution towards controlled states can shed new light into the precise nature of mutagenic therapy as a mechanism towards increasing tumor immunogeneicity. Such therapies have produced key results in the field of virology [48], but, within the context of cancer, recent insight seems to indicate that increasing the immunogeneicity of a tumor preludes evolution of subclonal neoantigen heterogeneity [49–51].

Implications on immune surveillance: the role of tumor size

Besides the possible implications for mutagenic therapy as a facilitator of immunotherapy effectiveness, the fact that genetic instability shapes the landscape of the cancer-immune interaction has further implications on the fate of tumor growth. Tumor size has been shown to be associated with response to immunotherapies [52], but several scenarios, from surveillance to evasion, are known to occur [31, 53, 54]. Is genetic instability related to the polymorphic nature of immunotherapy prognosis?

From Fig. 4a we know that, in conditions of low genetic instability, the large tumor equilibrium is globally asymptotically stable (GAS), and insufficient presentation of antigens implies that even small tumors can evade immune surveillance in the absence of a strong innate response through NK cells or macrophages. This could be the case of both initial microsatellite-stable malignancies or clones that have evolved low antigenicity through genome editing [45].

Increases in genetic instability result in a phase transition that creates a micro-tumor attractor (Fig. 4b-c). This state has been previously related to dormancy, where the adaptive immune system is able to control cancer growth [31]. However, the large cancer attractor is still present, and local asymptotic stability ensures that tumor sizes within its basin of attraction will stil grow towards it. The implications are revelant to therapy: small tumors of medium antigenicity can be controled, but large tumors will still grow towards larger disease. This result is consistent with the notion that therapy reducing tumor mass is often effective prior to immunotherapy [20, 55].

The second transition, consistent with experiments of immune surveillance after Mismatch-Repair Knockout [17], indicates the disappearance of the large cancer attractor (Fig. 4d). This implies that highly immunogenic tumors will always elicit a sufficiently effective immune response that will drive them towards microtumor control [31], no matter their initial size. However, the fact that there is no complete remission implies that evolutionary pressures still act on the remaining rogue population, and the small clone can eventually evolve immune evasion [45].

Mutagenic therapy remains a relevant actor on the cancer-immune ecology. However, without the cooperative effects of an innate response, through the constant recognition rate ρ_{NK}, or the buffering of immune migration m, the cancer-free equilibrium is only stable at very high genetic instability levels that do not seem attainable through mutagenic agents. What are the cooperative dynamics of genetic instability with these immune agents?

Effects of modulating immune migration and the innate response

Beyond the relevance of genetic instability as a driver of tumor antigenicity, the fact that the cancer free attractor becomes stable at very high mutational levels above 10^{-4} mutations per gene and division (at least for the data on adaptive immunity from [20]) implies that further considerations on therapy need to be taken into account. The overall condition for total disease eradication is

$$\frac{r(\mu)}{\rho_{NK} + \rho_T(\mu)} < \frac{m\delta}{d}. \tag{12}$$

If genetic instability alone does not suffice to fulfill this condition, what other therapeutic schemes are of relevance to our model? A first notion lies on understanding how does μ alter the minimal innate recognition ρ_{NK} necessary for complete disease remission, as defined by the condition

$$\rho_{NK}^* > \frac{r(\mu)d}{m\delta} - \rho(\mu) \tag{13}$$

For microsatellite stable tumors with $\mu << 10^{-5}$, the necessary recruitment rate of NK cells is within the 10^{-1}day-1 range, an order of magnitude larger than that measured in [27]. However, increasing genetic instability decreases ρ_{NK}^* in a quasi-linear way, so that after a possible MMR knockout, a recruitment rate within 10^{-2}day^{-1} would suffice for cancer clearance, indicating the possibility of a combination therapy enhancing both mutagenesis and NK cell activation [28].

Together with the role of innate immunity, another key observation is considering the rate of immune migration (m) as a measure of immune activation. The necessary flow of immune cells to the tumor to achieve complete remission is

$$m^* > \frac{r(\mu)d}{\delta(\rho_{NK} + \rho_T(\mu))} \tag{14}$$

Interestingly enough, the migration rate necessary for cancer clearance does not decay linearly with genome instability, as for ρ_{NK}^*, but in an exponential way, meaning that increases in genetic instability within the MMR knockout range rapidly decrease the condition for immune migration rate (Fig. 4f). This indicates a strong synergy between mutagenesis and immune activation therapies such as Adoptive Cell Therapy (ACT) [56], consistent with recent discussion on combination therapies [7, 19].

Moreover, by picturing the bifurcation diagram in standard μ and m regions as described in the Methods section (Fig. 4e), it is interesting to see how the first transition towards microtumor creation, μ_1^*, has a weak dependency on m, since the appearance of the intermediate attractors depends mostly on the denominator of nullcline 2 becoming null, so that $E_2(c)$ diverges at

$$\delta_E c + d - ((\rho_{NK} + \rho_T(\mu))c/(g + c)) = 0, \tag{15}$$

which is not a function of m. On the other hand, the transition to disappearance of the large-cancer attractor does depend on m, since m affects the width of $E_2(c)$, so that for higher m values $E_2(c)$ will go faster towards infinity and not cross $E_1(c)$. However, it seems intuitive from Fig. 4f that the role of genetic instability in increasing neoantigen production might be crucial even in the presence of high immune activation.

Mathematical work previous to our instability-driven model developed interesting considerations on derivation of cancer vaccines (see e.g. [57]), and introduced time dependent treatments [58] or time-delays in the immune response [59] based on the immune migration parameter, despite mathematical considerations remained somehow distant from clinical immunology and not many of the described behaviors after mathematically designed therapies have been observed in vivo [22].

Recent research has highlighted the importance of genetic instability as a marker for good prognosis in immune checkpoint inhibition therapies [14–16]. Its role in neoantigen production is acknowledged as crucial [10]. Our results describing μ as another driver towards surveillance complementing m and ρ_{NK} reinforce the relevance of genetic instability in the tumor-immune dynamics, further supporting the possibility of increasing tumor immunogeneicity by promoting T cell antigen presentation [7, 9].

Discussion

In the present work we have studied a minimal mathematical scenario describing how genetic instability, by means of enhancing tumor adaptation along with neoantigen production and immune recognition, can trigger sharp transitions towards tumor control and eradication.

Starting from basic considerations, we have asked ourselves about the ecological interactions between malignant cells and, in particular, effector immune cells able to respond after neoantigen recognition. Specifically, we consider how genetic instability, here as a mutation probability, shapes tumor adaptability and immune response.

Interestingly, genetic instability governs the possible outcomes of the system. Increasing mutational levels drive the system across two phase transitions. In the first one, two attractors are created involving smaller tumors coexisting with a larger population of T cells. This state has been characterized as a controlled, but not totally eliminated microtumor [30, 31]. The second transition accounts for the disappearence of the cancer-wins scenario, so that only solutions of immune control are present at large genetic instability levels.

Recent advances in the field of cancer immunology have proven that genetic instability is a key ingredient of the immune response [14–16], and particular research claims immune surveillance after MMR knockout follows from this causal relation between high mutational loads and neoepitope production [17]. In the context of this research, our model provides a conceptual and numerical description for how a transition between cancer growth and arrest can follow only from damaging DNA repair mechanisms. More generally, the fact that microsatellite instability levels govern transitions separating cancer growth from immune surveillance might be indicative of why highly unstable tumors are better respondants to immunotherapy [10]. Furthermore, we have used available data to calibrate the model parameters and to construct the immune recognition function. Using this information, we consistently explain phase transitions happening at microsatellite instability levels that resemble those of MMR knockout. However, even if these transitions could exist in the laboratory, we have discussed further aspects that need to be accounted when dealing with increasing tumor immunogeneicity through mutagenesis [49, 50].

We have also studied the roles of ρ_{NK}, the recruitment of NK cells, and m, a parameter refering to immune migration or an eventual immune therapy. The model indicates a cooperative effect between therapies affecting mutagenesis together with NK or migration buffering. The strength of this cooperative effect is linear for genetic instability and innate immune cell

recruitment, but the model also predicts that, when an innate response and T cell recognition alone cannot control tumor growth, cross-therapies modulating both m and μ might be exponentially effective in driving the tumor-immune interaction into a state of total disease eradication, thus indicating a mathematical validation for recent insight into combined immunotherapies [7]. We further suggest that the relevance of m in producing transitions to tumor arrest is low, while minor increases in genetic instability seem much more effective against large tumors. This indicates that cross therapies inducing DNA damage prior to immunotherapy might drive tumors to neoantigen-rich states [18, 19] before immune editing processes enter at play [45, 60]. We therefore postulate a possible mathematical description of recent discussions for novel perspectives on combination immunotherapy [7].

All the previous conclusions stem from a very minimal mathematical model, whereas the immune system is known to be complex [45, 61] Additionally, other interactions between immunotherapies and conventional therapies need to be taken into account [19]. In particular, several cooperative mechanisms between immune populations might play a role in non-antigenic T cell activation [27]. Further research should consider the possible nonlinear dynamics stemming from T cell sensitization after cancer-NK cellular interactions.

Finally, as a result from the lack of heterogeneity, our model does not yet capture immune editing, a phenomenom at the core of immunotherapy failure, in which the tumor might develope immune resistance by means of either buffering the growth of immunosilent cells or editing its genome to express fewer neoantigens [60]. Within this view, current research claims that tumor mutational burden might not be a sufficient biomarker [46, 50]. In the presence of an effective immune response, antigenic subclones can be negatively selected, giving rise to immuno-silent tumors despite its possibly high mutational load. Together with immune editing, recent studies highlight heterogeneity itself as a source for failure of the immune response [49, 51] as it directly affects the spatial and clonal distribution of neoantigens. Further modeling of the tumor-immune ecology could benefit from considering heterogeneous populations where antigen frecuencies are taken into account. Despite these considerations, our results on the cooperative roles of m and μ indicate that damage on DNA repair mechanisms prior to checkpoint blockade could render tumors immunogenic before a reactivated immune system pressures towards editing. Using an evolutionary framework such as adaptive dynamics [37], future work might help to characterize in which regimes do cancer subclones evade immune surveillance through evolving their neoantigen landscape [62].

Conclusions

This work provides a first effort towards modeling the double-edged effect of genetic instability in both cancer adaptation and immune surveillance with the goal of understanding the specific role of mutational load as a driver of immune attack. Two main results stem from the model. First, transitions towards tumor control follow from increases in mutational levels similar to those after MMR knockout. Second, genetic instability and immune activation have a cooperative effect in driving tumor elimination, indicating that combination therapies enhancing both might be key in the future.

Acknowledgements
The authors thank Elisa Beltran and Jordi Piñero for fruitful discussions and Blai Vidiella for detailed revisions, as well as the rest of Complex Systems Lab members and Pierre Menard for their inspiring ideas. G.A. thanks Noemi Andor and the team at the *Evolutionary Biology and Ecology of Cancer Summer School 2018* for interesting comments on the cancer-immune system interaction.

Authors' contributions
GA and RS developed the mathematical model, wrote the manuscript and created the figures. Both authors read and approved the final manuscript.

Funding
This work has been supported by the Botín-Foundation by Banco Santander through its Santander Universities Global Division, a MINECO grant FIS2015-67616-P (MINECO/FEDER, UE) fellowship, an AGAUR grant 2018 by the Universities and Research Secretariat of the Ministry of Business and Knowledge of the Generalitat de Catalunya and the European Social Fund and by the Santa Fe Institute.

Availability of data and materials
Not applicable

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Author details
[1]ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain. [2]Institut de Biologia Evolutiva (CSIC-UPF), Psg Maritim Barceloneta, 37, 08003 Barcelona, Spain. [3]Santa Fe Institute, 399 Hyde Park Road, 87501 Santa Fe, NM, USA.

Received: 12 June 2019 Accepted: 30 October 2019
Published online: 11 December 2019

References

1. Greaves M, Maley C. Clonal evolution in cancer. Nature. 2012;481:306–13.
2. Lengauer C, Kinzler K, Vogelstein B. Genetic instabilities in human cancers. Nature. 1998;396:643–9.
3. Negrini S, Gorgoulis V, Halazonetis T. Genomic instability – an evolving hallmark of cancer. Nat Rev Mol Cell Biol. 2010;11:220–228.
4. Andor N, Maley C, Ji H. Genomic instability in cancer: teetering on the limit of tolerance. Cancer Res. 2017;77:7–11.
5. Sottoriva A, Spiteri I, Piccirillo SGM, Touloumis A, Collins VP, Marioni JC, Curtis C, Watts C, Tavaré S. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proc Natl Acad Sci U S A. 2013;110:4009–14.

6. Miller J, Sadelain M. The journey from discoveries in fundamental immunology to cancer immunotherapy. Cancer Cell. 2015;27:439–49.
7. Zappasodi R, Merghoub T, Wolchok J. Emerging concepts for immune checkpoint blockade-based combination therapies. Cancer Cell. 2018;33: 581–98.
8. Rotte A. Combination of ctla-4 and pd-1 blockers for treatment of cancer. J Exp Clin Cancer Res. 2019;38(1):255. https://doi.org/10.1186/s13046-019-1259-z.
9. Looi C-K, Chung FF-L, Leong C-O, Wong S-F, Rosli R, Mai C-W. Therapeutic challenges and current immunomodulatory strategies in targeting the immunosuppressive pancreatic tumor microenvironment. J Exp Clin Cancer Res. 2019;38(1):162. https://doi.org/10.1186/s13046-019-1153-8.
10. Schumacher T, Schreiber R. Neoantigens in cancer immunotherapy. Science. 2015;348:69–74.
11. Efremova M, Finotello F, Rieder D, Trajanoski Z. Neoantigens generated by individual mutations and their role in cancer immunity and immunotherapy. Front Immunol. 2017;8:1679.
12. Vesely M, Schreiber R. Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. Ann NY Acad Sci. 2013;128:1–5.
13. Segal NH, Parsons DW, Peggs KS, Velculescu V, Kinzler KW, Vogelstein B, Allison JP. Epitope landscape in breast and colorectal cancer. Cancer Res. 2008;68(3):889–92.
14. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, Miller ML, Rekhtman N, Moreira AL, Ibrahim F, Bruggeman C, Gasmi B, Zappasodi R, Maeda Y, Sander C, Garon EB, Merghoub T, Wolchok JD, Schumacher TN, Chan TA. Mutational landscape determines sensitivity to pd-1 blockade in non–small cell lung cancer. Science. 2015;348(6230):124–8.
15. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, Walsh LA, Postow MA, Wong P, Ho TS, Hollmann TJ, Bruggeman C, Kannan K, Li Y, Elipenahli C, Liu C, Harbison CT, Wang L, Ribas A, Wolchok JD, Chan TA. Genetic basis for clinical response to ctla-4 blockade in melanoma. N Engl J Med. 2014;371(23):2189–99.
16. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Geukes Foppen MH, Goldinger SM, Utikal J, Hassel JC, Weide B, Kaehler KC, Loquai C, Mohr P, Gutzmer R, Dummer R, Gabriel S, Wu CJ, Schadendorf D, Garraway LA. Genomic correlates of response to ctla-4 blockade in metastatic melanoma. Science. 2015;350(6257):207–11.
17. Germano G, Lamba S, Rospo G, Barault L, Magrì A, Maione F, Russo M, Crisafulli G, Bartolini A, Lerda G, Siravegna G, Mussolin B, Frapolli R, Montone M, Morano F, de Braud F, Amirouchene-Angelozzi N, Marsoni S, D'Incalci M, Orlandi A, Giraudo E, Sartore-Bianchi A, Siena S, Pietrantonio F, Di Nicolantonio F, Bardelli A. Inactivation of dna repair triggers neoantigen generation and impairs tumour growth. Nature. 2017;552:409–13.
18. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, Wong F, Azad NS, Rucki AA, Laheru D, Donehower R, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Greten TF, Duffy AG, Ciombor KK, Eyring AD, Lam BH, Joe A, Kang SP, Holdhoff M, Danilova L, Cope L, Meyer C, Zhou S, Goldberg RM, Armstrong DK, Bever KM, Fader AN, Taube J, Housseau F, Spetzler D, Xiao N, Pardoll DM, Papadopoulos N, Kinzler KW, Eshleman JR, Vogelstein B, Anders RA, Diaz LA. Mismatch repair deficiency predicts response of solid tumors to pd-1 blockade. Science. 2017;357(6349):409–13.
19. Wang Y, Liu Z-G, Yuan H, Deng W, Li J, Huang Y, Kim BYS, Story MD, Jiang W. The reciprocity between radiotherapy and cancer immunotherapy. Clin Cancer Res. 2019;25(6):1709–17. https://doi.org/10.1158/1078-0432.CCR-18-2581. https://clincancerres.aacrjournals.org/content/25/6/1709.full.pdf.
20. Kuznetsov Y, Makalkin I, Taylor M, Perelson A. Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. Bull Math Biol. 1994;2:295–321.
21. Łuksza M, Riaz N, Makarov V, Balachandran V, Hellmann M, Solovyov A, Rizvi N, Merghoub T, Levine A, Chan T, Wolchok J, Greenbaum B. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. Nature. 2017;551:517–20.
22. Eftimie R, Bramson J, DJD E. Interactions between the immune system and cancer: A brief review of non-spatial mathematical models. Bull Math Biol. 2011;73:2–32.
23. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. Cell Rep. 2017;18:248–62.
24. Hackl H, Charoentong P, Finotello F, Trajanoski Z. Computational genomics tools for dissecting tumour-immune cell interactions. Nat Rev Genet. 2013;17:441–58.
25. Bindea G, Mlecnik B, Tosolini M, Kirilovsky A, Waldner M, Obenauf A, Angell H, Fredriksen T, Lafontaine L, Berger A, Bruneval P, Fridman W, Becker C, Pagès F, Speicher M, Trajanoski Z, Galon J. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. Immunity. 2013;39(4):782–95.
26. Ehrlich P. Collected Papers in Four Volumes Including a Complete Bibliography. London: Pergamon Press; Unknown Month 1956.
27. de Pillis LG, Radunskaya AE, Wiseman CL. A validated mathematical model of cell-mediated immune response to tumor growth. Cancer Res. 2005;65(17):7950–8. https://doi.org/10.1158/0008-5472.CAN-05-0564. https://cancerres.aacrjournals.org/content/65/17/7950.full.pdf.
28. Rezvani K, Rouce RH. The application of natural killer cell immunotherapy for the treatment of cancer. Front Immunol. 2015;6:578. https://doi.org/10.3389/fimmu.2015.00578.
29. Perelson A, Weisbuch G. Immunology for physicists. Rev Mod Phys. 1997;69:1219.
30. d'Onofrio A. A general framework for modeling tumor-immune system competition and immunotherapy: mathematical analysis and biomedical inferences. Physica D. 2005;208:220–35.
31. Koebel CM, Vermi W, Swann JB, Zerafa N, Rodig SJ, Old LJ, Smyth MJ, Schreiber RD. Adaptive immunity maintains occult cancer in an equilibrium state. Nature. 2007;450:903.
32. Matzavinos A, Chaplain M, Kuznetsov V. Mathematical modelling of the spatio-temporal response of cytotoxic t-lymphocytes to a solid tumour. Math Med Biol. 2004;21:1–34.
33. Gatenby RA, Gillies RJ. A microenvironmental model of carcinogenesis. Nat Rev Cancer. 2008;8(1):56–61. https://doi.org/10.1038/nrc2255.
34. Kirschner D, Panetta J. Modeling immunotherapy of the tumor-immune interaction. J Math Biol. 1998;37:235–52.
35. Solé R, Deisboeck T. An error catastrophe in cancer? J Theor Biol. 2004;228:47–54.
36. Solé RV, Valverde S, Rodriguez-Caso C, Sardanyés J. Can a minimal replicating construct be identified as the embodiment of cancer?. BioEssays. 2014;36(5):503–12. https://doi.org/10.1002/bies.201300098. https://onlinelibrary.wiley.com/doi/pdf/10.1002/bies.201300098.
37. Aguadé-Gorgorió G, Solé R. Adaptive dynamics of unstable cancer populations: The canonical equation. Evol Appl. 2018;11:1283–92.
38. Eyre-Walker A, Keightley P. The distribution of fitness effects of new mutations. Nat Rev Gen. 2007;8:610–8.
39. Eisenberg E, Levanon E. Human housekeeping genes, revisited. Trends Genet. 2013;29:569–74.
40. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science. 2013;339(6127):1546–58. https://doi.org/10.1126/science.1235122. https://science.sciencemag.org/content/339/6127/1546.full.pdf.
41. Tomlinson IPM, Novelli MR, Bodmer WF. The mutation rate and cancer. Proc Natl Acad Sci. 1996;93(25):14800–3. https://doi.org/10.1073/pnas.93.25.14800. https://www.pnas.org/content/93/25/14800.full.pdf.
42. Rooney M, Shukla S, Wu C, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell. 2015;160:48–61.
43. Alexandrov L, Nik-Zainal S, Wedge D, Aparicio S, Behjati S, Biankin A, Bignell G, Bolli N, Borg A, Børresen-Dale A, Boyault S, Burkhardt B, Butler A, Caldas C, Davies H, Desmedt C, Eils R, Eyfjörd J, Foekens J, Greaves M, Hosoda F, Hutter B, Ilicic T, Imbeaud S, Imielinski M, Jäger N, Jones D, Jones D, Knappskog S, Kool M, Lakhani SR, López-Otín C, Martin S, Munshi NC, Nakamura H, Northcott PA, Pajic M, Papaemmanuil E, Paradiso A, Pearson JV, Puente XS, Raine K, Ramakrishna M, Richardson AL, Richter J, Rosenstiel P, Schlesner M, Schumacher TN, Span PN, Teague JW, Totoki Y, Tutt ANJ, Valdés-Mas R, van Buuren MM, van 't Veer L, Vincent-Salomon A, Waddell N, Yates LR, Initiative APCG, Consortium IBC, Consortium IM-S, PedBrain ICGC, Zucman-Rossi J, Andrew Futreal P, McDermott U, Lichter P, Meyerson M, Grimmond SM, Siebert R, Campo

E, Shibata T, Pfister SM, Campbell PJ, Stratton MR. Signatures of mutational processes in human cancer. Nature. 2013;500:415–21.
44. Sottoriva A, Kang H, Ma Z, Graham T, Salomon M, Zhao J, Marjoram P, Siegmund K, Press M, Shibata D, Curtis C. A big bang model of human colorectal tumor growth. Nat Genet. 2015;47:209–16.
45. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and acquired resistance to cancer immunotherapy. Cell. 2017;168(4):707–23.
46. Lakatos E, Williams MJ, Schenck RO, Cross WCH, Househam J, Werner B, Gatenbee C, Robertson-Tessi M, Barnes CP, Anderson ARA, Sottoriva A, Graham TA. Evolutionary dynamics of neoantigens in growing tumours. bioRxiv. 2019. https://doi.org/10.1101/536433. https://www.biorxiv.org/content/early/2019/01/31/536433.full.pdf.
47. Hegan DC, Narayanan L, Jirik FR, Edelmann W, Liskay RM, Glazer PM. Differing patterns of genetic instability in mice deficient in the mismatch repair genes Pms2, Mlh1, Msh2, Msh3 and Msh6. Carcinogenesis. 2006;27(12):2402–8. https://doi.org/10.1093/carcin/bgl079. http://oup.prod.sis.lan/carcin/article-pdf/27/12/2402/7513214/bgl079.pdf.
48. Loeb LA, Essigmann JM, Kazazi F, Zhang J, Rose KD, Mullins JI. Lethal mutagenesis of hiv with mutagenic nucleoside analogs. Proc Natl Acad Sci. 1999;96(4):1492–7. https://doi.org/10.1073/pnas.96.4.1492. https://www.pnas.org/content/96/4/1492.full.pdf.
49. McGranahan N, Furness AJS, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, Watkins TBK, Shafi S, Murugaesu N, Mitter R, Akarca AU, Linares J, Marafioti T, Henry JY, Van Allen EM, Miao D, Schilling B, Schadendorf D, Garraway LA, Makarov V, Rizvi NA, Snyder A, Hellmann MD, Merghoub T, Wolchok JD, Shukla SA, Wu CJ, Peggs KS, Chan TA, Hadrup SR, Quezada SA, Swanton C. Clonal neoantigens elicit t cell immunoreactivity and sensitivity to immune checkpoint blockade. Science. 2016;351(6280):1463–9. https://doi.org/10.1126/science.aaf1490. https://science.sciencemag.org/content/351/6280/1463.full.pdf.
50. Helleday T. Making immunotherapy 'cold' tumours 'hot' by chemotherapy-induced mutations—a misconception. Ann Oncol. 2019;30(3):360–1. https://doi.org/10.1093/annonc/mdz013. http://oup.prod.sis.lan/annonc/article-pdf/30/3/360/28255092/mdz013.pdf.
51. Wolf Y, Bartok O, Patkar S, Eli GB, Cohen S, Litchfield K, Levy R, Jiménez-Sánchez A, Trabish S, Lee JS, Karathia H, Barnea E, Day C.-P., Cinnamon E, Stein I, Solomon A, Bitton L, Pérez-Guijarro E, Dubovik T, Shen-Orr SS, Miller ML, Merlino G, Levin Y, Pikarsky E, Eisenbach L, Admon A, Swanton C, Ruppin E, Samuels Y. Uvb-induced tumor heterogeneity diminishes immune response in melanoma. Cell. 2019;179(1):219–23521. https://doi.org/10.1016/j.cell.2019.08.032.
52. Hanson HL, Donermeyer DL, Ikeda H, White JM, Shankaran V, Old LJ, Shiku H, Schreiber RD, Allen PM. Eradication of established tumors by cd8+ t cell adoptive immunotherapy. Immunity. 2000;13(2):265–76. https://doi.org/10.1016/S1074-7613(00)00026-1.
53. Diefenbach A, Jensen ER, Jamieson AM, Raulet DH. Rae1 and h60 ligands of the nkg2d receptor stimulate tumour immunity. Nature. 2001;413(6852):165–71. https://doi.org/10.1038/35093109.
54. Kuznetsov VA, Knott GD. Modeling tumor regrowth and immunotherapy. Math Comput Model. 2001;33(12):1275–87. https://doi.org/10.1016/S0895-7177(00)00314-9.
55. Pillis LGD, Radunskaya A. A mathematical tumor model with immune resistance and drug therapy: an optimal control approach. J Theor Med. 2001;3(2):79–100. https://doi.org/10.1080/10273660108833067. https://www.tandfonline.com/doi/pdf/10.1080/10273660108833067.
56. Rosenberg SA, Restifo NP. Adoptive cell transfer as personalized immunotherapy for human cancer. Science. 2015;348(6230):62–8.
57. Gajewski TF. Failure at the effector phase: Immune barriers at the level of the melanoma tumor microenvironment. Clin Cancer Res. 2007;13(18):5256–61. https://doi.org/10.1158/1078-0432.CCR-07-0892. http://clincancerres.aacrjournals.org/content/13/18/5256.full.pdf.
58. Sotolongo-Costa O, Molina LM, Perez DR, Antoranz J, Reyes MC. Behavior of tumors under nonstationary therapy. Phys D Nonlinear Phenom. 2003;178(3–4):242–53.
59. Gałach M. Dynamics of the tumor—immune system competition—the effect of time delay. Int J Appl Math Comput Sci. 2003;13:395–406.
60. De Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. Nat Rev Cancer. 2006;6(1):24.
61. Chen DS, Mellman I. Elements of cancer immunity and the cancer–immune set point. Nature. 2017;541(7637):321.
62. Anagnostou V, Smith KN, Forde PM, Niknafs N, Bhattacharya R, White J, Zhang T, Adleff V, Phallen J, Wali N, et al. Evolution of neoantigen landscape during immune checkpoint blockade in non–small cell lung cancer. Cancer Discov. 2017;7(3):264–76.

Publisher's Note

INTERFACE

royalsocietypublishing.org/journal/rsif

Research

Cite this article: Aguadé-Gorgorió G, Solé R. 2020 Tumour neoantigen heterogeneity thresholds provide a time window for combination immunotherapy. *J. R. Soc. Interface* **17**: 20200736.
http://dx.doi.org/10.1098/rsif.2020.0736

Received: 10 September 2020
Accepted: 7 October 2020

Subject Category:
Life Sciences–Mathematics interface

Subject Areas:
biocomplexity, biomathematics, evolution

Keywords:
antigen diversity, immune escape, checkpoint blockade immunotherapy, neutral evolution

Authors for correspondence:
Guim Aguadé-Gorgorió
e-mail: guim.aguade@upf.edu
Ricard Solé
e-mail: ricard.sole@upf.edu

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5179186.

THE ROYAL SOCIETY PUBLISHING

Tumour neoantigen heterogeneity thresholds provide a time window for combination immunotherapy

Guim Aguadé-Gorgorió[1,2] and Ricard Solé[1,2,3]

[1]ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain
[2]Institut de Biologia Evolutiva (CSIC-UPF), Psg Maritim Barceloneta, 37, 08003 Barcelona, Spain
[3]Santa Fe Institute, 399 Hyde Park Road, Santa Fe NM 87501, USA

GA-G, 0000-0001-9467-3027; RS, 0000-0001-6974-1008

Following the advent of cancer immunotherapy, increasing insight has been gained on the role of mutational load and neoantigens as key ingredients in T cell recognition of malignancies. However, not all highly mutational tumours react to immune therapies, and initial success is often followed by eventual relapse. Heterogeneity in the neoantigen landscape of a tumour might be key in the failure of immune surveillance. In this work, we present a mathematical framework to describe how neoantigen distributions shape the immune response. The model predicts the existence of an antigen diversity threshold level beyond which T cells fail at controlling heterogeneous tumours. Incorporating this diversity marker adds predictive value to antigen load for two cohorts of anti-CTLA-4 treated melanoma patients. Furthermore, our analytical approach indicates rapid increases in epitope heterogeneity in early malignancy growth following immune escape. We propose a combination therapy scheme that takes advantage of preexisting resistance to a targeted agent. The model indicates that the selective sweep for a resistant subclone reduces neoantigen heterogeneity, and we postulate the existence of a time window before tumour relapse where checkpoint blockade immunotherapy can become more effective.

1. Introduction

The mechanisms that make cancer a major cause of human death are deeply rooted in the principles of evolution [1]. Through depletion of early mutations affecting mostly multicellular regulation and tissue homeostasis, cancer cells can overcome multiple selection barriers, making the human genome a pool for the evolution of a myriad possible phenotypes [2]. Highly diverse rogue populations within a tumour include cells that resist or evolve resistance to a vast array of therapy schemes [3,4].

Among other selection barriers, cancer cells often evolve the ability to evade immune surveillance [5] required to prevent the outgrowth of transformed cells [6] (figure 1*a*). Recent decades have seen a rapidly growing insight into the molecular details of immune silencing and tolerance [7], leading to the advent of checkpoint inhibition therapies able to target specific immune-evasion mechanisms in malignancies [8]. Once the immune system is back in place, the role of neoantigens, i.e. mutated surface proteins that elicit T cell recognition, has proven to be crucial, making genetic instability and mutational load valuable markers of eventual prognosis [9–13]. However, recent research highlights that not only the amount of neoantigens, but also their heterogeneity and distribution across the tumour determines the eventual fate of immune therapies [14,15]. In particular, it has been found that a high mutational burden might not be enough for immune recognition, as only clonal antigens elicit sufficient lymphocyte response [14].

Recent modelling efforts have explicitly introduced cancer neoantigens as key ingredients in the interaction between cancer and the immune system

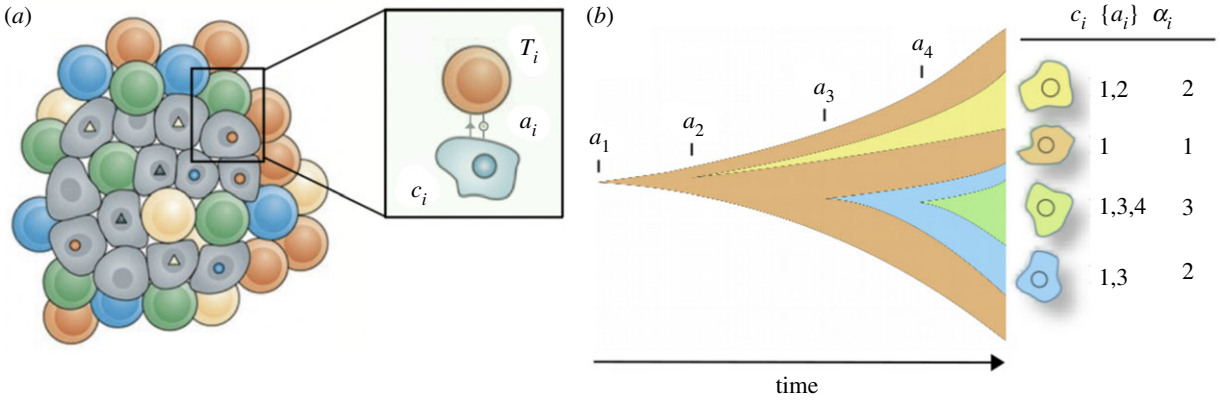

Figure 1. Mathematical framework for the cancer-immune ecology in heterogeneous antigen landscapes. T cell recognition of antigens a_i (*a*) induces a selective pressure towards immune evasion. In the lack of surveillance, neoantigens proliferate and evolve neutrally (*b*). After checkpoint blockade immunotherapy, subclonal death in our model is a function of neoantigen load α_i.

[16,17]. However, the exact role of neoantigen heterogeneity on the tempo and mode of the immune response is not fully understood. In this paper, we seek to understand, using theoretical models, how neoantigen heterogeneity evolves during immune evasion, and its effects on T cell recognition when lymphocyte function is back in place.

In this work, we introduce a mathematical framework (§2) that captures the effects of both neoantigen load and heterogeneity on immune surveillance. To this goal, two scenarios are considered, namely neutral antigen evolution in cancers that have evaded the immune system (figure 1*b*) and T cell recognition of heterogeneous tumours after immune reactivation (figure 1*a*). The model brings together knowledge from search processes in the immune response [18] and measures of epitope immunogenicity [17] to obtain a novel description of T cell reaction to heterogeneity. As shown below (§3), we predict that neoantigen diversity shapes an all-or-none transition separating growing from immune-controlled tumours. Moreover, to understand how diversity arises during malignancy progression, we infer the evolutionary pace of average antigen clonality in a growing tumour along with the evolution of the antigen-cold and -hot fractions of the malignancy.

Beyond heterogeneity, several evolutionary mechanisms are known to drive resistance to immunotherapy [19] providing novel opportunities for combination approaches to render immunotherapy more effective [20]. This includes the study of the possible timings of combining targeted agents with checkpoint blockade [21,22]. In this context, our model presents a qualitative scheduling approach that takes into account the evolutionary dynamics of neoantigen distributions when combination therapy is at work.

2. Mathematical framework

2.1. Neoantigen heterogeneity and the cancer-immune ecology

Here, we develop our approach to the ecology of tumour subclonal dynamics where each subclone comprises those cells harbouring the same epitope composition (figure 1*b*). Our toy model seeks to capture the minimal ecological dynamics of cancer subclones under immune surveillance, where each clone has a particular neoantigen set-up $\{a_i\}$ (figure 1*b*). Although several cell types participate in the immune response [23], the model focuses on those that react to antigen composition in order to capture the specific effect of neoantigen heterogeneity. Subclonal dynamics will be described by a set of coupled differential equations, namely

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = r_i c_i \left(1 - b\sum_j c_j\right) - \left(kD_iE_i\right)\frac{c_i}{g + \sum_j c_j}. \qquad (2.1)$$

The first term on the right-hand side is the logistic growth of each subclone c_i growing at rate r_i and with b indicating the inverse carrying capacity [24]. Malignant cells die at a rate (second term on the right-hand side) depending on their subclonal antigen composition, which implicitly introduces the adaptive immune system dynamics [16,17]. Here, k is the rate of killing by T cells once a target cell is recognized [25]. Subclonal death rate depends upon the dominance D_i of the immunogenic antigen in the subclone [26,27] and the efficiency E_i of the immune system to detect the antigen depending on its frequency. The most relevant novelty of model (1) is that it explicitly captures the effects of subclonal neoantigen composition by obtaining mathematical expressions for both D and E.

The model is built by considering several experimental sources. The first is that there is a causal correlation between tumour antigen load and immune response [12]. We know from studies on non-self antigens that only a few elicit immunological response [26]. Recent evidence on the overall lack of negative selection in tumours further supports that most antigens are not activating an immune attack [28]. We postulate that a higher antigen load increases the probability of presenting more immunogenic epitopes, thus increasing subclonal dominance D. Additionally, it has been shown that even highly recognizable antigens fail at inducing a T cell response if they are not present in a sufficient fraction of the tumour [14,15]. Following research on immune search mechanisms [18,29,30], we hypothesize that increased epitope heterogeneity leading to more private antigens will result in a loss of T cell efficiency E.

Finally, reduced T cell circulation [31], penetration into the solid tumour [32,33] and increased immunosuppressive factor production [34] are known to relate tumour size with reduced immune effectivity. A Michaelis–Menten saturation term (last term in equation (2.1)) is introduced following standard models of cancer–immune system dynamics [31]. The value of the Michaelis constant g indicates the cancer

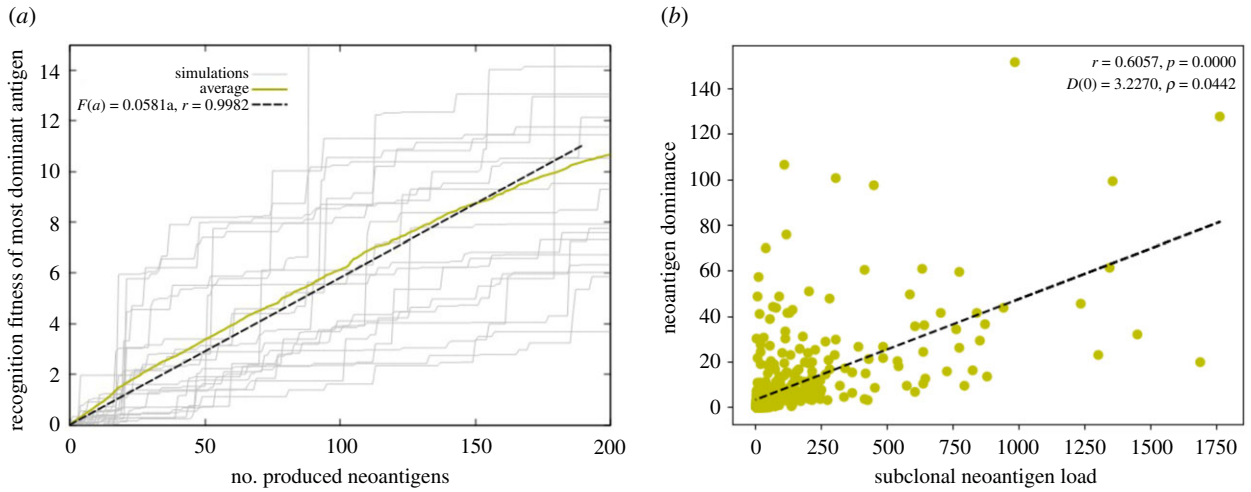

Figure 2. Correlation between subclonal neoantigen load α_i and dominance D. (*a*) Simulations show a linear increase $D_i = \rho\alpha_i$, as neoantigens are drawn and more dominant ones presented. Here, $\rho \simeq 0.06$ from our simulations on random neoantigen presentation. (*b*) Data from [17] also indicate a (weak) linear trend, with considerable data scattering.

population at which death is half of what it would be without size-related effects.

Finding the neoantigen composition $\{a_i\}$ in each subclone and estimating its immunogenicity and frequency relies on bulk data processing with sampling biases due to localized biopsies [35] and imperfect antigen prediction [36]. Mathematical techniques can bring up a simpler description to better understand how neoantigen landscapes affect prognosis by studying the underlying dynamics relating D and E to antigen composition. Below, we explain how the nature of these two key components of equation (2.1) can be determined.

2.1.1. Neoantigen dominance D

We aim at understanding the possible correlation between the immunogenicity of a subclone D_i and its antigen load α_i. Recent computational methods have been able to estimate the immunogenic capacity of cancer neoantigens [17]. The associated distribution is highly skewed: only a small subset of cancer epitopes elicits a T cell response. This result is consistent with the concept of *immunodominance* [26,27] and the overall lack of negative selection in the cancer genome [28]. Here, we assume that neoantigen load correlates with immune attack because a higher antigen burden increases the overall likelihood that a dominant one is presented. A novel and parameter-free stochastic framework is built to assess the hypothesis. In it, we measure the most dominant antigen of a cancer population as new mutations accumulate (see electronic supplementary material).

Simulations indicate that recognition potential of the most dominant antigen increases linearly with the total neoantigens of subclone α_i (figure 2*a*). This result is consistent with the notion that there is no known mechanism by which the probability of finding a more immunogenic antigen increases as more antigens are being found [12]. Since antigens are produced randomly, the probability that the next one is dominant remains stable, resulting in a linear relation $D_i \sim \alpha_i$. For larger antigen loads, maximum immunogenicity saturates as the specific epitope database is finite. This might be only due to limitations in sequencing depth [36,37]. In figure 2*b*, our simulation results can be compared with existing data on subclonal antigen burden and recognition potential from [17]. A similar trend can be obtained, although considerable scattering of the available data is at work (see electronic supplementary material for detailed descriptions of simulations and data analysis)

The previous results thus suggest approximating subclonal immunogenicity as a linear function of the antigen burden, namely

$$D_i = \rho\alpha_i. \tag{2.2}$$

Current research points to evidence for underestimation of mutational loads during next-generation sequencing [37,38], resulting in detection errors for antigens present in less than 5% of the tumour. This would imply a decrease in the steepness ρ in figure 2*b* (more mutations are actually accumulating to find the given antigens), but not a change in the linearity of the neutral accumulation dynamics, which actually governs the qualitative results of the present work.

2.1.2. Immune search efficiency E

The second part of our model definition requires a mechanistic description of the search efficiency term E in (1). The mechanism of T cell clonal selection is another key component in the immune response to cancers. Upon presentation of a given antigen by dendritic cells, helper T cells with the matching TCR replicate and release cytokines that eventually result in the expansion of a cytotoxic T cell clone [39]. This ensures efficient surveillance and further memory of previous antigenic encounters. Several search and migration processes underlie an efficient cascade [18].

The complete process is initiated by dendritic cells that recognize tumour antigens and present them to naive T cells in the lymph node [18]. Consequently, T cells become activated and migrate to the tumour site where they search for cancer cells expressing the same antigens [18]. We hypothesize here that there is a relation between neoantigen concentration in a tumour and the efficiency of the immune search processes involved. This might be key in the observed role of antigen heterogeneity in immunotherapy prognosis [14,15].

In order to make these relations explicit, we study a spatial simulation framework of a tumour surface built upon previous research [30] that minimizes the estimated parameters at play (see electronic supplementary material). In it, dendritic and T cell agents search for cancer antigens

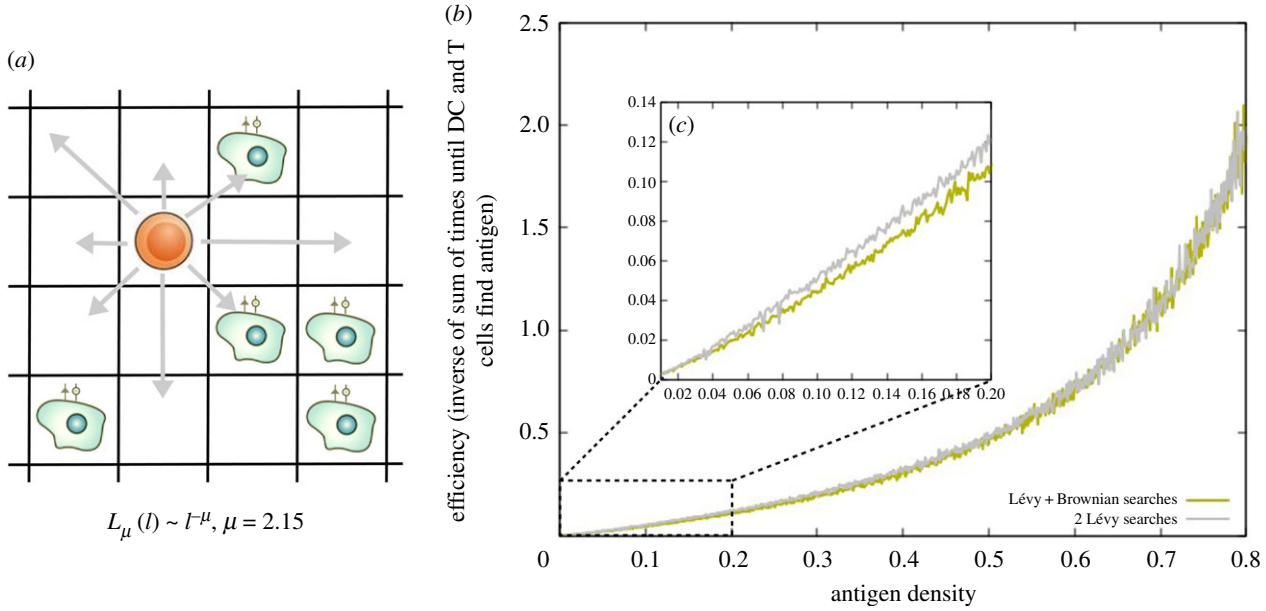

Figure 3. Modelling the effect of neoantigen frequency on immune search efficiency. (*a*) Neoantigens are distributed randomly on a two-dimensional grid according to their frequency. We sum the average times for a migrating dendritic cell and its activated T cell counterpart to find their cognate antigen distributed with density γ. (*b*) The efficiency of the immune search process scales linearly with antigen frequency γ_i for low antigen concentrations (*c*) as those in [17]. Following migration data from [29], simulations of Lévy-migrating T cells are presented for comparison, consistent with results indicating that Lévy strategies are more efficient than Brownian search [40]. The overall linear trend $E \sim \gamma$ is maintained across search strategies.

distributed according to their frequency γ (figure 3*a*). We explicitly compute the dependencies of search time and efficiency E (the inverse of search time) on neoantigen frequency (see electronic supplementary material).

This novel approach separates those steps that participate in the activation-and-killing process from the two searches directly related to neoantigen concentrations on the tumour surface: the motion of dendritic cells to find tumour antigens and the search of effector T cells that have migrated to the tumour site once activated.

The complete immune response process builds upon many layers of complexity. However, even for clonal antigens, the timescale of the search processes involved seems to be longer than those of immune cell activation or cancer cell removal (see electronic supplementary material), supporting the notion that the frequency-dependent processes play a determining role in the average speed of the immune response.

Efficiency diverges in the simulations when an antigen is found in most of the cells of the tumour surface. However, clinical antigen frequencies are much smaller, and a linear trend can be found for small γ levels similar to those found in tumours (figure 3*c*, [17]). This translates into immune efficiency E being a linear function of the frequency of the dominant antigen in the subclone, $E_i = s\gamma_i$. Here, s is the slope of the linear correlation, related to other molecular or non-antigenic processes that affect the timing of the search, such as immune cell activation or cancer cell removal [18], which we find to be rate-limited by the timescale of the antigen-search processes at play (see electronic supplementary material). By rewriting equation (2.1), we now map subclonal death rate with antigen load α_i and dominant epitope concentration γ_i

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = r_i c_i \left(1 - b\sum_j c_j\right) - k(\rho\alpha_i s\gamma_i)\frac{c_i}{g + \sum_j c_j}. \qquad (2.3)$$

In model (2.3), subclones will die at rates corresponding to their number of antigens (as more antigens increase the probability of presenting a dominant one). The frequency of their most dominant antigen will also increase the efficiency of dendritic search and T cell surveillance. Is it possible to gain insight from (2.3) into how T cells react at neoantigen distributions? Are there common patterns separating responding from non-responding patients, and what can we learn from them? In the Results section, we study the relation between this model and the experimentally observed correlation between neoantigen heterogeneity and poor prognosis.

2.2. Neutral evolution of neoantigen distributions

So far our mathematical framework introduces cancer cell death under a correctly functioning T cell cohort. However, the evolution of mechanisms to avoid T and B lymphocyte attack is recognized as a hallmark of cancer cells [6] and needs to be introduced too. The lack of an efficient immune response ensures that neoantigens are no longer negatively selected [28,41], making for tumours where neoantigen load and heterogeneity respond to neutral evolutionary dynamics [42]. What epitope landscape $\{\alpha_i, \gamma_i\}$ will T cells find when they are activated and equation (2.3) is back in place?

The role played by neutral evolution in cancer has seen increasing attention [28,37,42]. In a recent work, Lakatos and colleagues developed a computational approach for the evolutionary dynamics of neoantigens, and indicators of effectively neutral evolution are found in colorectal cancer exome sequencing data [16] consistent with the overall lack of negative selection across the cancer genome [28]. We present a coherent analytical approach to understand how neoantigen load and concentration evolve in the absence of T cell surveillance. Further analytical methods are provided to estimate the fraction of antigen-cold cells in a given tumour. Stochastic simulations of neoantigen evolution in

growing tumours are built to verify our analytical estimates (see electronic supplementary material for the mathematical and computational methods).

2.3. The effect of combination therapy on neoantigen landscapes

The evolution and composition of the neoantigen landscape is key to model and understand the interaction between cancer and the immune system and needs to be taken into account within our modelling approach. Recent research is focusing on the use of combination therapy to render checkpoint blockade more effective [20–22]. We here study the effects of combination therapy on neoantigen heterogeneity and propose a novel therapeutic approach able to modulate antigen landscapes.

Interdisciplinary approaches have already considered taking advantage of evolution to explore novel therapeutic designs. Mathematical models have studied the use of drug sequencing to avoid the evolution of cross-resistance [43] or the notion of an *evolutionary double bind* [44] to take advantage of its metabolic cost [45]. In this context, experimental evidence indicates that alkylating agents causing DNA damage can increase subclonal neoantigen burdens [14,46]. This results in a more heterogeneous neoantigen landscape. On the other hand, immunotherapy often results in selective pressures towards immune evasion through silencing of clonal antigens [19,47]. We here study other therapeutic combinations that do not increase neoantigen heterogeneity or reduce epitope presentation.

In particular, molecular-targeted therapies such as BRAF and MEK inhibitors for melanoma are being combined with immune checkpoint blockade in the search for optimal drug sequencing [20–22]. Knowing that resistance to targeted therapies mediated by preexisting or acquired mutations is common [48,49], we study the therapeutic designs that could take advantage of it. The stochastic simulations previously described are modified to include a generalistic targeted therapy and mutations driving resistance to it (see electronic supplementary material). We follow the dynamics of neoantigen distributions to understand the effects of drug resistance on tumour and epitope heterogeneity, and study which scenarios are better suited for checkpoint blockade efficiency.

3. Results

3.1. Before checkpoint blockade immunotherapy

The amount and distribution of neoantigens is key to predict the response of tumours to immunotherapy. However, early selective pressure following T cell recognition is known to drive immune evasion [41]. Until checkpoint blockade immunotherapy reactivates lymphocyte attack, neoantigen distributions evolve neutrally [28,42].

To address this problem, we use our model to study the evolution of average antigen load and clonality in growing tumours. Tumour growth and neoantigen production are known to slow down along progression. We can characterize the fastest neoantigen dynamics by studying the exponential phases of tumour growth. This will also useful for the proposed combination therapy design that takes into account small tumours and distant metastases.

In particular (see the electronic supplementary material for details), we find that the amount of antigenic mutations

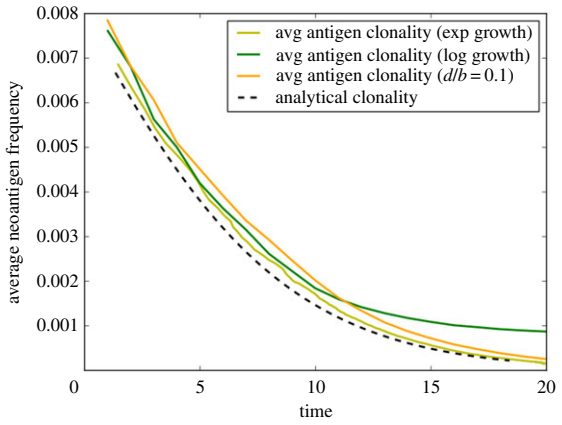

Figure 4. Evolution of antigen frequency during tumour growth. The analytical result (3.3) is compared with simulations of exponential growth, logistic growth under extreme spatial constraints ($K = 10^4$) and increased cell death ($d/b = 0.1$).

in the tumour, M_α will follow (see [42])

$$M_\alpha(t) = 2\mu p_\alpha \frac{b}{b-d} c_0(e^{(b-d)t} - 1). \tag{3.1}$$

where the tumour grows at an average rate $r = b - d$, μ stands for the overall mutational probability and $p_\alpha \sim 0.13$ is the approximate portion of mutations predicted to bind MHC1 indicating possible immunogenicity [32]. The two factor stems from the fact that both daughter cells can undergo errors at birth at rate $b\mu p_\alpha$. The result is consistent with logistically growing tumours, where slowing down of neoantigen production is only significative for very small carrying capacities (see electronic supplementary material). Moreover, another key element in therapy prognosis is the existence of cells that do not harbour antigens. It can be shown (see electronic supplementary material) that analytical results and simulations predict an exponential decay of the antigen-cold fraction of the tumour

$$c_{\alpha^-}(t) = e^{-2r\mu p_\alpha t}. \tag{3.2}$$

This result is indicative of the rapid dynamics at which microsatellite-unstable tumours become populated by antigen-hot cells, consistent with experimental insight [13]. The model seems to support the idea that, provided sufficient mutational load [37,46], it could be the excess of heterogeneity and not the lack of antigens that drives checkpoint blockade therapy failure [14,15].

We also study the dynamics of average neoantigen clonality as tumours grow (see electronic supplementary material). We find that the average clonality decays in time as

$$\langle \gamma(t) \rangle = \frac{1}{c_0(e^{(b-d)t} - 1)} [\psi^{(0)}(\theta e^{(b-d)t} + 1) - \psi^{(0)}(\theta + 1)] \tag{3.3}$$

where $\theta = 2\mu p_\alpha b\, c_0/(b-d)$ and $\psi^{(0)(x)}$ is the digamma function [50]. This result can be compared to simulations (figure 4*a*) and is a first analytical measure of how heterogeneity increases in the neoantigen landscape.

The predicted antigen clonality decay results from exponential growth in early tumours or small metastatic burdens, interesting for the therapeutic scheme proposed later on. However, tumour growth is known to slow down at later stages. Simulations indicate that logistic growth has a small effect on heterogeneity dynamics (figure 4), even

for very small carrying capacities of $K = 10^4$. This is because initial exponential growth ensures a very fast decay of average clonality (equation (3.3)) that remains low on later tumour growth phases. Increased cancer death (figure 4) or finite epitope landscape scenarios (see electronic supplementary material) are also consistent with the analytical estimate.

Neoantigen heterogeneity increases rapidly in growing tumours. This could be indicative of a decay in the efficiency of T cell surveillance, giving an explanation for the role of epitope clonality observed in immunotherapy prognosis [14,15]. Can our mathematical framework capture this effect?

3.2. After checkpoint blockade immunotherapy

These results predict the evolution of the neoantigen landscapes during immune evasion. After checkpoint blockade immunotherapy, T cell attack is back in place and equation (2.3) holds. The model is solvable for specific subclones. However, we aim to understand whole-tumour dynamics when immunotherapy is administered, so that several approximations are introduced to reduce the set of parameters at play.

On the one hand, neoantigen presentation does not affect the rate of cell division r_i [17]. Therefore, subclones characterized by any neoantigen composition are considered to replicate at an average rate $\langle r\rangle$. The final result stems from whole-tumour dynamics, depending only on $\langle r\rangle$ which can be measured from patient data [17].

Another similar assumption regarding antigen neutrality can be performed. Provided that neoantigens were isolated from immune attack prior to therapy [41], dominant antigens at the time of therapy are not negatively selected [28] and, on average, are as common as the rest. This allows us to write $\gamma_i \sim \langle\gamma\rangle$.

Inspired in the analysis of antigenic diversity thresholds in the evolution of HIV [51], we want to understand the conditions under which all subclonal growth is controlled by the immune surveillance mechanisms. This will occur provided that the following inequality:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} < 0 \tag{3.4}$$

holds for all $i = 1, \ldots, S$. A necessary condition for this to be true is that the sum of subclonal dynamics (2.3) over all S clones is negative

$$\sum_i^S \frac{\mathrm{d}c_i}{\mathrm{d}t} = \sum_i^S c_i \left[\langle r\rangle\left(1 - b\sum_j^S c_j\right) - \frac{k\rho\alpha_i s}{g + \sum_j^S c_j}\langle\gamma\rangle\right] < 0. \tag{3.5}$$

Even if (3.5) is not a sufficient condition (since a single clone could grow despite overall negative replication) it is required for immune surveillance to succeed: if the overall sum of growth rates is not negative, at least one of the clones will outgrow immune barriers. Since all clone populations c_i are either zero or positive, the sum of terms inside parentheses must be negative.

The number of elements in the sum S accounts for the number of subclones harbouring identical epitope configurations (figure 1*b*). Summing over (3.5) we find the following inequality:

$$\frac{\langle r\rangle}{k\rho s}\left(1 - b\sum_j^S c_j\right)\left(g + \sum_j^S c_j\right) < \langle\gamma\rangle\frac{\sum_i^S \alpha_i}{S}. \tag{3.6}$$

This result indicates the possibility of a threshold condition separating tumour growth from immune surveillance provided that tumour size $\sum_j c_j$ can be estimated. The left-hand side is the replication/predation ratio, that has to be larger than the tumour immunogenicity term of the right-hand side. Furthermore, the product of the logistic and Michaelis–Menten terms on the left-hand side defines to which extent the spatial constraints affect more tumour growth or else immune circulation. It can be seen that the $(1 - b\sum_j^S c_j)(g + \sum_j^S c_j)$ term can become large only for large carrying capacities and tumour sizes (see electronic supplementary material).

We are also interested in understanding (3.5) in the particular scenario of small tumours. On the one hand, we want to study the effects of a targeted therapy reducing tumour bulk prior to immunotherapy, so that the immune system will most probably face a small resistant subclone or distant metastases. On the other hand, because of the competitive release of resistant cells after therapy [45], we are looking at modelling a time window of fastest antigen production when tumour growth is unbounded, which happens when $\sum_j c_j$ is away from its carrying capacity and the growth dynamics are exponential. In this scenario, the nonlinear effects of large tumour masses can be neglected. Similar models considering unbounded growth in subclones are already in place [43], acknowledging that the effects of extracellular matrix barriers [33] or immunosuppressive factor production [34] are reduced. Now the threshold condition contains exclusively tumour-averaged parameters: average antigen frequency and average subclonal neoantigen load should outgrow the growth/recognition ratio

$$\langle\gamma\rangle\frac{\sum_i^S \alpha_i}{S} > \frac{\langle r\rangle g}{k\rho s}. \tag{3.7}$$

The existence a catastrophic shift separating tumour growth from extinction as a function of the average neoantigen clonality resembles recent clinical studies [14,15]. The parameters on the right-hand side of equation (3.7) are dependent on tumour type and microenvironment specificities. Producing a quantitative prediction for given cancer types and measures is away from the scope of the article. However, the threshold depends on a tumour immunogenicity value that we can study using existing data. To which extent does the immunogenicity marker of equations (3.6) and (3.7) correlate with patient prognosis after checkpoint blockade immunotherapy?

Using available neoantigen estimates from the database in [17] (see electronic supplementary material), we can study the correlation between our result and months of survival in anti-PD-1-treated patients with lung cancer [9] and anti-CTLA-4 treated patients with melanoma [10,11]. Without accessible data on tumour size $\sum_j c_j$, Kaplan–Meier curves and cumulative hazard ratios of overall survival of anti-CTLA-4 treated melanoma patients from [10,11] seem to indicate that our measure extracted from equation (3.7) could be a better biomarker for immunotherapy prognosis than total neoantigen load (figure 5), consistent with experimental results [14].

Modelling, simulations and data seem to consistently indicate that neoantigen heterogeneity might shape a threshold condition separating tumour growth from cure. Average antigen frequency decays rapidly as tumours grow, and advanced malignancies will probably be highly heterogeneous and hard to target by immune activation. Could we design a

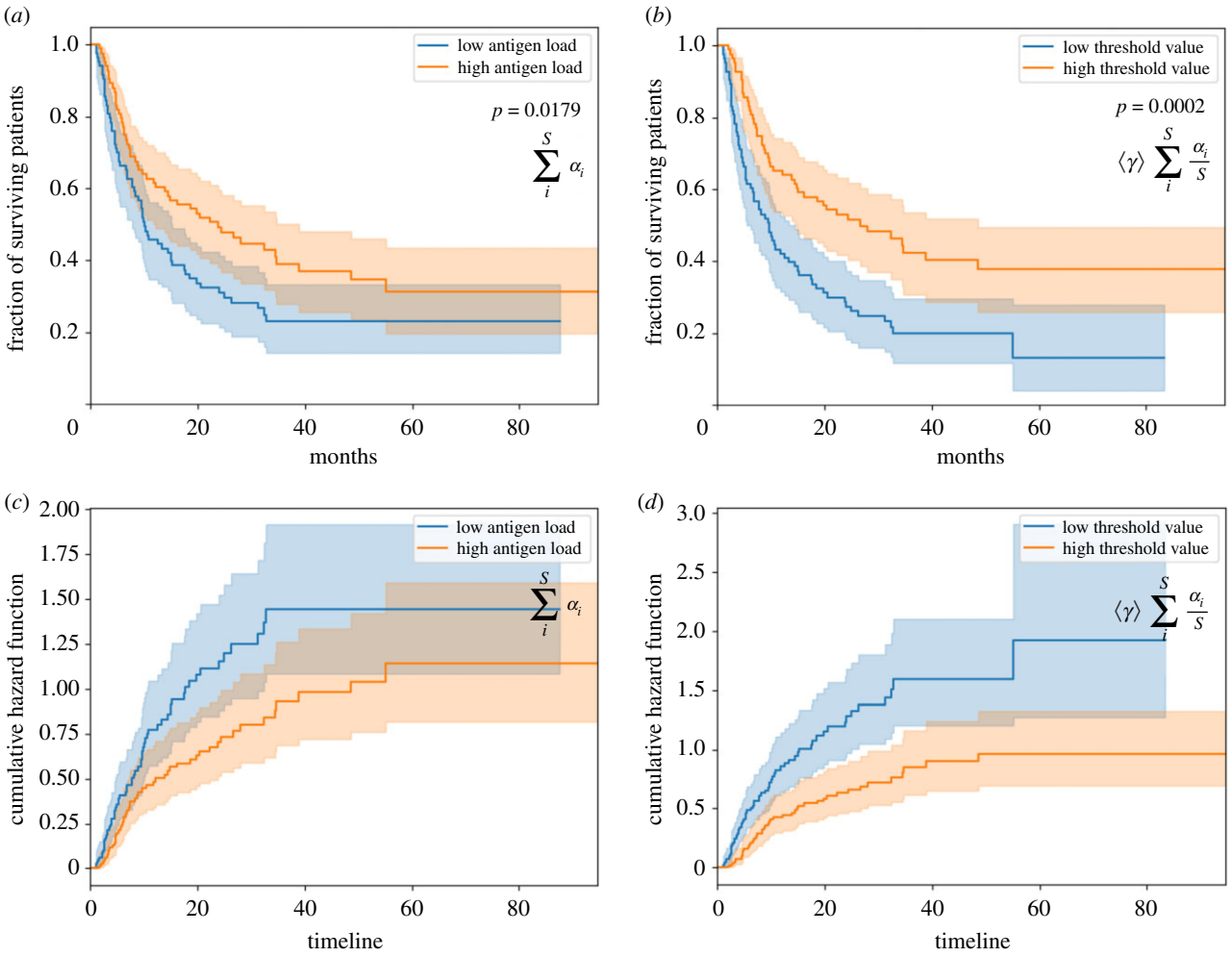

Figure 5. Correlation of patient biomarkers with months of survival after checkpoint inhibition therapy. The Kaplan–Meier curves for decay of survival probability are depicted, with 166 anti-CTLA-4 treated melanoma patients from [10,11] separated in two groups by the median of either neoantigen load (*a*, log-rank test *p*-value 0.0179) and our threshold value $\gamma \sum \alpha/S$ (*b*, log-rank test *p*-value 0.0002). Cumulative hazard ratios (*c*,*d*) also support that including average neoantigen clonality adds predictive capacity to the well-accepted neoantigen load marker. This suggests $\gamma \sum \alpha/S$ as a possible complementary biomarker for immunotherapy prognosis in melanoma. Including a cohort of 30 anti-PDL-1 treated lung cancer patients from [9] does not improve results (see electronic supplementary material, figure S6).

therapeutic scheme able to steer tumour evolution towards a more homogeneous neoantigen landscape?

3.3. Combination therapy could reduce neoantigen heterogeneity

The mathematical model predicts an heterogeneity threshold beyond which lymphocytes fail to reduce tumour growth, consistent with recent research [14,15]. Results also indicate that neoantigen frequency decreases rapidly during growth, meaning that therapy reactivating the immune system will not necessarily succeed. Now we study the (simulated) effect of administering a molecular-targeted therapy in the case that resistance is already present, following existing experimental efforts [20,22].

As therapy is administered, the death rate of sensitive cells increases, leading to a general decay in tumour cell number (figure 6*a*) that is commonly seen in combination approaches targeting tumour burden [20]. This decrease in cell size comes with a halt in the production of novel antigens, while total present antigens are reduced as some may go extinct (figure 6*b*).

When therapy is administered, the overall reduction of cell number produces a rapid increase in average antigen frequency. Knowing that resistance mutations are usually in place [37,49], can it be used in our favour? In spite of many antigens going extinct, those that belong to a surviving lineage (of cells that are resistant to therapy) increase their overall frequency (figure 6*c*).

Based on our *in silico* models, we postulate that resistance to therapy produces a selective sweep to a more homogeneous epitope landscape, which is easier to target by a reactivated immune system. Furthermore, previous results indicated that, along the growth process of the tumour, the fraction of antigen-cold cells decays rapidly. This is not altered by the targeted agent provided it has no alkylating effect, so that antigen accumulation continues until checkpoint blockade is administered. Until the resistant clone has grown into a full-size tumour, we observe a time window during which the average frequency of neoantigens is higher than when the cancer was of detectable size (figure 6*c*). This represents a first qualitative effort to understand combination therapy scheduling [20] when resistance mutations are in place.

4. Discussion

Immunological approaches to cancer therapy have seen remarkable advances in recent years [5,8,13]. However, not all

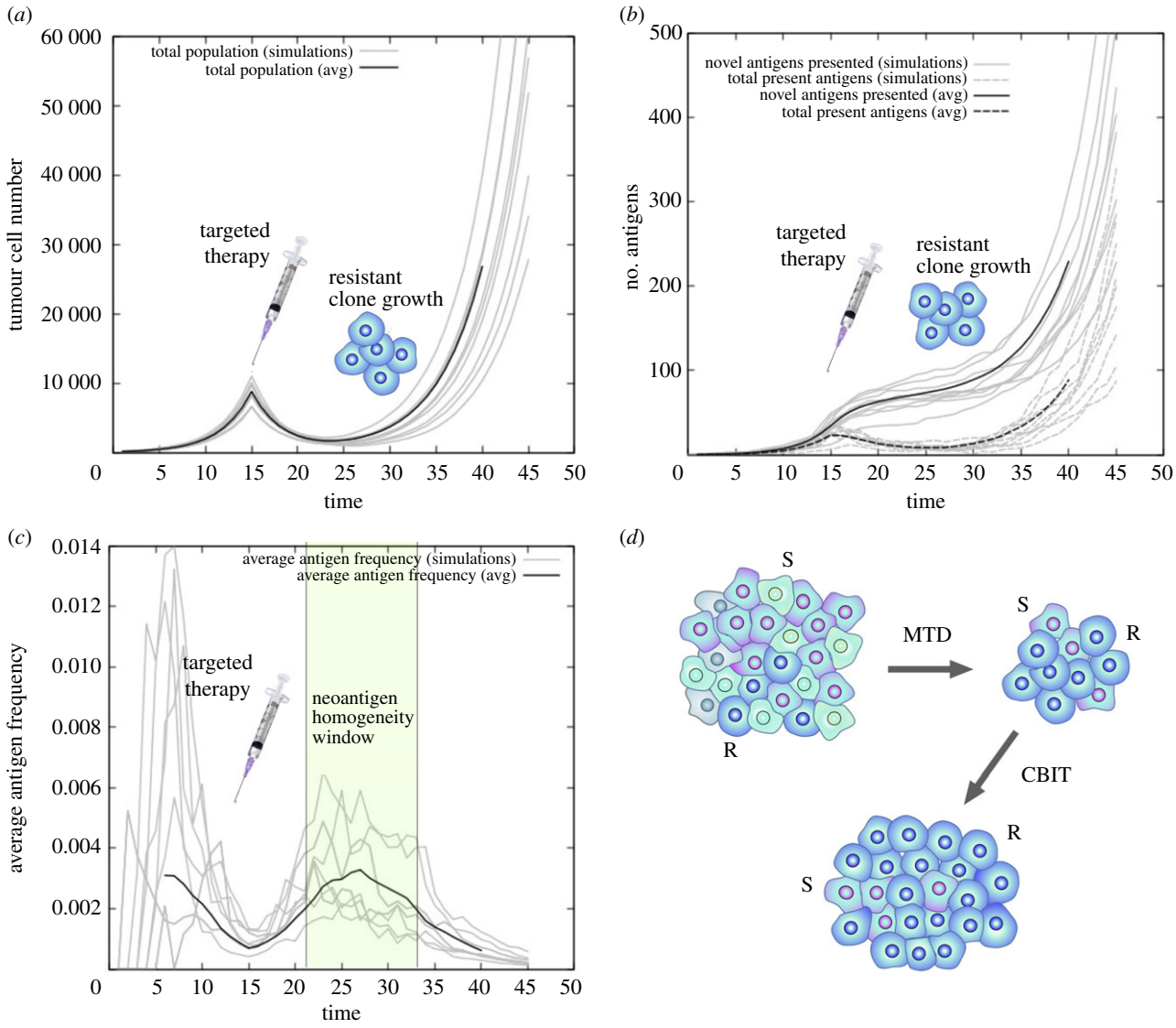

Figure 6. Evolutionary dynamics of neoantigens during therapy. (*a*) Cell number decreases after therapy until a resistant subclone repopulates the tumour. (*b*) Production of novel antigens slows down after therapy (black line) while total present antigens are reduced (dashed line). (*c*) Tumour size decay produces an increase in clonality of remaining antigens. Before complete relapse, increased average clonality creates a time window where the tumour is more neoantigen-homogeneous and T cells increase their efficiency (timescale in days). (*d*) A scheme of the population dynamics shows subclonal outgrowth after maximum tolerated dose (MTD) targeted treatment. We postulate that these processes elicit homogenization of the neoantigen landscape where checkpoint blockade immunotherapy (CBIT) can be more effective.

underlying mechanisms of immune surveillance and resistance are understood, and relapse is a common outcome [5,19]. A critical aspect of immune efficiency lies in the incidence of high neoantigen heterogeneity in poor therapy response [14–17]. To address some of these problems, in this paper, we have built a minimal model of heterogeneous cancer populations explicitly introducing their antigenic composition. In particular, neoantigen-related subclonal death takes into account antigen load and frequency across the tumour while keeping the number of model parameters small. Our framework contemplates computational methods and data in both epitope immunogenicity [17] and immune search efficiency [18,29,30] to translate previous evidence into a mathematical model able to produce insight into the underlying subclonal dynamics. Analytical results of the new model indicate that a heterogeneity threshold separates cancer growth from immune control, so that highly heterogeneous neoantigen landscapes might impair immune efficiency.

By incorporating our results into survival studies, predictive power could be added to immunotherapy prognosis in melanoma patients, providing a novel therapy biomarker consistent with experimental results [14,15]. To understand why tumours have heterogeneous neoantigen landscapes, the model studies the evolutionary dynamics of antigenic mutations. As shown by the model, the antigen-cold fraction of the tumour becomes rapidly smaller, meaning that the growing tumour is rapidly populated by antigens. However, analytical estimates and simulations indicate a fast decay of average antigen clonality during immune evasion phases of tumour growth. This is consistent with the correlation between increased antigen heterogeneity and poor response to checkpoint blockade immunotherapy.

Further layers of tumour complexity that do not depend on neoantigen composition have been taken out of the system to produce a treatable model, particularly in the immune search modelling. With this, we can compare the qualitative results of our approach with experimental evidence while keeping only with a minimal set of measurable parameters. Furthermore, we have studied specific scenarios where the model simplifies, such as immune activity in small tumours and

metastases, where nonlinear effects of tumour size can be neglected as in other therapeutic designs [43]. This also allows the use of existing data on neoantigen distributions where population counts are not available. Further non-antigenic events in the search and cytotoxic processes could be introduced to obtain a more quantitative approach, together with including extrinsic noise in the cancer–immune system interaction [52].

Mathematical approaches have shown how evolutionary dynamics can be used to obtain novel insight into the sequencing [43] or timing [45] of combination therapies. Ongoing research is currently exploring the possibility of combining targeted therapy with checkpoint blockade [21,22]. Our model highlights how evolutionary dynamics underlying resistance to a targeted agent can increase neoantigen homogenization. Although a refractory drug is not of interest in principle, resistance is so common that we propose a combination approach that takes advantage of it.

Simulations indicate that positive selection for a drug-resistant clone results in increased antigen clonality. Before complete relapse, a second hit with immunotherapy could be more effective as reactivated lymphocytes face a more homogeneous tumour. Continuation of immune attack after checkpoint blockade is known to drive selective pressure towards epigenetic silencing of targeted neoantigens [47]. Further analysis could result in more complex combination schemes that enhance neoantigen generation after prolonged T cell attack.

Data accessibility. This article has no additional data.

Authors' contributions. Both authors have contributed equally in this work.

Competing interests. We declare we have no competing interests.

Funding. This work has been supported by the Botín-Foundation by Banco Santander through its Santander Universities Global Division, a MINECO grant no. FIS2015-67616-P (MINECO/FEDER, UE) fellowship, and AGAUR FI grant by the Universities and Research Secretariat of the Ministry of Business and Knowledge of the Generalitat de Catalunya and the European Social Fund and by the Santa Fe Institute.

Acknowledgements. G.A. thanks Eszter Lakatos for interesting discussions on neoantigen evolution within the *Cancer Ecology and Evolution 2019* conference at the Wellcome Genome Campus. Both authors thank members of the Complex Systems Lab as well as the HEF debates for insightful discussions.

References

1. Greaves M, Maley CC. 2012 Clonal evolution in cancer. *Nature* **481**, 306–313. (doi:10.1038/nature10762)
2. Negrini S, Gorgoulis VG, Halazonetis TD. 2010 Genomic instability—an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* **11**, 220–228. (doi:10.1038/nrm2858)
3. Holohan C, Schaeybroeck SV, Daniel B, Longley DB, Johnston PG. 2013 Cancer drug resistance: an evolving paradigm. *Nat. Rev. Cancer* **13**, 714–726. (doi:10.1038/nrc3599)
4. Burell RA, Swanton C. 2014 Tumour heterogeneity and the evolution of polyclonal drug resistance. *Mol. Oncol.* **8**, 1095–1111. (doi:10.1016/j.molonc.2014.06.005)
5. Ribas A. 2015 Adaptive immune resistance: how cancer protects from immune attack. *Cancer Disc.* **5**, 915–919. (doi:10.1158/2159-8290.CD-15-0563)
6. Hanahan D, Weinberg RA. 2011 Hallmarks of cancer: the next generation. *Cell* **144**, 646–74. (doi:10.1016/j.cell.2011.02.013)
7. Xing Y, Hongquist KA. 2012 T-cell tolerance: central and peripheral. *Cold Spring Harb. Perspect. Biol.* **4**, 6. (doi:10.1101/cshperspect.a006957)
8. Ribas A, Wolchok JD. 2018 Cancer immunotherapy using checkpoint blockade. *Science* **359**, 1350–1355. (doi:10.1126/science.aar4060)
9. Rizvi NA *et al.* 2015 Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128. (doi:10.1126/science.aaa1348)
10. Snyder A *et al.* 2014 Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl J. Med.* **371**, 2189–2199. (doi:10.1056/NEJMoa1406498)
11. Van Allen EM *et al.* 2015 Genomic correlates of response to CTLA4 blockade in metastatic melanoma. *Science* **350**, aad0095. (doi:10.1126/science.aad0095)
12. Schumacher TN, Schreiber RD. 2015 Neoantigens in cancer immunotherapy. *Science* **348**, 69–74. (doi:10.1126/science.aaa4971)
13. Xiao Y, Freeman GJ. 2015 The microsatellite instable subset of colorectal cancer is a particularly good candidate for checkpoint blockade immunotherapy. *Cancer Disc.* **5**, 16–18. (doi:10.1158/2159-8290.CD-14-1397)
14. McGranahan N *et al.* 2016 Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351**, 1463–1469. (doi:10.1126/science.aaf1490)
15. Wolf Y *et al.* 2019 UVB-induced tumor heterogeneity diminishes immune response in melanoma. *Cell* **179**, 219–235. (doi:10.1016/j.cell.2019.08.032)
16. Lakatos E *et al.* 2019 Evolutionary dynamics of neoantigens in growing tumours. BioRxiv 209;536433. (doi:10.1101/536433)
17. Luksza M *et al.* 2017 A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* **551**, 517–20. (doi:10.1038/nature24473)
18. Krummel MF, Bartumeus F, Gérard A. 2016 T cell migration, search strategies and mechanisms. *Nat. Rev. Immunol.* **16**, 193. (doi:10.1038/nri.2015.16)
19. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. 2017 Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* **168**, 707–723. (doi:10.1016/j.cell.2017.01.017)
20. Zappasodi R, Merghoub T, Wolchok JD. 2018 Emerging concepts for immune checkpoint blockade-based combination therapies. *Cancer Cell* **33**, 581–598. (doi:10.1016/j.ccell.2018.03.005)
21. Ribas A *et al.* 2015 Phase I study combining anti-PD-L1 (MEDI4736) with BRAF (dabrafenib) and/or MEK (trametinib) inhibitors in advanced melanoma. *J. Clin. Oncol.* **33**, 3003. (doi:10.1200/jco.2015.33.15_suppl.3003)
22. Bendell JC *et al.* 2016 Clinical activity and safety of cobimetinib (cobi) and atezolizumab in colorectal cancer (CRC). *J. Clin. Oncol.* **34**, 3502. (doi:10.1200/JCO.2016.34.15_suppl.3502)
23. Bindea G *et al.* 2013 Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782–795. (doi:10.1016/j.immuni.2013.10.003)
24. Gatenby RA, Gillies RJ. 2008 A microenvironmental model of carcinogenesis. *Nat. Rev. Can.* **8**, 56–61. (doi:10.1038/nrc2255)
25. Budhu S *et al.* 2010 CD^{8+} T cell concentration determines their efficiency in killing cognate antigen-expressing syngeneic mammalian cells *in vitro* and in mouse tissues. *J. Exp. Med.* **207**, 223–235. (doi:10.1084/jem.20091279)
26. Yewdell JW, Bennink JR. 1999 Immunodominance in major histocompatibility complex class I-restricted T lymphocyte responses. *Annu. Rev. Immunol.* **17**, 51–88. (doi:10.1146/annurev.immunol.17.1.51)
27. Schreiber H, Wu TH, Nachman J, Kast WM. 2002 Immunodominance and tumor escape. *Semin. Cancer Biol.* **12**, 25–31.
28. Martincorena I *et al.* 2017 Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041. (doi:10.1016/j.cell.2017.09.042)
29. Harris TH *et al.* 2012 Generalized Lévy walks and the role of chemokines in migration of effector CD^{8+} T cells. *Nature* **486**, 545–548. (doi:10.1038/nature11098)
30. Macfarlane FR, Chaplain MA, Lorenzi T. 2019 A stochastic individual-based model to explore the role of spatial interactions and antigen recognition

in the immune response against solid tumours. *J. Theor. Biol.* **480**, 43–55. (doi:10.1016/j.jtbi.2019.07.019)

31. Kuznetsov VA, Makalkin IA, Taylor MA, Perelson AS. 1994 Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. *Bull. Math. Biol.* **56**, 295–321. (doi:10.1016/S0092-8240(05)80260-5)
32. Matzavinos A, Chaplain MA, Kuznetsov VA. 2004 Mathematical modelling of the spatio-temporal response of cytotoxic T-lymphocytes to a solid tumour. *Math. Med. Biol.* **21**, 1–34. (doi:10.1093/imammb/21.1.1)
33. Salmon H *et al.* 2012 Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *J. Clin. Invest.* **122**, 899–910. (doi:10.1172/JCI45817)
34. Wahl SM *et al.* 1988 Transforming growth factor-beta is a potent immunosuppressive agent that inhibits IL-1-dependent lymphocyte proliferation. *J. Immunol.* **140**, 3026–3032.
35. Stead LF, Sutton KM, Taylor GR, Quirke P, Rabitts P. 2013 Accurately identifying low-allelic fraction variants in single samples with next-generation sequencing: applications in tumor subclone resolution. *Hum. Mutat.* **34**, 1432–1438. (doi:10.1002/humu.22365)
36. Yadav M *et al.* 2014 Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* **515**, 572–576. (doi:10.1038/nature14001)
37. Loeb LA *et al.* 2019 Extensive subclonal mutational diversity in human colorectal cancer and its significance. *Proc. Natl Acad. Sci. USA* **116**, 26 863–26 872. (doi:10.1073/pnas.1910301116)
38. Roerink SF *et al.* 2018 Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature* **556**, 457–462. (doi:10.1038/s41586-018-0024-3)
39. Starr TK, Jameson SC, Hogquist KA. 2003 Positive and negative selection of T cells. *Annu. Rev. Immunol.* **21**, 139–176. (doi:10.1146/annurev.immunol.21.120601.141107)
40. Viswanathan GM, Buldyrev SV, Havlin S, Raposo EP, Stanley HE. 1999 Optimizing the success of random searches. *Nature* **401**, 911–914. (doi:10.1038/44831)
41. DuPage M, Mazumdar C, Schmidt LM, Cheung AF, Jacks T. 2012 Expression of tumour-specific antigens underlies cancer immunoediting. *Nature* **482**, 405–409. (doi:10.1038/nature10803)
42. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. 2016 Identification of neutral tumor evolution across cancer types. *Nat. Gen.* **48**, 238–244. (doi:10.1038/ng.3489)
43. Beckman RA, Schemmann GS, Yeang CH. 2012 Impact of genetic dynamics and single-cell heterogeneity on development of nonstandard personalized medicine strategies for cancer. *Proc. Natl Acad. Sci. USA* **109**, 14 586–14 591. (doi:10.1073/pnas.1203559109)
44. Gatenby RA, Brown J, Vincent T. 2009 Lessons from applied ecology: cancer control using an evolutionary double bind. *Cancer Res.* **69**, 7499–7502. (doi:10.1158/0008-5472.CAN-09-1354)
45. Gatenby RA, Silva AS, Gillies RJ, Frieden BR. 2009 Adaptive therapy. *Cancer Res.* **69**, 4894–4903. (doi:10.1158/0008-5472.CAN-08-3658)
46. Campbell BB *et al.* 2017 Comprehensive analysis of hypermutation in human cancer. *Cell* **171**, 1042–1056. (doi:10.1016/j.cell.2017.09.048)
47. Anagnostou V *et al.* 2017 Evolution of neoantigen landscape during immune checkpoint blockade in non-small cell lung cancer. *Cancer Disc.* **7**, 264–276. (doi:10.1158/2159-8290.CD-16-0828)
48. Corcoran RB *et al.* 2012 EGFR-mediated reactivation of MAPK signaling contributes to insensitivity of BRAF-mutant colorectal cancers to RAF inhibition with vemurafenib. *Cancer Disc.* **2**, 227–235. (doi:10.1158/2159-8290.CD-11-0341)
49. Diaz Jr LA *et al.* 2012 The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. *Nature* **486**, 537–540. (doi:10.1038/nature11219)
50. Abramowitz M, Stegun IA (eds). 1972 6.3 psi (Digamma) Function. In *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, pp. 258–259. New York, NY: Dover.
51. Nowak MA, May RM, Anderson RM. 1990 The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease. *Aids* **4**, 1095–1103. (doi:10.1097/00002030-199011000-00007)
52. d'Onofrio A. 2010 Bounded-noise-induced transitions in a tumor–immune system interplay. *Phys. Rev. E* **81**, 021923. (doi:10.1103/PhysRevE.81.021923)

Supplementary Material to:
Tumor neoantigen heterogeneity thresholds provide a time window for combination immunotherapy

Guim Aguadé-Gorgorió[*1,2] and Ricard Solé[†1,2,3]
[1]ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain
[2]Institut de Biologia Evolutiva, CSIC-UPF, 08003 Barcelona, Spain
[3]Santa Fe Institute, Santa Fe NM 87501, USA

I. MATHEMATICAL FRAMEWORK

A. Evolution of Neoantigen Distributions

We here study the evolution of average antigen load and clonality in growing tumors. Due to the scenarios of interest for the original manuscript, exponential growth is taken into account, so that the cancer subclone has dynamics $c(t) = c_0\exp((b-d)t)$. We know that neoantigen load will obey (see [1])

$$\frac{dM_\alpha}{dt} = 2\mu p_\alpha b c(t) \tag{1}$$

where M_α is the amount of antigenic mutations in the tumor, μ stands for the overall mutational probability and $p_\alpha \sim 0.13$ stands for the approximate portion of mutations predicted to bind MHC1 indicating possible immunogeneicity [2]. The 2 factor stems from the fact that both daughter cells can undergo errors at birth at rate $b\mu p_\alpha$. Introducing population dynamics we obtain an analytical expression consistent with simulations (Fig. S1). Comparison with logistic growth under even very small carrying capacities demonstrates the validity of the exponential assumption in the given setup, consistent with antigen load data from [3] were tumors are estimated to harbor around 10-1000 immunogenic epitopes. Further single-cell cloning analysis is likely to result in larger results for mutational burden [4].

$$M_\alpha(t) = 2\mu p_\alpha \frac{b}{b-d} c_0 \left(e^{(b-d)t} - 1\right). \tag{2}$$

Average mutation probability μ is known to evolve in time as further mutations affect DNA repair mechanisms [5]. This would speed up the rate at which neoantigens are produced, inducing a further nonlinearity in $M(t)$, resulting in faster frequency decays. A lower bound can still be found, and a finite genomic pool [6] and the difficulty of finding novel highly dominant epitopes [3] indicate that immune killing might plateau as $D(\alpha(\mu))$ saturates at high μ values. In the absence of cell death, the frequency of a given antigen is related to the population size at presentation, $\gamma(\alpha_i) = 1/c(\alpha_i)$ and remains constant [1].

Recent experimental evidence [6] indicates that mutational rates might be so high that the infinite-sites assumption from [1] might underestimate tumor mutational burden, meaning that a given loci will be mutated in more than one cell. This could imply that a given neoantigen can be produced in more than one cell independently increasing its overall frequency. The effect of this in our calculations will be later shown to be very small (Fig. S2). Together with this, cell death adds further deviation from $\gamma(\alpha_i) = 1/c(\alpha_i)$: when cells die out, this frequency can increase for antigens with surviving lineages [7] up to

$$\gamma(\alpha_i) = \frac{1-(d/b)^{c(t)}}{(1-d/b)c(t)} \tag{3}$$

which is a relevant bias from $1/c(t)$ when birth and death rates are similar. Thus, in order to infer the frequency of each of the M_α antigens, we need to know the tumor population when they were presented, and so their time of appearance. Each antigen appears when $M_\alpha = i$ is a positive integer, which happens for

* corresponding author
† corresponding author

$$t(\alpha_i) = \frac{1}{(b-d)} \log\left(\frac{(b-d)i}{2\mu p_\alpha c_0 b} + 1\right) \tag{4}$$

from where we infer that the population present when the i-th antigen is presented is

$$c\left(t(\alpha_i)\right) = c_0 e^{(b-d)t(\alpha_i)} = c_0 + \frac{(b-d)i}{2\mu p_\alpha b} \tag{5}$$

Both analytical estimates are consistent with simulations of exponentially growing tumors with random mutations (Fig. S1). However, an analytical sum of each antigen frequency (Eq. 9) using the population at appearance (Eq. 11) is not feasible. We still can, in the early tumor growth scenario, consider a low death-to-birth ratio and approximate the frequency of a mutation appearing at time t_i by $1/c(t_i)$ [1]. The deviation from the analytical estimate under this assumption is small (see Fig. 4 in the main text). The average antigen clonality results from adding the inverse of each population at antigen appearance (Eq. 11) for all present antigens at time t. This results in a decaying average antigen frequency

$$\langle\gamma(t)\rangle = \frac{1}{M_\alpha(t)} \sum_{i=1}^{M_\alpha(t)} \frac{1}{c_0 + \frac{(b-d)i}{2\mu p_\alpha b}} = \frac{1}{c_0\left(e^{(b-d)t} - 1\right)} \left[\psi^{(0)}\left(\theta e^{(b-d)t} + 1\right) - \psi^{(0)}\left(\theta + 1\right)\right] \tag{6}$$

where $\theta = 2\mu p_\alpha b c_0/(b-d)$ and $\psi^{(0)(x)}$ is the digamma function [8]. This result can be compared to simulations of several growth dynamics (Fig. 4 in the main text) and is a first analytical measure of the evolutionary increases of heterogeneity in the neoantigen landscape. The role of evolving instability levels μ is again constrained as $\langle\gamma(t)\rangle$ saturates rapidly for high values of μ observed experimentally [6]. A larger μ will ensure that antigens appear more frequently, producing early antigens that will be present across a larger tumor fraction, consistent with evidence on the role of mutational load in immune prognosis [9-12]. Together with this, an increase in the rate of mutation will also increase the elements in the sum of Eq. (9) each one smaller than the earlier. Together with dominance saturating at high neoantigen production, it appears that highly mutational tumors might initially be easier to target by the immune system, but eventual saturating effects indicate that further increases in μ will not be likely to enhance the immune response.

Another key element in therapy prognosis is the fraction of the tumor populated by cells that do not present antigens. To study how this fraction evolves as tumors grow, we model the dynamics of antigen-cold cells (c_{α^-}). These cells replicate correctly at rate $r(1-\mu p_\alpha)$, with $r = b - d$, and any mutation presenting antigens can happen at rate $r\mu p_\alpha$. The mean field dynamics are simply

$$\frac{dc_{\alpha^-}}{dt} = r(1-\mu p_\alpha)c_{\alpha^-} - r\mu p_\alpha c_{\alpha^-} \tag{7}$$

with solution

$$c_{\alpha^-}(t) = c_{\alpha^-}(0)e^{(r(1-2\mu p_\alpha)t)}. \tag{8}$$

The fraction of antigen-cold cells in the tumor is therefore

$$\frac{c_{\alpha^-}(t)}{c(t)} = \frac{c_{\alpha^-}(0)e^{(r(1-2\mu p_\alpha)t)}}{c_0 e^{(b-d)t}} = e^{-2r\mu p_\alpha t}. \tag{9}$$

This result is consistent with simulations (Fig. S3) and indicates that the relevance of neoantigen-cold cells in unstable tumors decays fastly. This is a mathematical support for the notion that it is excessive neoantigen heterogeneity and not lack of neoantigen presentation what could drive checkpoint blockade immunotherapy failure. This result is independent of other therapeutic drugs such as several targeted agents that are neutral to antigen load. Even in the combination therapy design of the original article, the fraction of antigen-hot cells in the tumor will increase exponentially until immunotherapy is administered.

B. Tumor size in the neoantigen diversity threshold

Summing the equations for all the subclonal populations in the tumor results in a condition for tumor control, Namely the following inequality:

$$\sum_i^S \frac{dc_i}{dt} = \sum_i^S c_i \left[\langle r\rangle \left(1 - b\sum_j^S c_j\right) - \frac{k\rho\alpha_i s}{g + \sum_j^S c_j}\langle\gamma\rangle\right] < 0 \tag{10}$$

A minimal condition necessary to fullfil growth control is that the sum of replication terms is negative

$$\sum_i^S \left[\langle r \rangle \left(1 - b \sum_j^S c_j \right) - \frac{k\rho\alpha_i s}{g + \sum_j^S c_j} \langle \gamma \rangle \right] < 0 \tag{11}$$

which results in the threshold condition as stated in the main text

$$\frac{\langle r \rangle}{k\rho s} \left(1 - b \sum_j^S c_j \right) \left(g + \sum_j^S c_j \right) < \langle \gamma \rangle \frac{\sum_i^S \alpha_i}{S} \tag{12}$$

The left hand side is the replication/predation ratio, that has to be larger than the tumor immunogeneicity term of the right hand side. Furthermore, the product of the logistic and Michaelis-Menten terms in the lhs defines to which extent the spatial constraints affect more tumor growth or else immune circulation.

Even if we are interested in the specific scenario of a small growing tumor, where spatial constraints can be neglected, we can discuss the overall role of non-linear spatial interactions in the threshold condition. The carrying capacity of a tumor is largely dependent on specific conditions and general conclusions cannot be taken for several tumor types or patient characteristics. It can be seen (Fig. S1) that, for tumors where carrying capacity is large, there will be a region where tumor growth exceeds immune predation. However, this happens for tumor sizes beyond the small populations under study in the main text. Furthermore, the article focuses on the rhs of equation 12 as a possible immunotherapy prognosis marker, leaving aside for the moment the exact computation of the parameters in the lhs that should incorporate many layers of quantiative estimates.

II. COMPUTER SIMULATIONS

A. Antigen dominance as a function of neoantigen load

Each neoantigen has a different capacity to elicit an immune response. To asses how the amount of neoantigens correlates with subclonal immunogeneicity, we build a minimal stochastic simulation based on an existing database. From [3] we obtain the repertoire of all detected neoantigens of three databases [9-11] and their estimated immunogeneicity, which is found to be very skewed, consistent with the concept of *immunodominance*: most analyzed antigens are of very low immunogeneicity while only a few elicit a considerable immune response (data not shown).

During immune evasion, tumor mutations that produce antigens evolve neutrally as no there is no lymphocyte attack. We build a computational framework in `c` to simulate the production of the antigens obtained from the Luksza et al. database. In the simulations, a single agent representing a cellular metapopulation acquires random mutations that select, one by one, random antigens from the database. At each mutation, we update the recognition capacity to that of the novel epitope if it is the most dominant antigen at place. This approach generates a first approximation of how average maximum dominance D increases as more neoantigens are being produced.

Many neoantigens have to be produced in order to find, on average, highly immunogenic ones. Simulations indicate that, for common subclones with about 1 to 200 antigens [3], recognition potential of the most dominant antigen scales linearly with the total neoantigens of the subclone, here termed α_i (Fig. 2a in the main text). For larger antigen loads maximum immunogeneicity in simulations saturates, as the given epitope database is finite and no novel surface proteins can be presented. This might be only due to limitations in sequencing depth [13,14], and it is probable that a complete repertoire is so large that saturation happens at even higher neoantigen loads.

B. Immune search efficiency as a function of neoantigen clonality

We also present a computational approach to explicitly determine the dependency of immune search efficiency on antigen clonality. The model is based on the fact that, even if the immune response has many layers of complexity, only two search processes are directly related with the amount and distribution of neoantigens on the tumor surface: the motion of dendritic cells to find tumor antigens and the search of effector T cells that have migrated to the tumor site once activated.

The computational framework in `python` includes a two-dimensional grid where antigens are distributed, representing an immunogenic tumor surface similar to that of CITE Macfarlane, but where specific collision between cellular agents are here modeled as searches of non-moving antigens. Antigens are therefore characterized by the fraction of

the cells they are present in, γ_i (Figure 3a in the main text). The immune response is simulated by dendritic cells that search for antigens following Lévy migration statistics. The length l of the Lévy flight with parameter μ is generated from [15] as

$$l_\mu = \frac{\sin\left((\mu-1)X\right)}{(\cos X)^{1/(\mu-1)}}\left(\frac{\cos\left((2-\mu)X\right)}{Y}\right)^{(2-\mu)/(\mu-1)} \tag{13}$$

where X is a uniform random variable on $[-\pi/2, \pi/2]$ and Y has a unit exponential distribution computed as $Y = -\ln Z$, with Z being a uniform random variable on $[0, 1]$.

Activated T cells undergo Brownian searches to find the same specific antigen at the tumor site [18]. With this, our minimal approach minimizes the number of parameters playing a role on final results to only antigen frequency γ and the Lévy exponent of inactive immune cells $\mu = 2.15$, that encapsulates the underlying effect of chemokine distributions on immune cell walks [15]. As originaly stated in [15], larger Lévy parameter values could result in less efficient search dynamics. Consistently, we find that larger values for μ result in longer search times, that result in a change of s, the steepness of the $E \sim s\gamma$ relation, but maintain the qualitative linear shape resulting from our approach (Fig. S4)

Both search processes happen in a lattice of 10μm squares with timesteps of $\delta t = 1$min, according to the estimated speed of immune cell movement ($10\pm5\mu$m min^{-1} [16]. We map the efficiency of the search (the inverse of average time until the T-α_i encounter) as a function of the antigen density γ_i (Figure 3b,c in the main text).

We assume that molecular processes that happen in the lymph node, as well as activation or removal rates, are not a direct function of antigen concentration [17]. They will not explicitly affect the neoantigen-response dynamics that our model aims at capturing and will only add up to a constant rate. In the model this can be encapsulated in the search constant s, which is not a function of antigen frequency γ_i. Several rates for these processes have been estimated in previous research, such as dendritic cell activation by antigens (0.07 cells min^{-1} [18]), Dendritic-T cell antigen presentation (0.12 cells min^{-1} [19]) or cancer cell apoptosis activation by T cells (0.038 cells min^{-1} [20]). All the molecular dynamics involved in these processes have a timescale of about $10^0 \sim 10^1$min. The timescale of our modeled search process is of up to 10^3min for subclonal antigens of small concentration (Fig. SX!!!), corresponding with the notion that it might be a rate-limiting step in the overall immune response.

Furthermore, a rough estimate for the characteristic timescale of antigen emergence results in

$$\tau_\alpha = \mu p_\alpha n r \approx 2x10^{-4} \text{ antigens cell day}^{-1} \tag{14}$$

where $\mu p_\alpha (\approx 10^-6$ is the rate of antigenic mutations per gene per division ($\mu \approx 10^{-5}$ for mutator phenotype cancers [4,6]), $n \approx 2\text{x}10^{-4}$ the number of genes in the genome and $r \approx 10^{-2}$ the replication rate per day. Therefore, even for highly mutable cells, the rate of antigen presentation is of a much slower timescale compared to the rest of the process, so that we consider antigens to be static during the immune search process. Long-term evolutionary dynamics of immune evasion or epigenetic silencing are discussed in the main text and not yet incuded in this work.

C. Evolution of Neoantigen Distributions and Combination Therapy

Computer simulations are built to understand how neoantigen load and heterogeneity evolve in the scenarios of tumor growth and cytotoxic therapy. The model consists in a birth-death Gillespie `python` simulation for the exponential growth of cancer cells, that replicate at rate b and die at rate d, with $r = b - d = 0.3\text{day}^{-1}$. At each replication event, both daughter cells can gain a set of mutations taken from a Poisson distribution with parameter $\mu = 1$. We here consider that further mutations in DNA repair mechanisms that alter either μ or the shape of the Poisson distribution would speed up the overall neoantigen production process, which eventually results in a faster pace of increased immunogeneicity. Specific dynamics for the evolution of mutational rates as in [21] have not yet been included. From the set of all mutations, $p_\alpha = 0.13$ are considered to produce a neoantigen [22]. In our simulations, each antigen is associated to the mutated cell and labeled with a natural number $a_1, a_2, a_3, a_4, ...$ as in figure 1b. We record the number of generated and present antigens and the frequency of each in the tumor by using the `Counter` tool from the `collections python` module. This last measure can be used to obtain the evolution of average antigen frequency along cancer growth.

To understand the effect of a generalistic targeted therapy, such as BRAF or MEK inhibitors in neoantigen distributions, we introduce a drug-resistance mutation acquired with a smaller probability $p_R = 10^{-4}$. This probability might be much smaller for given molecular domains, but we use a value large enough to see the appearance of a resistant subclone in a computable timespan. Once therapy is administered, sensitive cells die at rate $d_S = 0.9\text{day}^{-1}$, while resistant cells die at rate $d_R = 0.01\text{day}^{-1}$, much smaller. This values only aim at generating a qualitative image of the selective sweep for the resistant subclone. The introduction of these dynamics results in the reduction of

tumor size followed by the growth of a resistant subclone. Drug resistance is often present before targeted therapy is administered [23], together with experimental evidence for neoantigen evolution continuing after therapy [24]. This results in changes in the heterogeneity of the epitope landscape [24,25]. We follow the dynamics of neoantigens during therapy to understand the effects of resistance on tumor and epitope heterogeneity.

Due to the computational cost of simulating a tumor population of realistic size, the simulations study a toy model of a small tumor growth, where growth and probability of resistance mutations are scaled to the novel size. As previously discussed, the first stages of tumor growth are those where neoantigens accumulate faster and heterogeneity increases the most. We study the dynamics of these in the event of therapy resistance.

III. DATA ANALYSIS

A. Measures of subclonal neoantigen dominance

To understand the effect of neoantigen load on subclonal immunogeneicity, we designed a minimal stochastic model that indicated the possible existence of a linear trend $D \sim \alpha_i$ (Suppl. II.A, Fig. 2a in the main text). We use the data from [3] to study the consistency of this result.

Data from [3] contains the list and predicted immunogeneicity of the antigenic mutations of each subclone in a dataset of 198 patients from 3 different publications. For each subclone we store the number of antigens (α_i) and the estimated dominance of the subclone (D_i), that authors find it corresponds the immunogeneicity of the most dominant antigen. We plot both values for all subclones in all patients to study if there is a correlation. In particular, a linear trend can be obtained by studying the data from [3] (Fig. 2b in the main text), but results are very scattered and further experimental data should be collected to asses the confidence of our modeling approach of $D(\alpha_i)$.

B. Survival ratios for different biomarkers

Data from [3] contains an organized neoantigen phylogeny for each tumor in anti-CTLA-4 treated melanoma [10,11] and anti-PDL-1 treated lung cancer [9] datasets. This subclonal phylogeny makes it straightforward to compute, for each patient, the number of identified subclones S and the total neoantigens of that tumor, $\sum_i \alpha_i$. Furthermore, since neoantigens are labeled, we study the presence of a given neoantigen across subclones, because not all neoantigens are totally public or private and some are shared among specific subclones. The size of each subclone is also available in the original dataset. This allows us to compute γ_i and $\langle\gamma\rangle$, the frequency of each antigen and the average neoantigen frequency for each tumor.

For each patient, we add to the survival data from [3], both the load $\sum_i \alpha_i$ and the threshold $\langle\gamma\rangle \sum_i \alpha_i/S$ biomarkers. This allows for a study of how these markers predict therapy prognosis (Figures 5 in the main text and S6). Survival curves are obtained with `KaplanMeierFitter` from the `lifelines` python module. Log-rank tests are performed with `logrank_test` from the `lifelines.statistics` python module. Cumulative hazar ratios are computed with `NelsonAalenFitter` from the `lifelines` python module.

References

1. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. Nat Gen 2016;48:238–44.
2. Matzavinos A, Chaplain MA, Kuznetsov VA. Mathematical modelling of the spatioâtemporal response of cytotoxic Tâlymphocytes to a solid tumour. Mathematical Medicine and Biology 2004, 21(1), 1–34.
3. Luksza M, Riaz N, Makarov V, Balachandran VP, Hellmann MD, Solovyov A, et al. A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. Nature 2017;551:517–20.
4. Campbell BB, Light N, Fabrizio D, Zatzman M, Fuligni F, de Borja R., et al. Comprehensive analysis of hypermutation in human cancer. Cell 2017, 171(5), 1042–1056.
5. Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability–an evolving hallmark of cancer. Nat Rev Mol Cell Biol 2010;11:220–8.
6. Loeb, LA, Kohrn BF, Loubet-Senear KJ, Dunn YJ, Ahn EH, OâSullivan JN, et al. Extensive subclonal mutational diversity in human colorectal cancer and its significance. Proceedings of the National Academy of Sciences 2019, 116(52), 26863–26872.
7. Bozic I, Gerold JM, Nowak MA. Quantifying Clonal and Subclonal Passenger Mutations in Cancer Evolution. PLOS Comp Biol 2016;12(2):1–19.
8. Abramowitz M, Stegun IA (eds). "6.3 psi (Digamma) Function." Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. New York: Dover 1972:258–9.

9. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Mutational landscape determines sensitivity to PD-1 blockade in nonâsmall cell lung cancer. Science 2015;348:124–8.
10. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med 2014;371:2189–99.
11. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, et al. Genomic correlates of response to CTLA4 blockade in metastatic melanoma. Science 2015;aad0095.
12. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science 2015;348:69–74.
13. Stead LF, Sutton KM, Taylor GR, Quirke P, Rabitts P. Accurately Identifying LowâAllelic Fraction Variants in Single Samples with NextâGeneration Sequencing: Applications in Tumor Subclone Resolution. Hum Mutat 2013;34:1432–8.
14. Yadav M, Jhunjhunwala S, Phung QT, Lupardus P, Tanguay J, Bumbaca S, Franci C, et al. Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. Nature 2014;515:572–6.
15. Harris TH, Banigan EJ, Christian DA, Konradt C, Tait Wojno ED, Norose K, et al. Generalized Lévy walks and the role of chemokines in migration of effector CD8+ T cells. Nature 2012; 486(7404):545–48.
16. Boissonnas A, Fetler L, Zeelenberg IS, Hugues S, Amigorena S. In vivo imaging of cytotoxic T cell infiltration and elimination of a solid tumor. The Journal of experimental medicine 2007, 204(2), 345–356.
17. Krummel MF, Bartumeus F, Gérard A. T cell migration, search strategies and mechanisms. Nat Rev Immunol 2016, 16(3), 193.
18. Bianca C, Chiacchio F, Pappalardo F, Pennisi M. Mathematical modeling of the immune system recognition to mammary carcinoma antigen. In BMC bioinformatics 2012. (Vol. 13, No. S17, p. S21). BioMed Central.
19. Engelhardt JJ, Boldajipour B, Beemiller P, Pandurangi P, Sorensen C, Werb Z, Egeblad M, Krummel MF (2012) Marginating dendritic cells of the tumor microenvironment cross-present tumor antigens and stably engage tumor-specific T cells. Cancer Cell 21(3):402â417
20. Christophe C, MÃ¼ller S, Rodrigues M, Petit AE, Cattiaux P, DuprÃ© L, Gadat S, Valitutti S (2015) A biased competition theory of cytotoxic T lymphocyte interaction with tumor nodules. PLoS ONE 10(3):e0120,053
21. Aguadé–Gorgorió G, Solé R. 2018 Adaptive dynamics of unstable cancer populations: The canonical equation. Evolutionary applications, 11(8), 1283-1292.
22. Efremova M, Finotello F, Rieder D, Trajanoski Z. (2017). Neoantigens generated by individual mutations and their role in cancer immunity and immunotherapy. Frontiers in immunology, 8, 1679.
23. Diaz Jr LA, Williams RT, Wu J, Kinde I, Hecht JR, Berlin J, et al. The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature 2012, 486(7404), 537–540.
24. Anagnostou V, Smith KN, Forde PM, Niknafs N, Bhattacharya R, White J, et al. Evolution of neoantigen landscape during immune checkpoint blockade in nonâsmall cell lung cancer. Cancer discovery 2017, 7(3), 264–276.
25. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and acquired resistance to cancer immunotherapy. Cell 2017;168:707–23.
26. Kuznetsov VA, Makalkin IA, Taylor MA, Perelson AS. Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. Bulletin of mathematical biology 1994, 56(2), 295–321.
27. Aguadé-Gorgorió G, Solé R. Genetic instability as a driver for immune surveillance. Journal for ImmunoTherapy of Cancer 2019, 7(1), 1–13.
28. Gerlee P, Anderson AR. The evolution of carrying capacity in constrained and expanding tumour cell populations. Physical biology 2015, 12(5), 056001.

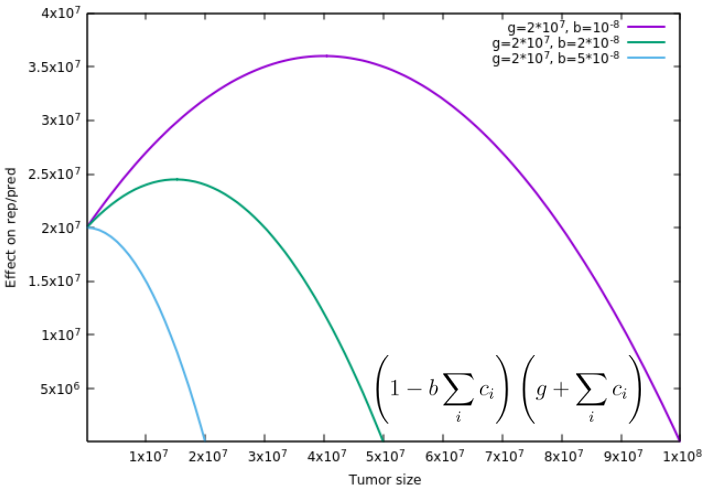

FIG. S1. The role of spatial constraints on the growth/predation ratio of the neoantigen threshold. A value for g is estimated in [26]. However, we know that carrying capacities depend on tumor type and microenvironment constraints, and also evolve in time [28]. In cancers where the carrying capacity is large enough (green and purple curves), the increase in tumor size results in a region where growth exceeds immune predation.

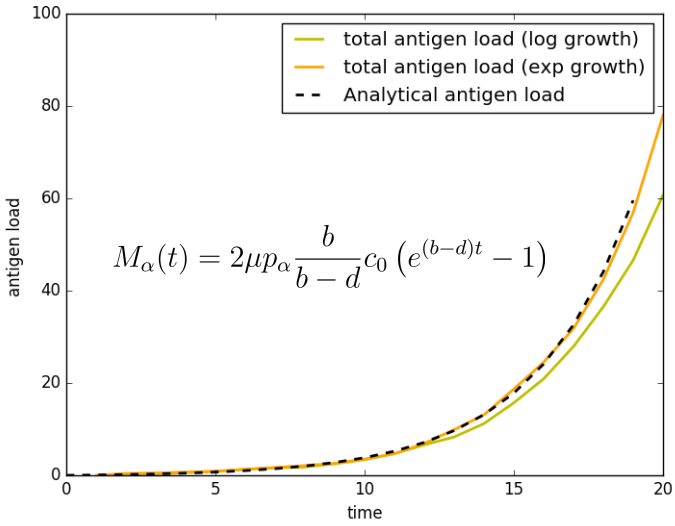

FIG. S2. Simulations and analytical estimates for the accumulation of antigens in exponentially and logistically growing tumors. Production of neoantigens slows down in logistic growth, but deviations from our estimate only start being representative at carrying capacities of $K = 2 * 10^5$, much smaller than those estimated in previous publications of $K = 10^7 \sim 10^9$ [26,27].

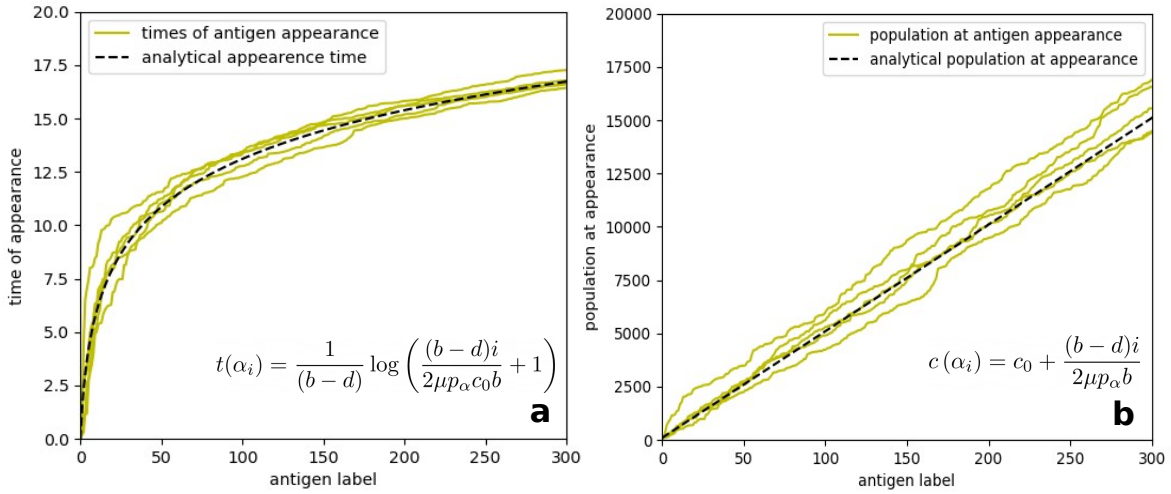

FIG. S3. Simulations and analytical estimates for the average time **(a)** and total tumor population **(b)** at the appearance of each novel antigen presented in an exponentially growing tumor.

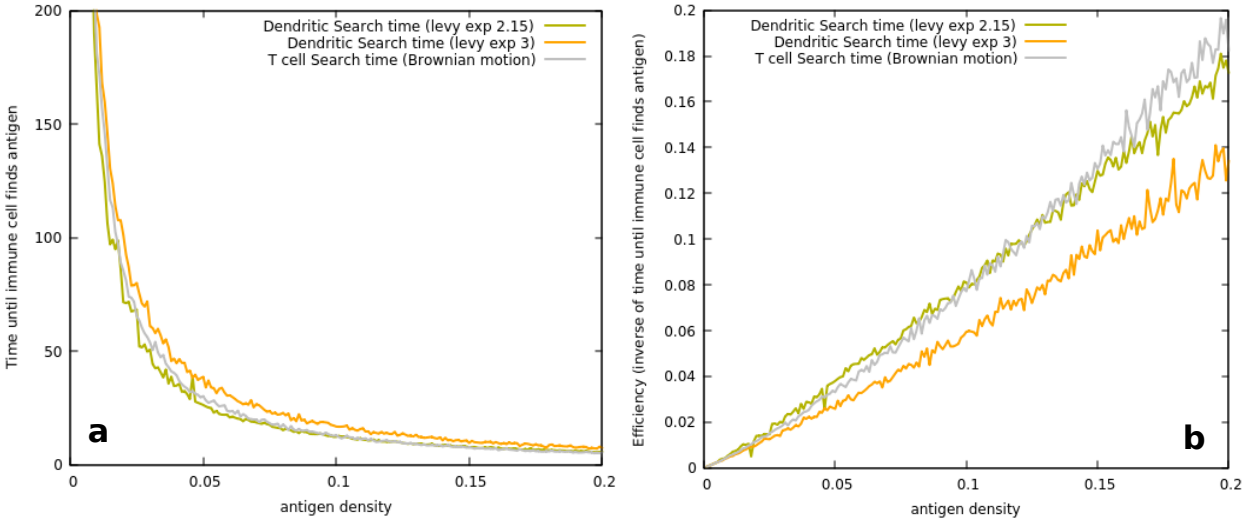

FIG. S4. Time distribution and Search efficiency (inverse of time until immune cell-antigen encounter) under different search processes. Immune agents search for antigens distributed with densities between 0.001 and 0.2, according to realistic antigen frequencies across tumor biopsies [3]. Brownian walks and alterations from the estimated Lévy flight parameter of T cells [15] result in diminished search efficiency, reducing s, but maintain the linear relation between antigen frequency and the efficiency of the search process.

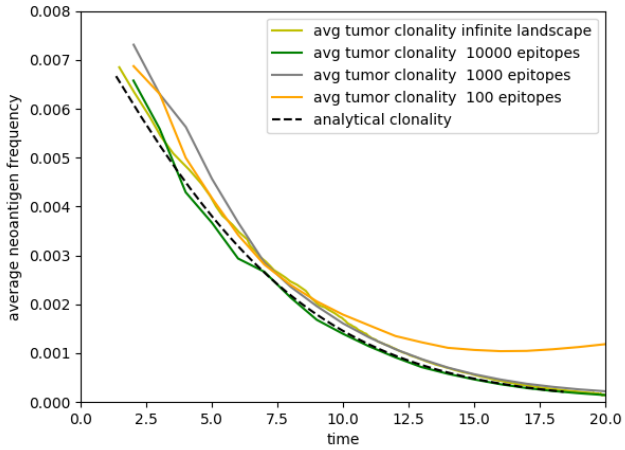

FIG. S5. Simulations and analytical estimates for the decay of average antigen frequency in the event of high mutational loads that make for a finite epitope landscape. Recent experimental evidence indicates that high mutational loads result in equivalent mutations occuring in different cells independently [6]. We compare the decay of antigen frequency in the assumption that mutations in the same loci could produce the same antigen in different cells. Deviance from our analytical results is only significant for unrealistically small genomes (timescale in days).

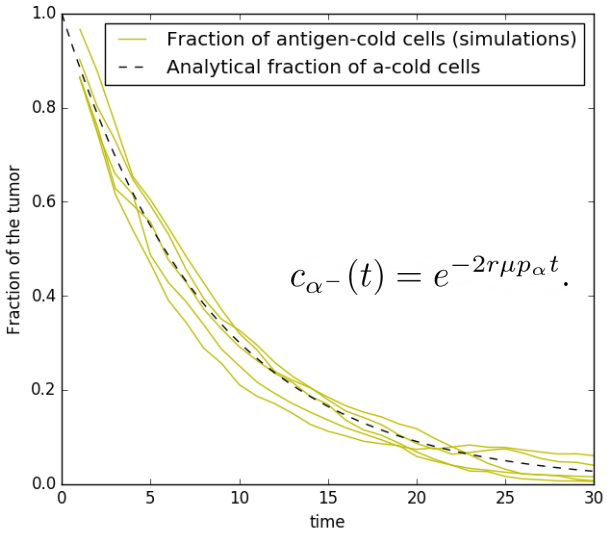

FIG. S6. Simulations and analytical estimates for the fraction of antigen-cold cells in an exponentially growing tumor. Results are unaltered by a targeted therapy that is neutral to antigen load.

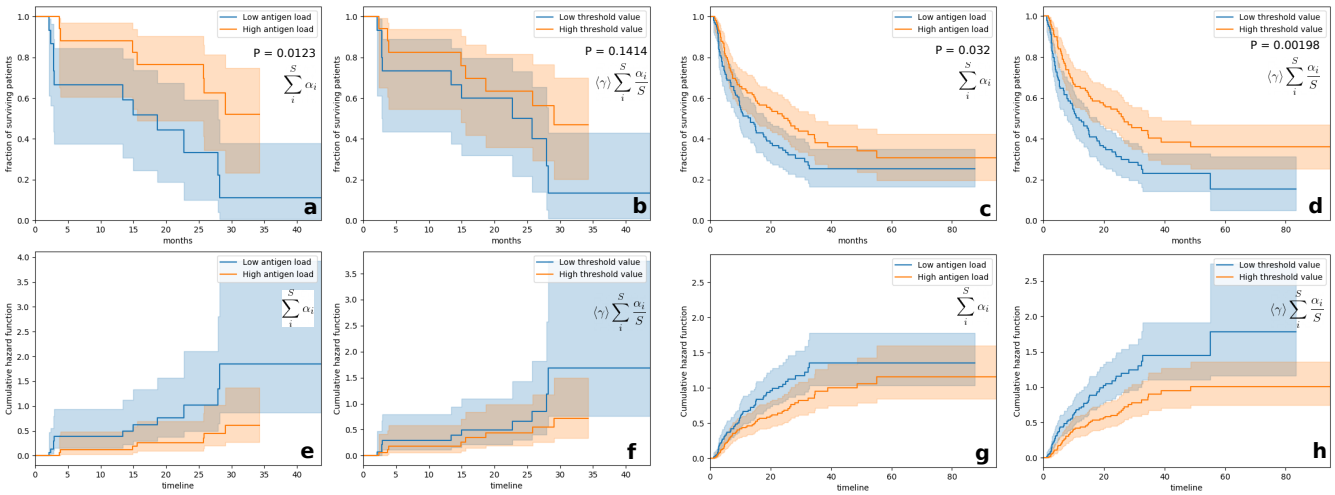

FIG. S7 Correlation of patient biomarkers with months of survival after checkpoint inhibition therapy. The Kaplan-Meier curves for decay of survival probability are depicted, with 30 anti-PDL-1treated Lung cancer patients from [9] separated in two groups by the median of either neoantigen load (**a**, log-rank test P value 0.0123) and our threshold value $\gamma \sum \alpha/S$ (**b**, log-rank test P value 0.1414). Cumulative hazard ratios (**e,f**) also support that including average neoantigen clonality does not add predictive capacity to the well-accepted neoantigen load marker. This suggests $\gamma \sum \alpha/S$ might be only a complementary biomarker for immunotherapy prognosis in Melanoma (see main text results). Including the cohort of 30 anti-PDL-1 treated Lung cancer patients from [9] does not improve overall results **(c,d,g,h)**. This might be also because of the limited size of the lung cancer cohort.

Journal of Theoretical Biology 511 (2021) 110552

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

The ecology of cancer differentiation therapy

Ricard Solé [a,b,c], Guim Aguadé-Gorgorió [a,b]

[a] ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Catalonia, Spain
[b] Institut de Biologia Evolutiva (CSIC-UPF), Psg Maritim Barceloneta, 37, 08003 Barcelona, Catalonia, Spain
[c] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501, USA

ARTICLE INFO

Article history:
Received 31 January 2020
Revised 21 October 2020
Accepted 17 November 2020
Available online 10 December 2020

Keywords:
Cancer
Evolution
Ecology
Differentiation therapy
Habitat loss
Metapopulations
Cancer stem cells

ABSTRACT

A promising, yet still under development approach to cancer treatment is based on the idea of differentiation therapy (DTH). Most tumours are characterized by poorly differentiated cell populations exhibiting a marked loss of traits associated to communication and tissue homeostasis. DTH has been suggested as an alternative (or complement) to cytotoxic-based approaches, and has proven successful in some specific types of cancer such as acute promyelocytic leukemia (APL). While novel drugs favouring the activation of differentiation therapies are being tested, several open problems emerge in relation to its effectiveness on solid tumors. Here we present a mathematical framework to DTH based on a well-known ecological model used to describe habitat loss. The models presented here account for some of the observed clinical and in vitro outcomes of DTH, providing relevant insight into potential therapy design. Furthermore, the same ecological approach is tested in a hierarchical model that accounts for cancer stem cells, highlighting the role of niche specificity in CSC therapy resistance. We show that the lessons learnt from metapopulation ecology can help guide future developments and potential difficulties of DTH.

1. Introduction

Cancer is a set of complex diseases, and the success of tumor progression (and the eventual death of its recipient organism) requires a number of changes to make cells capable of overcoming selection barriers. These changes provide the source of proliferative power that makes tumors able to expand and evolve (Weinberg, 2014). One particularly remarkable feature of cancer cells is the loss of molecular markers associated to the differentiated state. As the tumor evolves, some cancer cells appear to be in a de-differentiated state closer to early developmental stages, similar to that of normal stem cells, with increased potential for self-renewal and plasticity (Magee et al., 2012). To some extent, cancer is a disease of multicellularity: the cooperative order required to maintain organism's coherence is broken in favor of unicellular-like traits (Aktipis et al., 2015; Davies and Lineweaver, 2011).

The standard treatment of tumors has been grounded in the use of either specific cytotoxic drugs or radiotherapy, or a combination of both. The success of this approach has been discussed and even questioned over the last decades (Gatenby, 2009). Treatments involving a general mechanism of cell damage associated to toxicity are often inefficient and can trigger evolutionary pressures that select aggressive and resistant clones (Pepper et al., 2009). As a consequence, cytotoxic therapies can create undesirable side effects such as the development of metastasis. To a large extent, despite the undeniable success in our increasing understanding of the underlying molecular basis, cancer remains incurable. Because of these limitations, novel approximations have been proposed mainly from evolutionary and mathematical biology. They are based on the view of cancer as an ecological and evolutionary problem (Merlo et al., 2006; Korolev et al., 2014). In particular, ecological principles can guide alternative insights to cancer development and treatment (Basanta and Anderson, 2013).

One specially promising alternative to conventional cytotoxic agents is the use of so called *differentiation therapy*. Here the approach, early suggested more than 50 years ago (Pierce and Wallace, 1971; Pierce, 1983) is inspired in the observation that one hallmark of cancer is the loss or blocking of differentiation that leads to cells with increased potential for self-renewal and plasticity. Differentiation therapy (DTH) involves the use of diverse molecular agents able to induce differentiation in cancer cells. Since differentiated cell types are a terminal branch of development, the goal is to facilitate this process and remove cancer cells from the proliferative compartment. A growing family of DTH agents include neural growth factors, all trans retinoic acid, arsenic

https://doi.org/10.1016/j.jtbi.2020.110552
0022-5193/

trioxide, butyric acid or cAMP, which have been shown some degree of differentiation-inducing capability both in vitro and/or in vivo (de Thé, 2018). The success of DTH is well illustrated by the best known case study, namely its use in Acute Promyelocytic Leukemia (APL) by means of a combined cytotoxic therapy with all-trans retinoic acid (RA) (Huang et al., 1988).

A few numbers reveal some features of the impact of DTH. Again within the context of APL, before the use of DTH, cytotoxic-based therapies increased the likelihood of remission from 50 to 80% but with only a third of long-term survival. The combination with RA changed drastically the situation, with 90% remission and a 75% cure (see Cruz and Matushansky, 2012 and references therein). Interestingly, when DTH alone is used, despite cell differentiation perfectly well identified (it can actually be massive) only combination with standard cytotoxic agents seems to account for long-term disease remission (de Thé, 2018).

Over the last years, DTH agents have been also used for treating solid tumors. In contrast with the APL case study, the therapeutic effect of the differentiation-inducing agents on solid tumors is not strong when compared with that of conventional chemotherapeutic agents. However, because most of the differentiation-inducing agents can potentiate the effect of conventional chemotherapy or radiation therapy, combination therapy might be used as a second- or third-line therapy in patients with advanced cancer. Are the solid nature of the tumors, their genetic complexity or their hierarchical architecture leading factors for this limited success? Here too, a theoretical model can be helpful in interpreting the role of spatial competition effects and niche specificity in understanding the possibilities of DTH. The analysis of how differentiation therapy modulates the cancer habitat is based in an ecological approach to tumor dynamics inspired in well-established results from habitat loss and fragmentation in metapopulations (Moilanen and Hanski, 1998; Hanski, 1999).

2. Metapopulation model of tumor differentiation therapy

The simplest mathematical approach taken here is based on the assumption that two different therapies act together on the growth of a cancer cell population. As a first dynamical description able to characterize a wide set of disease types we use a generic logistic growth model (Gatenby and Gillies, 2008), where cellular replication saturates as the tumor population approaches the carrying capacity K of the micro-environment. Sigmoidal growth has been proven to capture the tumor microenvironment effects of spatial constraints, resource limitations or intercellular growth inhibition in a wide range of malignancies such as breast cancer (Norton, 1988), colorectal cancer (Misale et al., 2015) or chronic lymphocytic leukemia (Gruber et al., 2019) The cancer cell population is also inhibited in two different ways. The first corresponds to standard therapies, based on cancer-targeted cytotoxic drugs. In this first scenario, a population of cancer cells C follows:

$$\frac{dC}{dt} = rC\left(1 - \frac{C}{K}\right) - \delta C \tag{1}$$

where for simplicity the carrying capacity will be normalized to one ($K = 1$) and thus C can be understood in terms of the fraction of host tissue occupied by the tumor. The last term in the rhs indicates the linear decay caused following from cellular death rate δ that can be increased under chemotherapy. This is equivalent to the well-known Levins model, where growth and decay would be related to colonization and extinction (Levins, 1969). The analysis of this system reveals that two equilibrium states C^* are possible: extinction $C^* = 0$ and $C_1^* = 1 - \delta/r$. Tumor growth will occur when $r > \delta$, i. e. if growth overcomes the negative impact of treatment.

How can differentiation treatment be introduced in this approach? The impact of DTH is dynamically very different. Previous research has studied mathematical modeling of tissue hierarchies by introducing differentiation as a rate at which progenitor cells transition into non-cycling phenotypes (Dingli and Michor, 2006; Werner et al., 2016). However, before getting into this multi-layer approach (Fig. 1b), we aim at understanding a specific effect of DTH that has not been included in previous models: what is the effect of non-cycling, differentiated cells as they occupy the habitat that the tumor needs for expansion? (Fig. 1b, dashed box). Lessons from habitat fragmentation have shown that these effects might be potentially key in extinction of colonizing species (Moilanen and Hanski, 1998; Hanski, 1999). In this context, the DTH scenario discussed below is inspired in the nonlinear behavior how fragmented landscapes (fig. 1c). Reductions of habitat size, along with stochastic death, define viability thresholds that we will connect with tumor extinction dynamics. In particular, these models reveal a somewhat counterintuitive feature that is relevant to our paper: a reduction of habitat does not simply imply a similar population reduction. Instead ad, extinction can occur despite the existence of a remaining habitat.

As a fraction of cancer cells gets differentiated, they have an impact in population dynamics as they contribute to the overall population and thus limit the potential carrying capacity of the system. Studies on control networks indicate that not only spatial or resource constraints but also cellular signaling contributes to

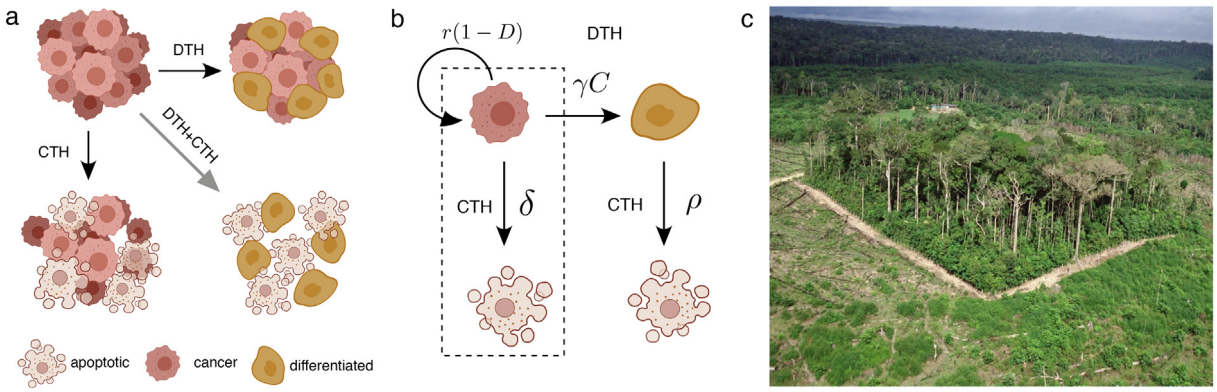

Fig. 1. A metapopulation model of tumor differentiation therapy. In tackling alternative treatments to cancer progression, DTH exploits the potential of blocking tumor growth by activating differentiation pathways. A combination of cytotoxic therapy (CTH) and differentiation therapy (DTH) as shown in **(a)** can successfully kill the tumor when Both separately cannot. In **(b)**, both the complete model including the dynamics of the differentiated compartment and the minimal habitat-loss approach (dashed box) is shown. The DTH + CTH model is inspired in studies of habitat fragmentation **(c)**, where habitat reduction, along with stochastic mortality, can trigger species extinction (image courtesy of Mark Moffett). Drawings were made with Biorender.

these dynamics (Vainstein et al., 2012; Yang et al, 2017; Komarova and van den Driessche, 2018). If D weights the effectiveness of the DTH the simplest extension of the previous model incorporates the amount of

$$\frac{dC}{dt} = rC(1 - D - C) - \delta C \tag{2}$$

The ecological equivalent here is the extended Levins model incorporating habitat loss (Bascompte and Solé, 1996, Fig. 1b, red box). In habitat loss models, the D term is associated to the amount of habitat that has been degraded thus being unavailable to colonization. The interpretation within the context of DTH is easy: the fraction of cancer cells that have become differentiated introduce a shift $1 \rightarrow 1 - D$ in the maximum available habitat.. The non-trivial fixed point is now:

$$C_1^*(D, \delta) = 1 - D - \frac{\delta}{r} \tag{3}$$

In this case, cancer decay will be expected provided that the fraction of differentiated habitat (and thus the efficiency of DTH) is larger than a critical value:

$$D > D_c = 1 - \frac{\delta}{r} \tag{4}$$

One particularly relevant and non-obvious result is that even if there's apparently room for further growth, the dynamics of the system reveal a transition from cancer growth to cancer decay. Once the critical point D_c is reached, tumor dynamics faces extinction.

In Fig. 2a we show a diagram for D against δ where the critical line $D = D_c$ has been used to separate the two phases associated to cancer progression and cancer decay. The lower axis indicates the efficiency of single cytotoxic therapy in the absence of DTH. A threshold is found for $\delta_C = r$ as defined from model (1). By adding the second axis (differentiation) we can see that lower levels of chemotherapy are required to achieve tumor decay. This is at the core of our explanation for the success of DTH: the combination of both treatments can successfully achieve remission when the right combination of chemotherapy and differentiation is used. Since toxicity can be reduced provided that D is large enough, the diagram supports the observed success and long-term remission in APL. On the other hand, the levels of differentiation that are required for small δ can be very large (perhaps unrealistically large). An important point needs to be made here: could DTH only also achieve remission? The model in this case reads:

$$\frac{dC}{dt} = rC(1 - D - C) \tag{5}$$

which can be shown to behave always in the same way: a logistic growth towards an intermediate level $C^* = 1 - D$ with no threshold value. This implies, and seems consistent with clinical evidence, that DTH alone will fail to succeed given the lack of a remission threshold.

We have used a specific functional form for population growth, and the previous analysis deals with steady states involving large populations. However, our results are robust, as shown by looking at early tumor growth in a differentiated environment, where $C \ll 1 - D$, provides interesting differences between the dynamical impacts of cytotoxic therapy and the possibilities of modifying environmental carrying capacity through DTH. The model described by Eq. (2) now reads

$$\left(\frac{dC}{dt}\right)_{C \ll 1-D} \sim [r(1 - D) - \delta]C \tag{6}$$

with exponential growth solution, i. e. starting from an initial population $C(0)$,

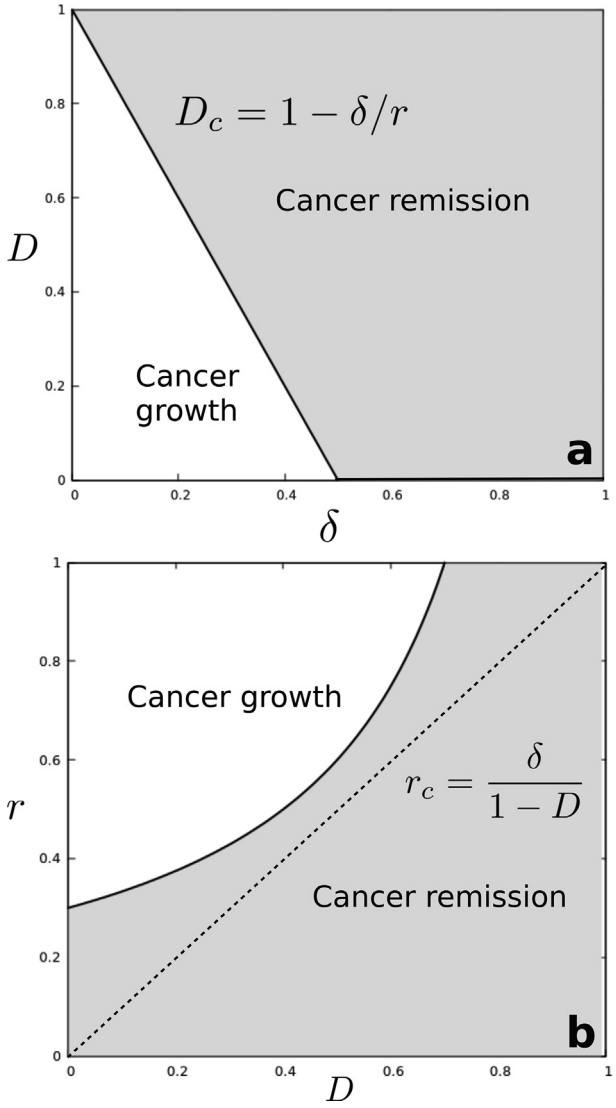

Fig. 2. Phase space of cytotoxic-differentiation combination therapies. In **(a)**, the use of DTH can induce cancer remission even for death rates smaller than the tumor cells replication rate. In **(b)**, the nonlinear effect of DTH is pictured. For $C \ll 1 - D$, the replication rate necessary for cancer outgrowth r_c grows linearly with δ (dashed line), while increasing D imposes a stronger condition (curved black line).

$$C(t) \approx C(0)e^{[r(1-D)-\delta]t} \tag{7}$$

which gives cancer expansion only if the growth rate of the cancer cells is larger than a threshold value r_c, namely

$$r > r_c = \frac{\delta}{1 - D} \tag{8}$$

We can appreciate here the difference between the impact of cytotoxic therapy (acting linearly) and DTH (acting in a nonlinear fashion). In Fig. 2b we summarize these results by displaying cancer growth rates against the efficiency of DTH. In the absence of DTH, the tumor will grow if $r > \delta$, but increasing habitat differentiation results in a nonlinear increase of the proliferation rate that tumor cells need to survive. A quick comparison with standard epidemiology models shows that this corresponds to epidemics suppression through vaccination: as more individuals are vaccinated and thus moved out from the pool of potentially infected individuals, the pathogen requires an increase in infectivity that might not be achievable.

How robust are the previous results? The model described above lacks an obvious dynamical layer: the amount of differentiated habitat follows in fact from cancer cells that have become differentiated. This can be clearly described by introducing an explicit, dynamical compartment of differentiated cells:

$$\begin{aligned} \frac{dC}{dt} &= rc(1-D-C) - \delta C - \gamma C\phi(C,D) \\ \frac{dD}{dt} &= \gamma C\phi(C,D) - \rho D \end{aligned} \tag{9}$$

where γ is now the rate at which cancer cells differentiate, and $\phi(C,D)$ describes the possible control network relating differentiation to the amount of populations at play (Vainstein et al., 2012; Yang et al, 2017; Komarova and van den Driessche, 2018). Furthermore, differentiated cells appear as they are produced by C, and disappear at rate ρ due to cellular death and removal (efferocytosis).

Cellular fate decisions, and differentiation in particular, are known to be controlled by regulatory signals related to different cell populations (Yang et al, 2017). However, the exact control networks governing differentiation processes are widely unknown. Knowing the clinical particularities of differentiation therapy, namely its major failure when administered as a single anticancer drug, can we infer possible expressions for $\phi(C,D)$?

Among the possible functional forms for differentiation control $\phi(C,D)$, we expect a minimal function that is, at least, $f(C)$. If not, stability of the $(0,0)$ attractor would follow from

$$\frac{\partial}{\partial C}\left(\frac{dC}{dt}\right) = r - \delta - \gamma\phi(D=0) \tag{10}$$

In this case, DTH alone would be able to eradicate tumor growth without chemotherapy ($\delta = 0$), provided that

$$\gamma > \frac{r}{\phi(D=0)} \tag{11}$$

which contradicts, as discussed above, clinical observations (de Thé, 2018). Our minimal assumption, consistent with computational results in (Vainstein et al., 2012) is that regulation of differentiation is orchestrated by the amount of surrounding C cells, not considering secondary regulation effects by D cells:

$$\begin{cases} \frac{dC}{dt} = rC(1-D-C) - \delta C - \gamma C^2 \\ \frac{dD}{dt} = \gamma C^2 - \rho D \end{cases} \tag{12}$$

namely, differentiation increases due to cell–cell interactions, as introduced by the term γC^2. This self-regulation term, similar to the one captured by a cellular automata in (Vainstein et al., 2012), is needed to capture that a larger differentiation rate alone will not be able to stop tumor progression (Fig. 3a). This assumption relies on a phenomenological approach to modeling, as the precise regulatory mechanisms of cellular fate remain largely unknown (Yang et al, 2017).

In this model, stability of the cancer-free state is only controlled by chemotherapy, provided that $\delta > r$. What is the role of DTH here? This can be seen by obtaining the attractor states as a function of γ:

$$C^*(\gamma) = \frac{r(1-D^*(\gamma)) - \delta}{r+\gamma} \tag{13}$$

$$D^*(\gamma) = \frac{1}{2r^2}\left[-\beta(\gamma) \pm \left(\beta(\gamma)^2 - 4r^2(r-\delta)^2\right)^{1/2}\right] \tag{14}$$

with

$$\beta(\gamma) = -2r(r-\delta) - \rho\frac{(r+\gamma)^2}{\gamma} \tag{15}$$

Here, modifying γ alone cannot modify the existence of at least one attractor state that is real and positive (see Fig. 3a). However, because of the combined effect of differentiation in both γ and tumor growth constraints $r(1-D)$, agents incrementing γ will reduce overall tumor size, thus reducing the time and dose of a secondary cytotoxic agent needed to completely erase the tumor (Fig. 3b).

Several interesting points arise from describing cancer differentiation under the perspective of habitat-loss ecological models. First of all, the minimal model of DTH as habitat loss captures

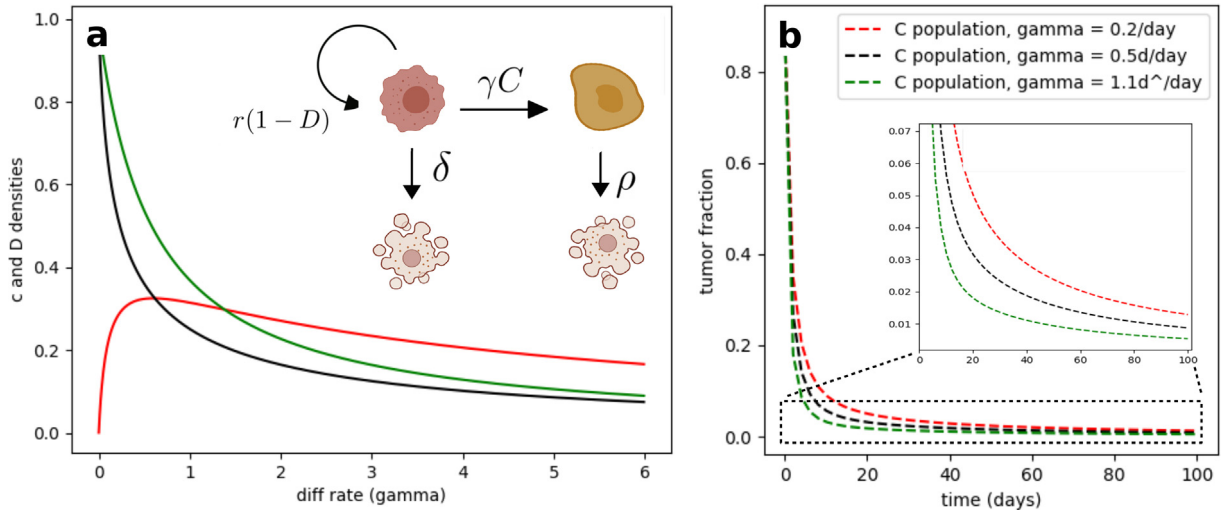

Fig. 3. DTH and chemotherapy in the C,D model. In **a**, increasing levels of γ do not totally eradicate the tumor, but reduce the overall C (dark line) D (red line) populations, as it can be seen by computing the analytical attractor states (see SM). The green line indicates how the equilibrium C population would evolve in the absence of habitat-loss effects, indicating the relevance of considering the ecological dynamics of DTH. This is indicative of the possibilities of DTH when administered with cytotoxic therapies. In **b**, three different values of increasing γ show that a tumor targeted by CHT needs shorter chemotherapy interventions as DTH increases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the existence of a differentiation threshold for tumor arrest. This highlights the role of habitat in tumor growth, with differentiation therapy driving an all-or-none response similar to that seen in APL treatment outcomes (Huang et al., 1988). When introducing the cellular dynamics of the differentiated compartment D, the sharp threshold gets diluted, but the model still shows the opportunities of DTH provided it is not delivered as a single agent.

Both model versions indicate that DTH is only effective when combined with cytotoxic therapies directly targeting the cellular death rate δ (Fig. 1). This could explain why arsenic, that triggers p53-driven senescence apart from differentiation (Ablain et al., 2014), is functional as a single-agent therapy, while retinoic acid and other differentiation drugs that do not target cell death specifically require combined cytotoxic therapy to success (Dos Santos et al., 2013). The study of differentiation and tumor habitats becomes even more relevant in the context of Cancer Stem Cells (Meacham and Morrison, 2013). How do results change when a tumor seeding population resides in a different habitat?

3. Differentiation therapy in hierarchical tissues

A broad range of cancer types are hierarchically organized, with a population of *cancer stem cells* (CSC) driving tumor growth and plasticity (Meacham and Morrison, 2013). Besides the relevance of this in radio- and chemotherapy resistance (see e.g. Dean et al., 2005), we are interested in understanding if the hierarchical architecture specific to a stem cell compartment is related with the fact that most solid and genetically complex tumors do not show valuable responses to differentiation therapy (Cruz and Matushansky, 2012; de Thé, 2018).

A wide range of mathematical models have been powerful in highlighting the sometimes undercover role of cancer stem cells (see e.g. Michor et al., 2005; Michor, 2008 and references therein). We here consider a minimal view of the accepted modeling of tissue architecture as a set of hierarchically organized cancer subpopulations (Michor et al., 2005; Dingli and Michor, 2006, Solé et al., 2008, Fig. 4). Following the previous approximations, we here add a seeding CSC compartment, S:

$$\begin{cases} \frac{dS}{dt} = r_s S\varphi(S,C,D) - \gamma_S S^2 - \delta_S C \\ \frac{dC}{dt} = rC(1-D-C) + \gamma_S S^2 - \delta C - \gamma C^2 \\ \frac{dD}{dt} = \gamma C^2 - \rho D \end{cases} \tag{16}$$

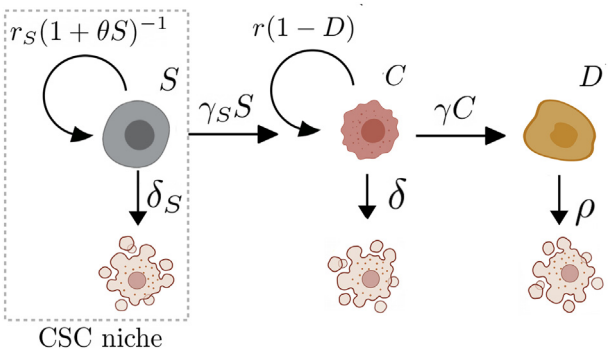

Fig. 4. Tissue architecture and the ecology of tumor differentiation therapy. The minimal hierarchical model involves a cancer stem cell compartment S that replicates under the constraints of a well-separated niche, dies under cytotoxic therapy with diminished effectivity ($\delta_S \ll \delta$) and seeds a progenitor cancer population C..

Here C indicates the progenitor compartment, that differentiates following self-regulated mechanisms and feels the crowding (spatial) effects of these terminal cells D. This population, in turn, is seeded by a particular CSC compartment S, that replicates at rate r_S. The effect of chemotherapy is captured by δ_S, which is in generally much smaller than that of non-stem cancer cells due to plasticity or quiescence potential of CSC. Differentiation rate into the general progenitor compartment is captured by γ_S and in a first approximation is believed to follow autocontrol regulation as in (Vainstein et al., 2012). Replication is again constrained by how populations occupy space, here $\varphi(S,C,D)$. Stem cells (and CSCs in particular) are known to inhabit in a well-differentiated spatial niche (Plaks et al., 2015; Voog and Jones, 2010), such as the bone marrow for hematopoietic stem cells (Adams and Scaden, 2006). This indicates that, most often, CSC replication –in contrast with differentiation– might not be regulated by the density of the tumor populations. Here we have $\varphi(S,C,D) = \varphi(S)$.

Furthermore, crowding effects observed in the bone marrow, together with niche and carrying capacity modulation (Dingli and Michor, 2006; Gerlee and Anderson, 2015), indicate that a logistic growth model with absolute saturation ($r(C=K)=0$) might not be accurate to describe CSC plasticity. A more general saturating function φ considers an adimensional CSC sensitivity to crowding θ (see Dingli and Michor, 2006). In our case, the model is written as

$$\begin{cases} \frac{dS}{dt} = r_s S\frac{1}{1+\theta S} - \gamma_S S^2 - \delta_S C \\ \frac{dC}{dt} = rC(1-D-C) + \gamma_S S^2 - \delta C - \gamma C^2 \\ \frac{dD}{dt} = \gamma C^2 - \rho D \end{cases} \tag{17}$$

with the particularity that CSC replication does not totally stop, rather it slows down as the CSC niche becomes populated. This crowding effect cannot be correctly captured by cellular automata models studying DTH (such as Vainstein et al., 2012), as the number of neighbouring cells is considered constant. What is, once again, the effect of space in DTH? Could this CSC niche-specificity explain why some cancers are resistant to differentiating agents?

It can be seen how DTH, even if considered in a totally symmetric fashion where $\gamma_S = \gamma$, does not have the same effect in a hierarchical tumor if CSCs inhabit a different niche (Fig. 3a). In particular, the model shows that, for low differentiation rates, the general cancer population needs a similar time until eradication as with the logistic growth model (Fig. 5a, red curves). However, as γ increases, not only CSCs residing in a different niche, but also the rest of the tumor becomes harder to eradicate (black curves). The model therefore captures how stem cell resistance to differentiation approaches not only follows from resides in cellular plasticity (Foo et al., 2009; Meacham and Morrison, 2013), but also from the role of habitat ecology in niche construction and independence (Adams and Scaden, 2006).

Along with CSCs death and differentiation, sensitivity to CSC density θ also plays a role in tumor extinction. Increasing θ results in a linear decrease in the time of tumor eradication under chemotherapy (Fig. 5b). This result, together with the overall indications of this work regarding the role of spatial effects of DTH, supports the application of *niche therapy* for CSCs (Plaks et al. 2015). Previous efforts had already considered the possibility of disrupting the CSC niche, such as through VEGF inhibition reducing blood vessel production (Calabrese et al., 2007). Together with this, our results propose that CSCs become weaker to DTH when increasing their sensitivity to the rest of tumor populations by physical disruption of their niche.

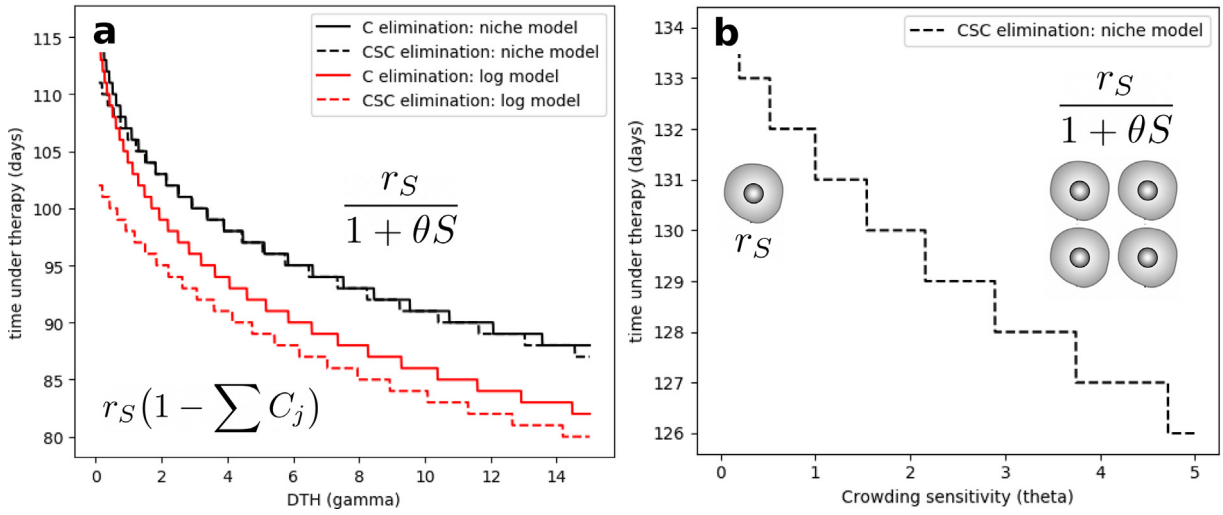

Fig. 5. The role of DTH and spatial constraints in hierarchical tumor architectures. In **(a)**, the duration of therapy (for $\delta_i > r_i$) until tumor eradication is computed as a function of DTH effectivity (γ). In particular, simulations compare the duration of therapy under a generic model (red) and a model that captures how CSCs inhabit a specific niche, and therefore do not feel the spatial effects of an increasingly differentiated habitat. In **(b)**, the effect of crowding sensitivity (θ) in the CSC model is studied. It is easily seen how an increasing sensitivity to surrounding cells results in reduced therapy duration. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

In this paper we have shown how ecological models of habitat loss can shed light into several aspects of differentiation therapy in cancer. This is done by using habitat loss as a surrogate of differentiated patches, while an independent extinction term is matched by the effects of cytotoxic therapy. Early models of ecological decline due to habitat loss show that a well-defined threshold exist: once a given critical loss is present, no viable populations are allowed, despite that some amount of habitat is still present (Levins, 1969; Bascompte and Solé, 1996). Within the cancer context, when a critical amount of cancer cells have been differentiated, remission results in a similar fashion, provided that standard cytotoxic therapy is also present as seen in the clinics of APL (de Thé, 2018). However, the dynamical nature of differentiation implies that not a sharp threshold, but rather a progression towards cancer eradication result from DTH. In order to test the generality of the approximation, both an homogeneous metapopulation model and an extension considering the specificity of a cancer stem cell compartment and its niche have been explored. The models consistently explain several qualitative observations concerning the impact of DTH.

On the one hand, approaching differentiation as an ecological process for a simple population and its terminally differentiated surrogate can predict interesting dynamics in genetically simple cancers such as APL where DTH has been successful. Our model predicts a well defined threshold for the amount of differentiated habitat, beyond which a malignant population is not able to progress. The fact that eradication with differentiating agents is totally dependent on cytotoxic therapy is consistent with studies on DTH for leukemia, where arsenic, that triggers p-53 driven senescence as well as differentiation, is effective as a single-agent therapy, while other agents might require combined cytotoxicity. Further learning from the single population model indicate that DTH becomes much more effective than chemotherapy for small, growing tumors away from the carrying capacity of their tissue. This result opens novel questions on the role of DTH as an early therapeutic scheme.

An extension of the ecological model introduces a minimal architecture to understand the possible role of a cancer stem cell niche in sensitivity to DTH. In particular, we aim at understanding how niche specificity in CSCs might be playing a role in resistance to DTH agents for certain tumor types. The model indicates how a hierarchical tumor seeded by a CSC population inhabiting a separated niche might be much more difficult to eradicate through DTH than in well-mixed approaches. This indicates that not only the cellular characteristics of cancer stem cells, but also the ecological interactions building their niche (namely, crowding and spatial competition) might be key in understanding their resistance to this kind of therapeutic approaches.

A large body of literature concerning mathematical modeling of cancer tissue hierarchies has grown in the course of recent years (Foley and Mackey, 2009), from the understanding of cell fate decisions (Sun et al., 2015), the molecular basis of cytokine and division dynamics (Stiehl and Marciniak-Czochra, 2012; Stiehl et al., 2015) or the study of control networks regulating CSC differentiation (Yang et al, 2017; Komarova and van den Driessche, 2018), among many others. Despite using a similar fundamental background, our approach using the learning of habitat ecology is able to mathematically describe the phenomenological role of spatial interactions and how these, and the anatomical niches, might result in success or failure of DTH combined with other agents. Furthermore, the role of spatial distribution of CSC-niche interactions, and the possibility of disrupting the barriers of this distribution, appears as a possible improvement to the cytotoxic + DTH combination approach when cancer stem cells are in place.

Several shortcomings, potential extensions and implications of this work can be outlined. First of all, the model involves the most minimal set of rules and assumptions, sacrificing the details of the population description in favor of an ecological picture that can be intuitively interpreted. Real tumours include several layers of complexity, such as cell–cell interactions or micro-environmental cues that we have captured through phenomenological modeling alone. Moreover, models of habitat loss including noise reveal the importance of considering several sources of disturbance, from demographic stochasticity to large catastrophes (Casagrandi and Gatto, 2002).

Further exploration would require considering, for example, the heterogeneous spatial organization and its Impact on tumor growth (Sottoriva et al., 2013). However, the need for a more realistic model does not invalidate the key findings of our study. In fact, a similar criticism could be raised in relation to the simplicity of habitat loss models derived from Levins equation. Despite their

minimalism, the simplest models (also ignoring the details of species-specific metabolic or physiologic features) have been extremely valuable in understanding the problem as well as how to prevent its consequences (Lande, 1988; Hanski, 1999).

CRediT authorship contribution statement

Ricard Solé: Conceptualization, Formal analysis, Methodology, Writing - original draft. **Guim Aguadé-Gorgorió:** Conceptualization, Formal analysis, Methodology, Writing - original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors thank the Complex Systems Lab members for fruitful discussions. Special thanks to Alvah Bassie for his inspiring ideas. This work was supported by the Botín Foundation by Banco Santander through its Santander Universities Global Division, the Spanish Ministry of Economy and Competitiveness, grant FIS2016-77447-R MINECO/AEI/FEDER, an AGAUR FI 2018 grant, and the Santa Fe Institute where most of this work was done.

References

Ablain, J. et al., 2014. Activation of a promyelocytic leukemia-tumor protein 53 axis underlies acute promyelocytic leukemia cure. Nature Medicine 20, 167–174.
Adams, G.B., Scadden, D.T., 2006. The hematopoietic stem cell in its place. Nature Immunology 7 (4), 333–337.
Aktipis, C.A., Boddy, A.M., Jansen, G., Hibner, U., Hochberg, M.E., Maley, C.C., Wilkinson, G.S., 2015. Cancer across the tree of life: cooperation and cheating in multicellularity. Philosophical Transactions of the Royal Society B: Biological Sciences 370 (1673), 20140219.
Basanta, D., Anderson, A.R., 2013. Exploiting ecological principles to better understand cancer progression and treatment. Interface Focus 3 (4), 20130020.
Bascompte, J., Solé, R.V., 1996. Habitat fragmentation and extinction thresholds in spatially explicit models. Journal of Animal Ecology. 465–473.
Calabrese, C., Poppleton, H., Kocak, M., Hogg, T.L., Fuller, C., Hamner, B., Frank, A., 2007. A perivascular niche for brain tumor stem cells. Cancer Cell 11 (1), 69–82.
Casagrandi, R., Gatto, M., 2002. Habitat destruction, environmental catastrophes, and metapopulation extinction. Theoretical Population Biology 61, 127–140.
Cruz, F.D., Matushansky, I., 2012. Solid tumor differentiation therapy–is it possible?. Oncotarget 3 (5), 559.
Davies, P.C., Lineweaver, C.H., 2011. Cancer tumors as Metazoa 1.0: ta ing genes of ancient ancestors. Physical Biology 8, (1) 015001.
Dean, M., Fojo, T., Bates, S., 2005. Tumour stem cells and drug resistance. Nature Reviews Cancer 5, 275–284.
de Thé, Hugues, 2018. Differentiation therapy revisited. Nature Reviews Cancer 18, 117–127.
Dingli, D., Michor, F., 2006. Successful therapy must eradicate cancer stem cells. Stem Cells 24, 2603–2610.
Dos Santos, G.A., Kats, L., Pandolfi, P.P., 2013. Synergy against PML-RARa: targeting transcription, proteolysis, differentiation, and self-renewal in acute promyelocytic leukemia. Journal of Experimental Medicine 210, 2793–2802.
Foley, C., Mackey, M.C., 2009. Dynamic hematological disease: a review. J Math Biol. 58 (1–2), 285–322. https://doi.org/10.1007/s00285-008-0165-3.
Foo, J., Drummond, M.W., Clarkson, B., Holyoake, T., Michor, F., 2009. Eradication of chronic myeloid leukemia stem cells: a novel mathematical model predicts no therapeutic benefit of adding G-CSF to Imatinib. PLOS Computational Biology 5, (9) e1000503.
Gatenby, R.A., 2009. A change of strategy in the war on cancer. Nature 459 (7246), 508.
Gatenby, R.A., Gillies, R.J., 2008. A microenvironmental model of carcinogenesis. Nature Reviews Cancer 8 (1), 56–61.
Gerlee, P., Anderson, A.R., 2015. The evolution of carrying capacity in constrained and expanding tumour cell populations. Physical Biology 12, (5) 056001.
Gruber, M., Bozic, I., Leshchiner, I., Livitz, D., Stevenson, K., Rassenti, L., Fernandes, S. M., 2019. Growth dynamics in naturally progressing chronic lymphocytic leukaemia. Nature 570 (7762), 474–479.
Hanski, I., 1999. Metapopulation Ecology. Oxford University Press.
Huang, M.E., Ye, Y.C., Chen, S.R., Chai, J.R., Lu, J.X., Zhoa, L., Gu, L.J., Wang, Z.Y., 1988. Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. Blood 72 (2), 567–572.
Komarova, N.L., van den Driessche, P., 2018. Stability of control networks in autonomous homeostatic regulation of stem cell lineages. Bulletin of Mathematical Biology 80 (5), 1345–1365. https://doi.org/10.1007/s11538-017-0283-4.
Korolev, K.S., Xavier, J.B., Gore, J., 2014. Turning ecology and evolution against cancer. Nature Reviews Cancer 14 (5), 371.
Lande, R., 1988. Genetics and demography in biological conservation. Science 241 (4872), 1455–1460.
Levins, R., 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. American Entomologist 15 (3), 237–240.
Magee, J.A., Piskounova, E., Morrison, S.J., 2012. Cancer stem cells: Impact, heterogeneity and uncertainty. Cancer Cell 21, 283–296.
Meacham, C.E., Morrison, S.J., 2013. Tumour heterogeneity and cancer cell plasticity. Nature 501, 328–337.
Merlo, L.M., Pepper, J.W., Reid, B.J., Maley, C.C., 2006. Cancer as an evolutionary and ecological process. Nature Reviews Cancer 6 (12), 924.
Michor, F., 2008. Mathematical models of cancer stem cells. Journal of Clinical Oncology 26 (17), 2854–2861.
Michor, F., Hughes, T.P., Iwasa, Y., Branford, S., Shah, N.P., Sawyers, C.L., Nowak, M.A., 2005. Dynamics of chronic myeloid leukemia. Nature 435, 1267–1270.
Misale, S., Bozic, I., Tong, J., Peraza-Penton, A., Lallo, A., Baldi, F., Di Nicolantonio, F., 2015. Vertical suppression of the EGFR pathway prevents onset of resistance in colorectal cancers. Nature Communications 6 (1), 1–9.
Moilanen, A., Hanski, I., 1998. Metapopulation dynamics: effects of habitat quality and landscape structure. Ecology 79 (7), 2503–2515.
Norton, L., 1988. A Gompertzian model of human breast cancer growth. Cancer Research 48 (24 Part 1), 7067–7071.
Pepper, J.W., Scott Findlay, C., Kassen, R., Spencer, S.L., Maley, C.C., 2009. Synthesis: cancer research meets evolutionary biology. Evolutionary a Lications 2 (1), 62–70.
Pierce, G.B., 1983. The cancer cell and its control by the embryo. Rous-Whi le Award lecture. The American Journal of Pathology 113 (1), 117.
Pierce, G.B., Wallace, C., 1971. Differentiation of malignant to benign cells. Cancer Research 31 (2), 127–134.
Plaks, V., Kong, N., Werb, Z., 2015. The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? Cell Stem Cell 16 (3), 225–238.
Solé, R.V., Rodríguez-Caso, C., Deisboeck, T.S., Saldaña, J., 2008. Cancer stem cells as the engine of unstable tumor progression. Journal of Theoretical Biology 253, 629–637.
Sottoriva, A., Spiteri, I., Piccirillo, S.G.M., Touloumis, A., Collins, P., Marioni, J.C., Curtis, C., Watts, C., Tavaré, S., 2013. Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics. Proceedings of the National Academy of Sciences 110 (10), 4009–4014.
Stiehl, T., Marciniak-Czochra, A., 2012. Mathematical modelling of leukemogenesis and cancer stem cell dynamics. Mathematical Modeling of Natural Phenomena 7, 166–202.
Stiehl, T., Baran, N., Ho, A.D., Marciniak-Czochra, A., 2015. Cell division patterns in acute myeloid leukemia stem-like cells determine clinical course: a model to predict patient survival. Cancer Research 75 (6), 940–949.
Sun, X., Zheng, X., Zhang, J., Zhou, T., Yan, G., Zhu, W., 2015. Mathematical modeling reveals a critical role for cyclin D1 dynamics in phenotype switching during glioma differentiation. FEBS Letter 589 (18), 2304–2311. https://doi.org/10.1016/j.febslet.2015.07.
Vainstein, V., Kirnasovsky, O.U., Kogan, Y., Agur, Z., 2012. Strategies for cancer stem cell elimination: insights from mathematical modeling. Journal of Theoretical Biology 298, 32–41. https://doi.org/10.1016/j.jtbi.2011.12.016.
Voog, J., Jones, D.L., 2010. Stem cells and the niche: a dynamic duo. Cell Stem Cell 6 (2), 103–115.
Weinberg, R.A., 2014. The biology of cancer. Garland Science.
Werner, B., Scott, J.G., Sottoriva, A., Anderson, A.R.A., Traulsen, A., Altrock, P.M., 2016. The cancer stem cell fraction in hierarchically organized tumors can be estimated using mathematical modeling and patient-specific treatment trajectories. Cancer Research 76 (7), 1705–1713.
Yang, J., Axelrod, D.E., Komarova, N.L., 2017. Determining the control networks regulating stem cell lineages in colonic crypts. Journal of Theoretical Biology 429, 190–203. https://doi.org/10.1016/j.jtbi.2017.06.033.

Further reading

Solé, R.V., Alonso, D., Saldaña, J., 2004. Habitat fragmentation and biodiversity collapse in neutral communities. Ecological Complexity 1 (1), 65–75.

Transition therapy: tackling the ecology of tumor phenotypic plasticity

Guim Aguadé-Gorgorió,[1,2] Stuart Kauffman,[3] and Ricard Solé*[1,2,4]
[1]*ICREA-Complex Systems Lab, Universitat Pompeu Fabra, 08003 Barcelona, Spain*
[2]*Institut de Biologia Evolutiva, CSIC-UPF, 08003 Barcelona, Spain*
[3]*Institute for Systems Biology, Seattle WA 98109, USA*
[4]*Santa Fe Institute, Santa Fe NM 87501, USA*

Phenotypic switching in cancer cells has been found to be present across tumor types. Recent studies on Glioblastoma report a remarkably common architecture of four well-defined phenotypes coexisting within high levels of intra-tumour genetic heterogeneity. Similar dynamics have been shown to occur in breast cancer and melanoma, and are likely to be found across cancer types. Given the adaptive potential of phenotypic switching (PHS) strategies, understanding how it drives tumor evolution and therapy resistance is a major priority. Here we present a mathematical framework uncovering the ecological dynamics behind PHS. The model is able to reproduce experimental results, and mathematical conditions for cancer progression reveal PHS-specific features of tumors with direct consequences on therapy resistance. In particular, our model reveals a threshold for the resistant-to-sensitive phenotype transition rate, below which any cytotoxic or switch-inhibition therapy is likely to fail. The model is able to capture therapeutic success thresholds for cancers where larger PHS architectures are in play, such as glioblastoma or melanoma. This results in a set of conditions for combination therapies targeting replication and phenotypic transitions at once. Following our results we discuss *transition therapy* as a novel scheme to target not only combined cytotoxicity but also the rates of phenotypic switching.

Keywords: Cancer ecology, phenotypic switching, epigenetic plasticity, combination therapies, transition therapy.

I. INTRODUCTION

Phenotypic plasticity is a widespread phenomenon across the tree of life. From bacteria to multicellular development, epigenetic pathways generate a population of diverse phenotypes from homogeneous, stable genomes [1-4]. Phenotypic switching (PHS) is a stochastic phenomenon known to maintain population diversity in unicellular organisms as a means to survive in fluctuating environments [5,6]. This mechanism can also be found to boost non-genetic heterogeneity in a special multicellular context: cancer cell populations [7]. In this context, tumors can take advantage of already existing differentiation hierarchies to promote unlimited self-renewal or senescence and drug resistance with no need of selecting somatic mutations [8,9].

Phenotypic switching is a source for non-genetic heterogeneity in cancer beyond Cancer Stem Cells hierarchies [7,10,11]. The most recent example comes from Glioblastomas, where tumor cells are found to organize around four well-defined meta-modules resembling -though aberrant- healthy brain cell lines [12]. This arrangement is highly robust: tumors initiated by single cells from a biopsy evolve towards the previous phenotypic composition, regardless of the specific phenotype of the original cell, showing that stochastic transitions happen between all of the four phenotypes. Similar dynamics have been described in breast cancer [13], as well as in melanoma [14,15] and prostate cancer [16], and are nowadays considered key in the observation of non-Darwinian evolution of adaptive resistance across cancer types [17-19].

The existence of phenotypic plasticity in tumors has important consequences for therapy. Tumor relapse after therapy is usually acknowledged to be a consequence of pre-existing or acquired resistance mutations, present in a given subclone that survives and repopulates the tumor (see e.g [20]). This image is often correct, yet further mechanisms in many therapeutic settings, from stem cell senescence [21] to immunological editing [22] prove that a wider scope is key when trying to understand therapeutic failure. The stochastic nature of switching between rogue cellular phenotypes allows robust and plastic tissue architectures, resulting in an adaptive mechanism that might be even harder to target [17]. How does this affect therapeutic strategies? Models of phenotypic switching have helped to explore cancer invasion [23-25] or the possible role of plasticity in maintaining one or more resistant phenotypes in place [19,26].

Here we present a toy model to study the characteristics of phenotypic plasticity in cancer by exploring the population dynamics of competing replicators exhibiting transitions among them (Fig. 1). The model allows in particular to analyze the rise of the switching populations and the equilibrium conditions for stable heterogeneity, as well as the requirements to tumor extinction with implications on novel therapeutic approaches when more than two phenotypes are in place.

II. PHENOTYPIC SWITCHING DYNAMICS

In this section we explore several features exhibited by different versions of a toy model of cancer cell pop-

*corresponding author

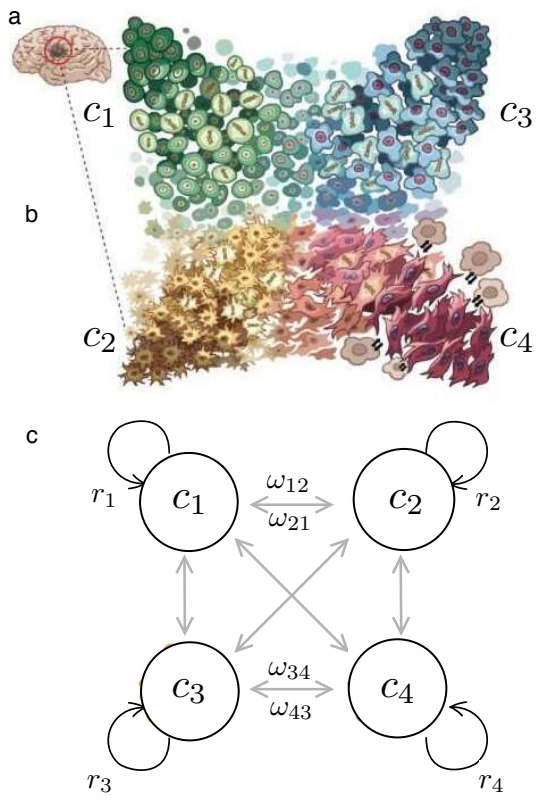

FIG. 1 **Phenotypic switching in cancer.** Genetic analysis reveals four transitioning phenotypes in Glioblastoma (a) and thus a set of cancer cell populations (b, after Neftell et al., 2019). Different transitions occur, linking phenotypes C_k by means of a matrix of transition rates, as sketched in (c).

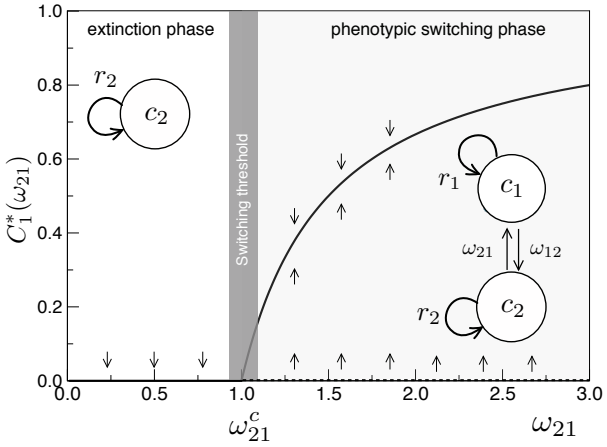

FIG. 2 Bifurcation diagram for the reduced $N = 2$ PHS model with two strains, as defined by equation (5) where the C_1 population is analyzed under CPC. This diagram represents the fixed points C_1^* against the transition rate ω_{21}. A critical switching threshold is defined here for a given ω_{21}^c separating a heterogeneous phase (gray) from a homogeneous one. Here $r_1 = 1, r_2 = 3/2$ and $\omega_{12} = 1/2$, which gives a critical value $\omega_{21} = 1.0$ (equation 8).

ulations exhibiting PHS. Our goal is to provide some basic bounds to the response of these systems to cytotoxic or targeted agents acting on the switching dynamics. Ecological models of heterogeneous cancer populations can be represented by means of a set of replicator equations [27]. Consider a set of N phenotypes, where $\mathbf{C} = (C_1, ..., C_N)$. The i-th cancer cell type population will change in time following:

$$\frac{dC_i}{dt} = \Gamma_i(\mathbf{C})C_i + \sum_{k\neq i}\omega_{ki}C_k - \sum_{k\neq i}\omega_{ik}C_i - C_i\phi(\mathbf{C}) \quad (1)$$

with $(i, k = 1, ..., N)$. Here $\Gamma_i(\mathbf{C})$ indicates the functional form of the replication rate associated with the $i-$th clone, which in general will be a nonlinear function of clone or tumor size [28]. The three last terms in the rhs correspond to (1) the phenotypic transitions from other phenotypes to phenotype C_i (i. e. $C_k \to C_i$) (2) the complementary transitions from C_i to the rest (i. e. $C_i \to C_k$) and (3) an outflow term that allows introducing competition and resource limitation effects. The previous set of equations can be re-written as follows:

$$\frac{dC_i}{dt} = \left(\Gamma_i(\mathbf{C}) - \sum_{k\neq i}\omega_{ik}\right)C_i + \sum_{k\neq i}\omega_{ki}C_k - C_i\phi(\mathbf{C}) \quad (2)$$

By aggregating those terms affecting C_i we can appreciate the fact that the effective growth rate of C_i involves a trade-off between intrinsic replication and the likelihood that it shifts to a different cell type. However, a negative balance can be counterbalanced by the net inflow from the rest of the phenotypes holding C_i in place. As a first approximation, a constant replication rate is associated to each phenotype (i. e. $\Gamma_i(\mathbf{C}) = r_i$).

What is the impact of PHS on potential therapeutic approximations? Are there novel attractors or alternative pathways to avoid targeted death? Relevant insight can be obtained by considering a minimal system, where a finite set of cancer clones replicate at rate r_i, defined as the effective difference $r_i = b_i - d_i$ between birth b_i and death d_i rates, and that can be negative when cytotoxic therapy is effective (increasing death beyond birth, see Fig. 3a). In this section we consider the simplest models of PHS in cancer populations.

A. Predictable heterogeneity in PHS tumors

Experimental evidence in cancer populations exhibiting PHS shows that a secondary tumor evolves to the original phenotypic distribution of the primary malignancy, regardless of the initiating cell type [12,13]. This is an interesting outcome of PHS: the system has the potential to reliably restore population diversity in a pre-

dictable fashion. Instead of the often unpredictable heterogeneity driven by somatic mutations, we have here a surrogate of developmental dynamics driven by epigenetic changes. A first mathematical approach and its consequences are easily derived considering a population of two switchers ($N = 2$) under a constant population constraint (CPC) [6]. Such CPC constraint allows for direct analysis of population fractions or densities $c_i = C_i/\sum_\mu C_\mu$, and writes

$$\frac{dc_1}{dt} = (r_1 - w_{12})c_1 + w_{21}c_2 - c_1\phi(\mathbf{C}) \quad (3)$$

$$\frac{dc_2}{dt} = (r_2 - w_{21})c_2 + w_{12}c_1 + -c_1\phi(\mathbf{C}) \quad (4)$$

This equation reduces to a simple competition model when $\omega_{ij} = 0$. Darwinian selection would then be decided by the highest r_i, eliminating the possibility for heterogeneity.

Assuming constant population, the competition term reads $\phi(\mathbf{C}) = r_1c_1 + r_2c_2$ and considering that c_i are here densities and $c_1 + c_2 = 1$, this is in fact the average replication rate, i.e. $\phi(\mathbf{C}) = \langle r\rangle$. Using this result, it is possible to reduce the system to a one-dimensional ordinary differential equation for the fraction of one of the populations, say c_1:

$$\frac{dc_1}{dt} = \gamma c_1(1 - c_1) - w_{12}c_1 \quad (5)$$

with $\gamma = (r_1 - r_2 - w_{21})$. This model displays two fixed points, namely $c_1^* = 0$ (extinction) and the heterogeneous point (where both populations persist) given by

$$c_1^* = 1 - \frac{w_{12}}{\gamma} \quad (6)$$

Interestingly, the presence of an heterogeneous attractor that is not dependent on initial phenotypic composition can be compared to experimental evidence of cell growth recapitulating original clonal distributions [12,13]. In particular, it can be seen that the attractor for population distributions, c_1^*/c_2^*, is consistent with the long-term stable distribution in the absence of intrinsic competition, $\lim_{t\to\infty} C_1(t)/C_2(t)$, because the CPC assumption is equivalent to formulating the model in terms of population concentrations (see SM). This result is consistent both analytically and through computer simulations, so that the minimal model is able to generate the basic *in vitro* properties of phenotypic switching. This, in turn, indicates that experimental observations of phenotypic distributions can be used to estimate the switching parameters that hold the heterogeneous cellular architecture, as previously seen in [13,19,29].

Under which conditions is the system able to maintain heterogeneity beyond the pressure of strictly-competitive Darwinian selection? The stability analysis of this system shows that heterogeneity will persist (i.e. $c_1^*, c_2^* > 0$) and any initial condition will recapitulate the whole attractor distribution provided that

$$\omega_{21} - \omega_{12} > r_2 - r_1. \quad (7)$$

This inequality has an interesting, intuitive interpretation: c_1 will be positive, even if $r_2 > r_1$, provided that the difference between transition rates is larger than the difference between growth rates, highlighting the ability of PHS to maintain tumor heterogeneity (Fig. 2). This allows defining a threshold value: heterogeneity will be observed when

$$\omega_{21}^c = \omega_{12} + (r_2 - r_1) \quad (8)$$

which determines the threshold condition for the switching rate ω_{21} required to sustain C_1, being other parameters fixed. The basic bifurcation diagram associated to this model is shown on figure 2. Two phases are indicated. The first is associated to the diverse switching phenotypes (for $\omega_{21} > \omega_{21}^c$, gray area). Here a single attractor exists, which can be reached from any initial condition. Another, homogeneous phase occurs for $\omega_{21} < \omega_{21}^c$ where only the fastest replicating population persists.

The transition defines a tipping point that is determined (with other parameters fixed) by the rate of recovery provided by the PHS mechanism. The diagram is obtained under unfavorable replication: we use $r_1 < r_2$ which, in the absence of PHS, would inevitably lead to the extinction of C_1. The presence of a heterogeneous phase indicates that phenotypic populations can persist even in unfavorable competition scenarios. How does the system evolve when these populations are targeted by therapy?

B. PHS in the Sensitive-Resistant scenario

A first instance of PHS in cancer is observed in tumors deploying temporary resistant cell subpopulations [17]. In certain settings, such drug-tolerant phenotypes can arise in the absence of resistance-conferring alterations [30,31], indicating the role of non-Darwinian epigenetic plasticity in generating and maintaing tolerant phenotypes in place [29]. Modeling PHS can uncover the underlying dynamics of sensitive-resistant populations, proposing specific therapeutic outlines.

In order to formulate this model, we remove the competition term $c_i\phi(\mathbf{C})$ in the previous equations (3-4) and consider phenotypic populations away from their carrying capacity. Now C_i are not densities, but actual population counts. We study the following linear system

$$\frac{dC_1}{dt} = (r_1 - w_{12})C_1 + w_{21}C_2 \quad (9)$$

$$\frac{dC_2}{dt} = w_{12}C_1 + (r_2 - w_{21})C_2 \quad (10)$$

The unbounded system does not admit a single-population solution: the tumor either gets extinct or both $C_1(t)$ and $C_2(t)$ undergo exponential growth. As previously discussed, long-term phenotypic composition C_1/C_2 is still predictable and independent from initial

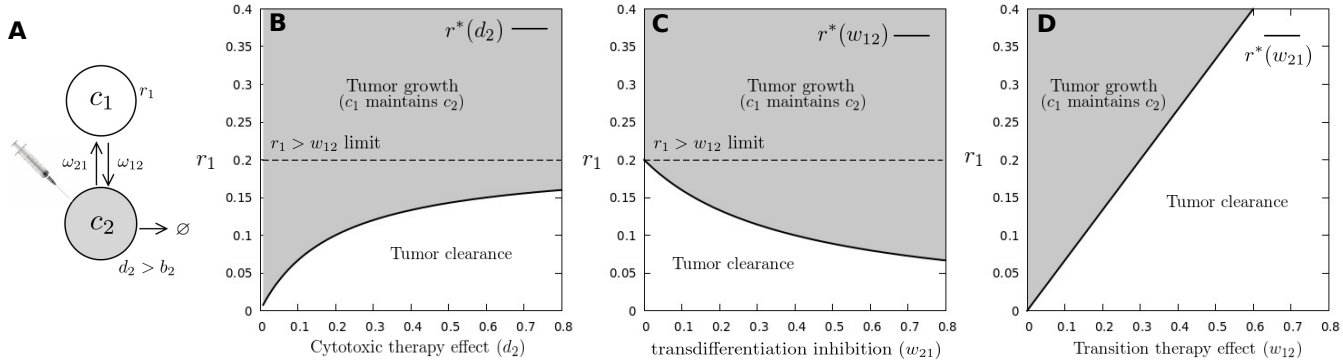

FIG. 3 **Transition therapy.** Targeting proliferation of a single phenotype in a switching tumor (**A**). In the presence of PHS strategies, the resistant population r_1 is able to maintain tumor growth. Targeting sensitive cell death (d_2, **B**) or inhibiting transitions towards resistance (w_{21}, **C**) is likely to fail provided resistant cells replicate faster than they transition into the sensitive phenotype ($r_1 > w_{12}$). PHS modeling indicates that only therapies draining c_1 into c_2 are effective across parameter settings (**C**).

conditions (see SM), as observed in experimental setups [12,13]. We know that the $(0,0)$ attractor is stable if both *effective* growth rates are negative. Since $r_i = b_i - d_i$, this can be true if death rates for both cell types are increased beyond their birth rates by means of two different drugs. However, provided C_1 is a drug-tolerant state [17], chemotherapy will only increase death rates of the C_2 population.

Let us introduce a nomenclature for cytotoxic-sensitive and -resistant phenotypes. Assume that cell type C_1 has a positive replication rate $r_1 > 0$ under chemotherapy. In this setting, the drug-resistant phenotype will be labeled C_R, growing at rate r_R. The death rate of cell type C_2 can be increased by means of a cytotoxic therapy, so that $r_2 = b_2 - d_2$ could shift from be positive to negative (Fig. 3A), and be labeled C_S, with $r_S = b_S - d_S < 0$. The sensitive-resistant system now writes

$$\frac{dC_R}{dt} = (r_R - w_{RS})C_R + w_{SR}C_S \tag{11}$$

$$\frac{dC_S}{dt} = (r_S - w_{SR})C_S + w_{RS}C_R \tag{12}$$

Stability analysis of the tumor-free attractor results in a threshold replication rate for C_R (see SM),

$$r_R^* = \frac{w_{RS}}{1 + \left(\dfrac{w_{SR}}{|r_S|}\right)} \tag{13}$$

If C_R replicates faster than this threshold level, it will repopulate the tumor and maintain the sensitive population C_S (Fig. 3). This is consistent with recent analytical results from [26] for the progression of a tumor in the presence of a drug-tolerant phenotype.

This result uncovers several potential therapeutic implications. In the setting that C_R is a drug-tolerant phenotype, therapy could focus on increasing d_S, the death rate of the sensitive phenotype [32], decreasing w_{SR}, the rate at which the sensitive phenotype becomes resistant [33], or increasing w_{RS}, the rate at which the resistant phenotype transdifferentiates into drug-sensitivity [29]. All approaches could potentially drive tumor extinction (Fig. 3).

However, if the drug-tolerant phenotype replicates faster than its transition rate ($r_R > w_{RS}$), which is a plausible setting considering measured w_{ij} rates in some cellular substates [13], any efforts on d_S or w_{SR} will fail at eliminating the tumor (Fig. 3B, 3C). Mathematically, equation (11) implies a minimal resistant-sensitive transition rate, below which the resistant population persists:

$$w_{RS}^* = r_R\,(1 + \theta_S) \tag{14}$$

with $\theta_S = w_{SR}/|r_S|$ being the transition-to-death ratio of the sensitive population. In very effective therapy settings, $\theta \sim 0$ and $w_{RS}^* = r_R$. The only path to eliminating the drug-resistant tumor is by increasing its transition rate beyond the threshold cycling rate.

This threshold has potential implications on switching inhibition, in that therapies targeting inhibition of sensitive-resistant transitions ($w_{SR} \sim 0$) are likely to fail unless the same drug alters r_R or w_{RS}. This is a key result regarding therapeutic options targeting EMT inhibition to prevent metastases [34,35].

Another particular example here is provided by the discovery of sensitive transient states in chemotherapy experiments on breast cancer [29]. In them, resistance to first-line chemotherapy implies a transition to a transient phenotype T that can be resensitized by a second drug. Initial chemotherapy increases w_{RT}, while the second drug resensitizes this transient state to initial chemotherapy, inducing w_{TS}. The overall effect is that of a combination scheme that increases w_{RS}. In the specific setting of Goldman *et al.*, the measured transition rates

from stem-cells to the induced state is $w_{RT} \approx 0.96\text{day}^{-1}$, while $r_R \approx 0.5\text{day}^{-1}$, so that therapeutic efficacy correlates with the *transition threshold* condition (14). To which extent is this specific therapeutic approach robust across cancer types?

Our results highlight the potential limitation to be accounted for when designing such PHS therapeutic strategies: increasing w_{RS}, the rate at which C_R switches to C_S, can drain the replicative phenotype into the one we can kill by cytotoxic therapy (Fig. 3D), only if it overcomes C_R replication. Transition to a sensitive state will only be effective if the resistant state cannot persist and maintain the PHS architecture.

A therapeutic corollary of this is that a most effective combination therapy in a sensitive-tolerant setting would contemplate increasing w_{RS} while also decreasing r_R to facilitate the threshold condition. Even if initial cytotoxic efforts might not slow down C_R replication, other specific microenvironmental cues, in the form of antiangiogenic [36] or dormancy-inducing [37] drugs targeting cell cycling rate are likely to help the overall *transition therapy* scheme.

C. Targeting PHS in larger architectures ($N > 2$)

We have used the $N = 2$ case to illustrate the concept of cancer growth with switching and how different growth-transition trade-offs can influence therapeutic outcome in simple Sensitive-Resistant scenarios. But tumor architectures often include more than two coexisting phenotypes [12,13] beyond the effects of chemotherapy. Given a larger system with N phenotypes that switch stochastically, can our mathematical framework define the limits of PHS resilience? The analytical approach for $N > 2$ independent phenotypes becomes harder as we add dimensions, and results now depend on N^2 parameters. However, certain average effects of given therapy schemes can be predicted under symmetry assumptions.

Let us here suppose a common therapeutic scheme, where certain phenotypes are sensitive to a first drug, while others tolerate it. This could be the scenario encountered in the development of adaptive resistance to docetaxel (DTX) in breast cancer (N=3, [13,29]) or the targeting of either EGFR, PDGFRA, or CDK4 only affecting one out of four phenotypes in Glioblastoma (N=4, [12]).

The problem can be tackled as follows. Let us first consider the $N = 3$ case, as indicated in figure 3a. In order to reduce the complexity of our calculations, we consider a coarse-graining assumption: all resistant and sensitive cells do so at equal rates, r_R and r_S respectively, and transition rates between replicating and dying cells are also homogeneous. This is summarized in figure 3b.

In this scenario, suppose a system with two phenotypes that replicate at $r_R > 0$ and hold a sensitive phenotype

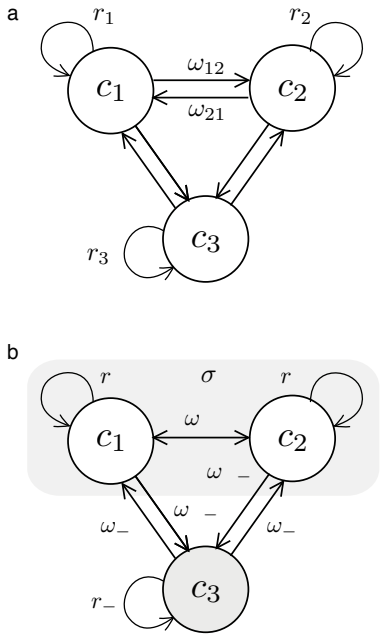

FIG. 4 **Transitions for N=3 phenotypes**. For a $N = 3$ case study, the flow diagram (a) indicates all the transition and replication rates. In order to determine the requirements for successful therapy when a cytotoxic drug is used against C_3, a homogeneous model (b) is used. + subscripts refer to therapy tolerant or resistant phenotypes C_R, while − indicate those phenotypes with negative effective replication under the action of a drug.

$r_S < 0$:

$$\frac{dC_1}{dt} = (r_R - w_{RR} - w_{RS})C_1 + w_{RR}C_2 + w_{SR}C_3 \quad (15)$$

$$\frac{dC_2}{dt} = (r_R - w_{RR} - w_{RS})C_2 + w_{RR}C_1 + w_{SR}C_3 \quad (16)$$

$$\frac{dC_3}{dt} = (r_S - 2w_{SR})C_3 + w_{RS}(C_1 + C_2) \quad (17)$$

Let us now indicate by σ_R the total population of resistant cells, I. e. $\sigma_R = C_1 + C_2$ (figure 3b). In this case, the system reduces to

$$\frac{d\sigma_R}{dt} = \sigma_R r_R - \sigma_R w_{RS} + 2w_{SR}C_3 \quad (18)$$

$$\frac{dC_3}{dt} = C_3 r_R - 2C_3 w_{SR} + w_{RS}\sigma_R \quad (19)$$

For this two-compartment system, it can be shown that the minimal threshold for the resistant population replication rate is:

$$r_R^* = \frac{w_{RS}}{\left(1 + 2\frac{w_{SR}}{|r_S|}\right)}. \quad (20)$$

This calculation, under our homogeneity assumptions, can be done in a systematic way for a switching population with of N cell types (see SM). Specifically, we can

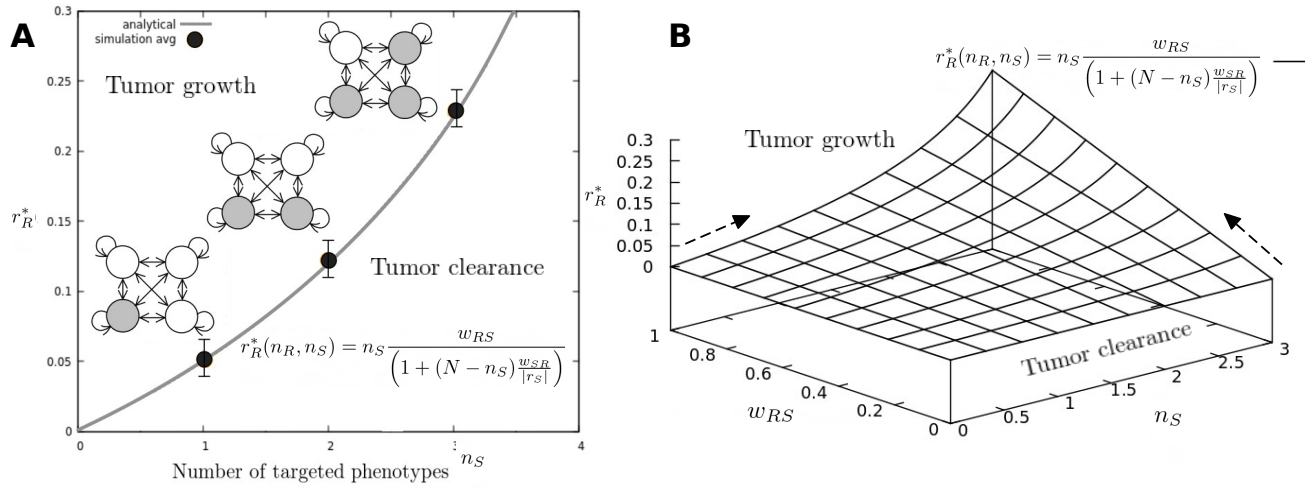

FIG. 5 **PHS therapy in larger architectures ($N > 2$).** In **A**, Replicating phenotypes (empty circles) maintain drug-targeted phenotypes (gray) through stochastic switching. Therapies targeting replication through targeting of genetic drivers such as EGFR in GBM are likely to increase non-linearly the cost for resistance phenotypes to maintain diversity (continuous line, equation displayed in the inset). Stochastic Gillespie simulations result in a certain degree of deviation, where extinction in smaller populations can eventually happen for values of $r > r_R$. Filled dots indicate the value for which 95% of the computational experiments result in population extinction, with errors bars indicating 5% deviations from this value (see SM for computational details). In **B**, the effect of combined transition therapies draining R phenotypes into S is captured by a *therapeutic efficacy* landscape. It can be seen that the effect of adding a transition therapy (w_{RS}) or a novel targeted agent (n_S) is captured by the gradients (dark arrows), directly resulting in an analytical scheme to compute best scenarios for combination therapy in PHS.

consider n_R replicators with a positive effective growth rate r_R and n_S sensitive cell types targeted by therapy, so that their death rate increases beyond birth and $b_S - d_S = r_S < 0$.

By aggregating the two different populations in σ_R and σ_S compartments, the problem of a tumor with N switching phenotypes can be studied (see SM). It can be shown that the minimal growth rate for the positive replicators to sustain the tumor is

$$r_R^*(w_{RS}, n_S) = n_S \frac{w_{RS}}{\left(1 + (N - n_S)\frac{w_{SR}}{|r_S|}\right)}. \qquad (21)$$

Complete cancer eradication can happen if all phenotypes are targeted. Targeting less than four phenotypes can prove useless if the other cell types maintain diversity by replicating faster than (3) (Figure 5). Through sequentially targeting several phenotypes, we can increase n_S and decrease $n_R = N - N_S$ accordingly. This results in a nonlinear increase in the pressure to maintain diversity and growth (Fig. 5A).

The existence of a threshold relating replication (*i.e* drug sensitivity), targeted phenotypes and phenotypic transitions to overall therapy effectivity is consistent with results in [19], where several combinations of drugs inhibiting plasticity-mediated resistance are tested in BRAF mutant melanoma. There is direct correlation between the effect of drugs on transition rates and overall cellular growth, with failure of vemurafenib-only therapy related with r_R overcoming the threshold (21) for all plastic populations [19].

For a GBM setting, the threshold could be potentially exploited by a multi-gene, multi-drug approach able to target the 3 main genetic pathways of the AC-like, OPC-like and NPC-like populations through EGFR, PDGFRA and CDK4 respectively [12]. Each novel target is likely to induce a strong pressure for replication to resistant phenotypes σ_R, eventually resulting in the mesenchymal phenotype alone bearing the pressure of the replication threshold ($n_S = 3$, Fig. 5A). This is a specially relevant result, since it provides a rough estimate of the potential obstacles to replication-oriented therapy posed by the presence of N-dimensional switching.

What is the role of transition rates in therapeutic schemes for $N > 2$? We know from the smaller system (11)-(12) that increasing C_R draining beyond $w_{RS} \geq r_R$ is a necessary condition for tumor eradication. When N phenotypes are at place, the condition for *Transition therapy* to success writes:

$$w_{RS} \geq \frac{r_R}{n_S}\left[1 + (N - n_S)\frac{w_{SR}}{|r_S|}\right] \qquad (22)$$

The multiple phenotypes architecture threshold differs from equation (14) when n_S is considered. This result implies a novel combination therapy landscape, able to characterize overall therapeutic effectivity as a function of the parameter changes occurring after each drug hit

(Fig. 5B). The landscape offers the possibility of computing the gradients (Fig. 5B), dark arrows) that indicate the pressure on r_R^* exerted by either increasing w_{RS} or n_S. In given therapy settings, computing this landscape and its gradients result in a preliminary indicator on choosing if the next drug should focus on draining the untargeted phenotype (w_{RS}) or targeting a novel sensitive phenotype (n_S). Overall, this could improve targeting of multi-phenotype plastic networks where w_{ij} is only targeted so far through inhibition and not increase [19].

The gradients of $\partial_{w_{RS}} r^*$ and $\partial_{n_S} r^*$ therefore indicate a key evolutionary ingredient for combination therapeutic designs

$$\frac{\partial r_R^*}{\partial w_{RS}} = \frac{n_S}{\left(1 + (N - n_S)\frac{w_{SR}}{|r_S|}\right)}, \tag{23}$$

$$\frac{\partial r_R^*}{\partial n_S} = \frac{w_{RS}\left(1 + N\frac{w_{SR}}{|r_S|}\right)}{\left(1 + (N - n_S)\frac{w_{SR}}{|r_S|}\right)^2}. \tag{24}$$

With given parameters, adding single agents should follow from which gradient of both is larger. If not, using drugs that induce small gradient effects on r_R^* is likely to allow resistant phenotypes a window to explore escape mechanisms in the lack of strong drug activity [38].

III. DISCUSSION

Several considerations on therapy design arise directly from the previous results (and our simplifying assumptions). A well-adapted population can maintain non-adapted cell types, provided replication and transition rates are tuned accordingly. Evidence for skewness in experimental transition rate values [13, 19] could indicate the evolution of PHS networks towards enhancing well-adapted phenotypes. PHS offers therefore an alternative pathway to cancer heterogeneity and consequent drug resistance [39]. In this context, single-phenotype strategies are likely to fail, steering tumor evolution towards other phenotypes instead of providing a cure. Our mathematical framework provides a qualitative understanding of such failure for N-dimensional PHS architectures.

In them, what is to be tackled is diversity itself: if only one phenotype can be targeted, the model indicates that others can be drained by increasing the rates at which they transition to the dying one. A key implication here is that inhibition of resistant-phenotype transitions is not necessarily a successful approach if drug-tolerant cell types are not specifically drained towards sensitivity.

Therapeutic strategies that target differentiation pathways are already in place [40], and much is known about dedifferentiation and reprogramming across cell types [40,42]. Clinical and experimental evidence points to differentiation-regulating genes as potential targets of *transition therapy*. Potential examples are TBX3 affecting inter-phenotype switching [13] or SFK/Hck regulating chemotolerance [29] in breast cancer cell lines. Epigenetic drugs targeting DNA methylation are nowadays another therapeutic opportunity [43,44], and combinatorial antibody libraries as regulators of cell fate [45] or stem cell transdifferentation [46] might provide further options to induce phenotypic transdifferentation as a therapeutic strategy. Recent evidence indicates the relevance of obtaining a clear portrait of the underlying Gene Regulatory Networks (GRNs) driving plasticity in order to target possible feedback loops or hysteresis mechanisms of PHS [33,47]. Furthermore, the possibility that phenotypic switching can be targeted beyond oncogenic phenotypes [48] opens up the Waddington landscape to be explored by *transition therapy* [49].

When more than two phenotypes coexist it is likely that several cell types have evolved oncogenic advantage [12,19]. Our approach indicates that targeting of several phenotypes increases non-linearly the pressure for tumor survival. Drug combination targeting multiple cell types together with transition rates to drain non-targeted phenotypes could result in increased benefits for patient survival if specific PHS threshold conditions are fulfilled.

Sequential therapy schemes are known to drive tumor evolution by inducing pressures that drive clonal selection [50]. Even in tumors where phenotypes show self-renewal capacity after cytotoxic therapy, our modeling approach is a predictive tool for the resulting phenotypic trajectories. Since we can compute the stable phenotypic composition for any combination of parameters, knowing how they change after therapy results in a quantitative prediction of the new tumor state.

This can prove helpful to understand tumor evolution after each drug [29,51]. It has been studied for clonal evolution tumor schemes [52], but accumulated knowledge indicates that epigenetic plasticity introduces novel conditions for eradication of resistant cell types [39]. The ability to push the system towards equilibria predicted by our model puts forward the opportunity of directing evolution to pre-sensitize the tumor to a second drug [53]. Following the notion of cancer attractors and combination therapy [54], increasing (instead of only inhibiting) transition rates offers new ways of thinking in how to tackle PHS-driven heterogeneity under a more "developmental" perspective. Future extensions might need to be considered, including gene network regulation, spatially explicit structure, niche architecture and tissue hierarchy. Each extra layer will undoubtedly modify our basic bounds, but we conjecture that the ways PHS influences tumor responses will be basically the same.

Acknowledgments

G. Aguadé-Gorgorió thanks the SFI 2019 crew and all the CSL lab members. S. Kauffman thanks Sui Huang for inspiring conversations on the cancer attractors and the open

possibilities of transdifferentation. Special thanks to J.C. Scott for his inspiring ideas. This work was supported by the Botín Foundation by Banco Santander through its Santander Universities Global Division, the Spanish Ministry of Economy and Competitiveness, grant FIS2016-77447-R MINECO/AEI/FEDER, an AGAUR FI 2018 grant, and the Santa Fe Institute where most of this work was done.

Bibliography

1. Sultan SE. 2000 Phenotypic plasticity for plant development, function and life history. Trends Plant Sci. 5, 537–542. (doi: 10.1016/S1360-1385(00)01797-0)
2. Piggliuci M. 2001 Phenotypic plasticity: beyond nature and nurture. JHU Press, 328pp.
3. Margueron R, Reinberg D. 2010. Chromatin structure and the inheritance of epigenetic information. Nat. Rev. Genet. 11, 285–296.
4. Balalszi G, Oudenaarden AV, Collins JJ. 2011 Cellular Decision Making and Biological Noise: From Microbes to Mammals Cell 144, 910–925. (doi: 10.1016/j.cell.2011.01.030)
5. Kussell E, Leibler S. 2005 Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. Science 309, 2075–2078. (doi:10.1126/science.1114383)
6. Balaban NQ, Merrin J, Chait R, Kowalik L, Leibler S. 2004 Bacterial persistence as a phenotypic switch. Science 305, 1622–1625.
7. Flavahan WA, Gaskell E, Bernstein BE. 2017 Epigenetic plasticity and the hallmarks of cancer Science 357, eaal2380. (doi:10.1126/science.aal2380)
8. Dean M, Fojo T, Bates S. 2005 Tumour stem cells and drug resistance. Nat. Rev. Can. 5, 275–284.
9. Shackleton M, Quintana E, Fearon ER, Morrison SJ. 2009 Heterogeneity in cancer: cancer stem cells versus clonal evolution. Cell 138, 822–829.
10. Marusyk A, Almendro V, Polyak K. 2012 Intra-tumor heterogeneity: a looking glass for cancer? Nat. Rev. Can. 12, 323–334.
11. Brock A, Chang H, Huang S. 2009 Non-genetic heterogeneity–a mutation-independent driving force for the somatic evolution of tumours. Nature Reviews Genetics 10, 336–342.
12. Neftel C, *et al.* 2019 An integrative model of cellular states, plasticity, and genetics for glioblastoma. Cell 178, 835–849.
13. Gupta PB, Fillmore CM, Jiang G, Shapira SD, Tao K, Kuperwasser C, Lander ES. 2011 Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell 146, 633–644.
14. Quintana E, *et al.* 2008 Efficient tumour formation by single human melanoma cells. Nature 456,593–598.
15. O'connell MP, Weeraratna AT. 2013 Change is in the air: the hypoxic induction of phenotype switching in melanoma. Jour. Inves. Dermatol. 133, 2316–2317.
16. Jolly MK, Kulkarni P, Weninger K, Orban J, Levine H. 2018 Phenotypic plasticity, bet-hedging, and androgen independence in prostate cancer: Role of non-genetic heterogeneity. Front. Oncol. 8, 50.
17. Sharma SV, *et al.* 2010 A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell 141,69–80
18. 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. Nature communications, 4(1), 1-11.
19. Su, Y., Wei, W., Robert, L., Xue, M., Tsoi, J., Garcia-Diaz, A., ... & Heath, J. R. (2017). Single-cell analysis resolves the cell state transition and signaling dynamics associated with melanoma drug-induced resistance. Proceedings of the National Academy of Sciences, 114(52), 13679-13684.
20. Diaz Jr LA, *et al.* 2012 The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers. Nature 486, 537–540.
21. Jordan CT, Guzman ML, Noble M. 2006 Cancer stem cells. New Engl. Jour. Med. 355, 1253–1261.
22. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. 2017 Primary, adaptive, and acquired resistance to cancer immunotherapy. Cell 168, 707–723.
23. Gerlee P, Nelander S. 2012 The impact of phenotypic switching on glioblastoma growth and invasion. PLoS computational biology, 8(6).
24. Jolly MK, Ware KE, Gilja S, Somarelli JA, Levine H. 2017 EMT and MET: necessary or permissive for metastasis?. Molecular oncology, 11, 755–769.
25. Mathis RA, Sokol ES, Gupta PB. 2017 Cancer cells exhibit clonal diversity in phenotypic plasticity. Open biology, 7(2), 160283.
26. Gunnarsson E, De S, Leder K, Foo J. 2020 Understanding the role of phenotypic switching in cancer drug resistance. Journal of Theoretical Biology, 110162.
27. Nowak MA. 2006 Evolutionary dynamics: exploring the equations of life. Harvard University Press, 363pp.
28. Roose T, Chapman SJ, Maini PK. 2007 Mathematical models of avascular tumor growth. SIAM review 49, 179–208.
29. Goldman, A., Majumder, B., Dhawan, A., Ravi, S., Goldman, D., Kohandel, M., ... & Sengupta, S. (2015). Temporally sequenced anticancer drugs overcome adaptive resistance by targeting a vulnerable chemotherapy-induced phenotypic transition. Nature communications, 6(1), 1-13.
30. Talpaz, M., Silver, R. T., Druker, B. J., Goldman, J. M., Gambacorti-Passerini, C., Guilhot, F., ... & Sawyers, C. L. (2002). Imatinib induces durable hematologic and cytogenetic responses in patients with accelerated phase chronic myeloid leukemia: results of a phase 2 study. Blood, The Journal of the American Society of Hematology, 99(6), 1928-1937.
31. Berrieman, H. K., Lind, M. J., & Cawkwell, L. (2004). Do beta-tubulin mutations have a role in resistance to chemotherapy?. The lancet oncology, 5(3), 158-164.

32. Tracqui, P., Cruywagen, G. C., Woodward, D. E., Bartoo, G. T., Murray, J. D., & Alvord Jr, E. C. (1995). A mathematical model of glioma growth: the effect of chemotherapy on spatioâtemporal growth. Cell proliferation, 28(1), 17-31.

33. Hari, K., Sabuwala, B., Subramani, B. V., La Porta, C. A., Zapperi, S., Font-Clos, F., & Jolly, M. K. (2020). Identifying inhibitors of epithelialâmesenchymal plasticity using a network topology-based approach. NPJ systems biology and applications, 6(1), 1-12.

34. Tripathi, S., Chakraborty, P., Levine, H., & Jolly, M. K. (2020). A mechanism for epithelial-mesenchymal heterogeneity in a population of cancer cells. PLoS computational biology, 16(2), e1007619.

35. Ramesh, V., Brabletz, T., & Ceppi, P. (2020). Targeting EMT in cancer with repurposed metabolic inhibitors. Trends in Cancer.

36. Ledzewicz, U., & SchÃ¤ttler, H. (2007). Antiangiogenic therapy in cancer treatment as an optimal control problem. SIAM Journal on Control and Optimization, 46(3), 1052-1079.

37. Goss, P. E., & Chambers, A. F. (2010). Does tumour dormancy offer a therapeutic target?. Nature Reviews Cancer, 10(12), 871-877.

38. Liau, B. B., Sievers, C., Donohue, L. K., Gillespie, S. M., Flavahan, W. A., Miller, T. E., ... & Bernstein, B. E. (2017). Adaptive chromatin remodeling drives glioblastoma stem cell plasticity and drug tolerance. Cell stem cell, 20(2), 233-246.

39. Easwaran H, Tsai HC, Baylin SB. 2014 Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. Mol. cell 54, 716–727.

40. de Thé, H. 2018 Differentiation therapy revisited. Nat. Rev. Can. 18, 117.

41. Jopling C, Boue S, Belmonte JCI. 2011 Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. Nature reviews Mol. Cell Biol. 12, 79–89.

42. Huang S. 2009 Reprogramming cell fates: reconciling rarity with robustness. Bioessays 31, 546–560.

43. Gore SD, Baylin S, Sugar E, Carraway H, Miller CB, Carducci M, *et al.* 2006 Combined DNA methyltransferase and histone deacetylase inhibition in the treatment of myeloid neoplasms. Cancer research, 66(12), 6361–6369.

44. Juergens RA, Wrangle J, Vendetti FP, Murphy SC, Zhao M, Coleman B, *et al.* 2011 Combination epigenetic therapy has efficacy in patients with refractory advanced non–small cell lung cancer. Cancer discovery, 1(7), 598–607.

45. Lerner RA, Grover RK, Zhang H, Xie J, Han KH, Peng Y, Yea K. 2015. Antibodies from combinatorial libraries use functional receptor pleiotropism to regulate cell fates. Quarterly reviews of biophysics, 48(4), 389-394.

46. Xie J, Zhang H, Yea K, Lerner RA. 2013. Autocrine signaling based selection of combinatorial antibodies that transdifferentiate human stem cells. Proceedings of the National Academy of Sciences, 110(20), 8099-8104.

47. CeliÃ -Terrassa, T., Bastian, C., Liu, D. D., Ell, B., Aiello, N. M., Wei, Y., ... & Kang, Y. (2018). Hysteresis control of epithelial-mesenchymal transition dynamics conveys a distinct program with enhanced metastatic ability. Nature communications, 9(1), 1-12.

48. Ishay-Ronen, D., Diepenbruck, M., Kalathur, R. K. R., Sugiyama, N., Tiede, S., Ivanek, R., ... & Christofori, G. (2019). Gain fatâlose metastasis: converting invasive breast cancer cells into adipocytes inhibits cancer metastasis. Cancer cell, 35(1), 17-32.

49. Waddington CH (1957) The Strategy of Genes. Allen Unwin, London.

50. Gatenby R, Brown J. 2018 The evolution and ecology of resistance in cancer therapy. Cold Spring Harb. Perspect. Med. 8, a033415.

51. Amirouchene-Angelozzi N, Swanton C, Bardelli A. 2017 Tumor evolution as a therapeutic target. Cancer discov. 7, 805–817.

52. Bozic I *et al.* 2013. Evolutionary dynamics of cancer in response to targeted combination therapy. eLife 2,e00747. (doi:10.7554/eLife.00747)

53. Basanta D, Gatenby RA, Anderson AR. 2012 Exploiting evolution to treat drug resistance: combination therapy and the double bind. Molecular pharmaceutics, 9(4), 914–921.

54. Huang S, Ernberg I, Kauffman S. 2009 Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. Sem. cell develop. biol. 20, 869–876.

55. Huang S, Kauffman S. 2013 How to escape the cancer attractor: rationale and limitations of multi-target drugs. Sem. Can. Biol. 23, 270–278.

4 Discussion

The work presented here has tried to shed some light across several open problems in cancer progression. We have seen how cancer is, among all, an ecological, evolutionary and developmental process, and that complexity abounds along these three axes. In the light of pervasive therapy resistance in advanced tumors there is an urgent need to use the mathematical models of complex biological systems to understand cancer adaptation. As presented along the Objectives and Results sections, we have focused on a wide range of different problems that span the fields of immunology, genome architecture, cancer-microenvironment interactions or epigenetic plasticity. The general framework to approach these issues, however, has departed from the same goal of using minimal models able to capture universal patterns observed in the clinics. Furthermore, the resulting timeline of scientific research has followed a relatively linear process, in which each specific modeling result eventually led to the formulation of new questions.

In a general sense, we can separate the two major research lines of the thesis –and the presented results– in two: the study of genetic instability and its implications on neoantigen landscapes and immune surveillance, and the approach to developmental issues in cancer through the use of ecological models (Fig. 13). These fields, namely immunotherapy and cancer epigenetics, are two main research areas in today's oncology, and it is likely that related therapeutical opportunities arise in the immediate future. In this context, mathematical approaches are needed as a conceptual background for treatment design and prognosis prediction.

By the end of the PhD funding period, we have focused on finding a way to understand how ecology, evolution and development are intertwined while shaping malignant transformation. In particular, we know that cancer is a heterogeneous disease, but the extent of this heterogeneity is somehow vague and spans across different metrics. Cancer is heterogeneous at the genome level, in the corresponding microenvironmental organization or in the way it uses different epigenetic hierarchies to progress [Marusyk et al., 2012]. Are there well-defined motifs behind inter-cancer heterogeneity? And if so, can

we provide a cartography to map the possible pathways of tumor progression? To analyze this, we are building a *morphospace* of human cancers, able to characterize them by how they explore aberrant complexity across ecological, evolutionary and developmental paths (unpublished Results).

4.1 Genetic instability and immune surveillance

We know nowadays that cancer does not make sense *except in the light of evolution* [Dobzhansky, 2013, Greaves and Maley, 2012]. Within this context, evolutionary dynamics follow from the random accumulation of genome alterations that eventually result in a phenotype with differential fitness. A sufficient rate at which individual replicators accumulate changes at the DNA level is, therefore, a key parameter for evolution to take place, and so genome instability is considered nowadays as a basic enabling characteristic for tumorigenesis [Hanahan and Weinberg, 2011]. However, early studies on the so-called *Quasispecies model* established that there are limits, constrained by overall genome size, to the maximum mutation rate for a replicator to stay in place and do not loose self-identity [Eigen and Schuster, 1977]. A prominent example here is that of viruses, entities with minimum genome size that are known to live close to their critical instability levels [Solé and Elena, 2018].

Many cancers are known to present mutational rates surprisingly higher than those of healthy cells [Loeb, 2001]. In this context, research of the early 2000s was able to postulate the possibility that, as for viruses, cancer cells evolved towards critical instability levels, optimizing an eventual trade-off between fast evolution and genome maintainance [Solé and Deisboeck, 2004]. In the lack of recombination mechanisms (except for rare evidence of cell fusion events [Duelli and Lazebnik, 2003]), the effect of the Muller's Ratchet should pervade both the accumulation of deleterious mutations and deleterious instability levels. In this scenario, as tumors progress and more genome stability and repair mechanisms are lost, the question arises on how do tumor cells avoid the error catastrophe.

To understand this, we have used the formalism of *adaptive dynamics* (AD) [Diekmann, 2002], that captures the evolutionary dynamics of stochastic phenotypic traits as a fitness landscape is explored. The crucible of the problem here is that genetic instability, an otherwise fixed parameter in AD models, becomes the evolving phenotypic trait. Mutations in genes accounting for DNA stability result in an increased mutation rate. To understand this nonlinear process, we modify the so-called *canonical equation* of AD to establish a dynamical description of instability rates in cancer, and how these evolve as a function of the underlying mutational landscape involved [Aguadé-Gorgorió and Solé, 2018]. Two potential results arise from solving the canonical equation in different landscape settings. On the one hand, even at high instability levels, sufficient selection for optimal instability values can maintain tumor populations away from the error catastrophe. On the other hand, the equation provides insight into possible mutagenic therapies, and captures basic dynamical differences before and beyond the threshold that indicate possible benefits of mutagenic therapy against trying to hamper instability [Aguadé-Gorgorió and Solé, 2018].

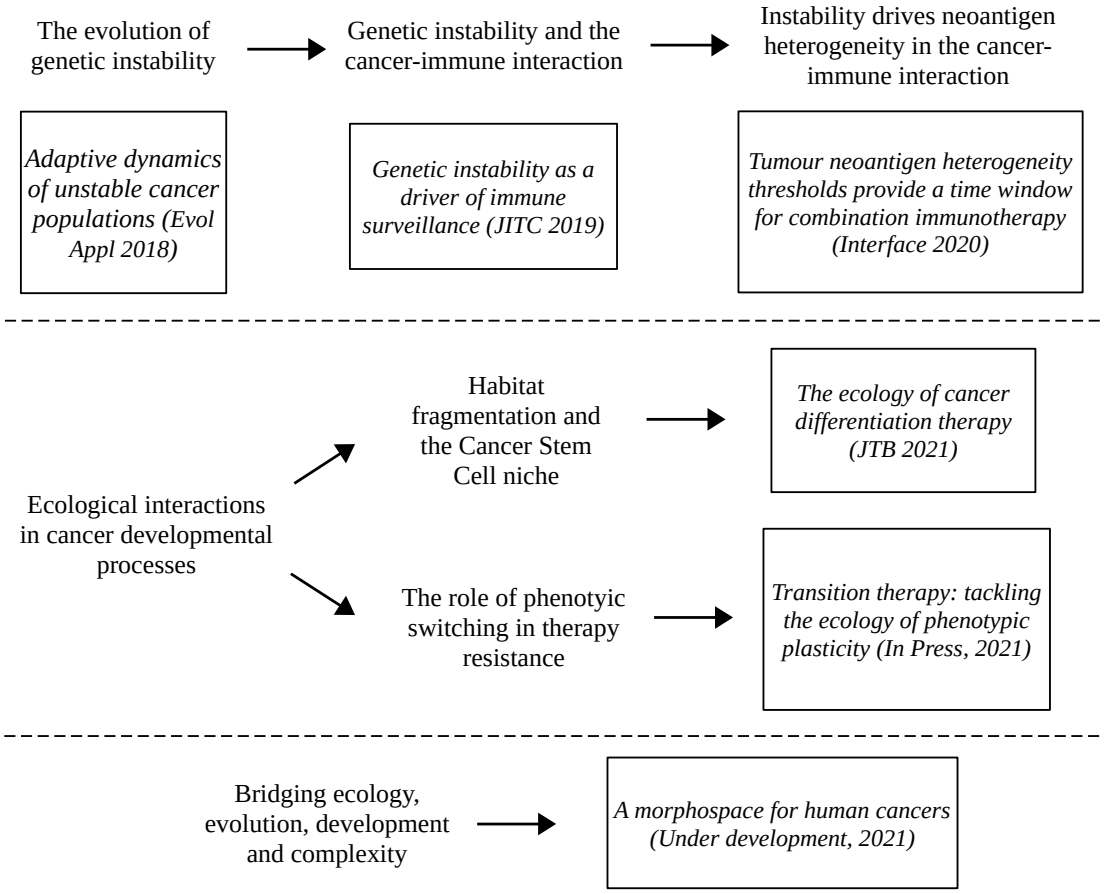

Figure 13: **Research lines within the present thesis.** Mathematical models have been used to approach two major fields in oncology: genetic instability and its implications in immune surveillance, and the role of ecological interactions in epigenetic-related phenomena such as cancer stem cell differentiation therapy or phenotypic switching. By the end of the research period, a comprehensive perspective on the complete eco-evo-devo picture of oncology is being developed, by means of constructing a *morphospace* for human cancer classification.

As I was presenting the AD research project at the *Cancer ecology and evolution* 2018 summer school, a hallway discussion was the evidence that neoantigens played a key role in the immune surveillance of tumors [Schumacher and Schreiber, 2015]. Neoantigens are mutated surface proteins that allow for T cell recognition of cancer cells as *nonself*. As so, the amount of neoantigens in a tumor is in direct correlation with its mutational load, and therefore the underlying genetic instability. This opened up the possibility of another basic trade-off in cancer instability: the same mutations that lead to cancer evolution are targeted by immune cells. How could this translate to a mathematical model?

To face the present question, we used existing research of basic predator-prey approaches that model cancer-T cell interactions [Eftimie et al., 2011]. As for the Lotka-Volterra model, the system here is made of a prey, cancer cells, that explores replication parameters through modulating its instability levels. On the other hand, the predator T

cells recognize cancer cells as these increase their mutation rate [Aguadé-Gorgorió and Solé, 2019]. Taken together, the system shows a set of multiple attractors, out of which we studied cancer clearance, cancer control (by T cells) and large cancer masses. Two main results arise from the model. First, critical transitions from cancer growth to cancer control happen at a mutation rate consistent with experiments of immune surveillance after MMR knockout. This means that T cell efficiency could be enhanced with mutagenic agents. Second, total cancer eradication cannot happen at realistic instability values, and other agents need to be used in combination. A possibility here is that mutagenic compounds are used together with adoptive cell therapy, increasing both T cell efficiency and overall number [Aguadé-Gorgorió and Solé, 2019].

Despite consistent in several settings, the *instability-neoantigens-recognition* pathway might be overly simple in account of cancer complexity. A landmark clinical finding here is that heterogeneity plays a key role in T cell efficiency [McGranahan et al., 2016]. This means that it is not only neoantigens, but the way they are distributed across the tumor, that elicits a sufficient immune response. To study this, we developed a multifaceted research project that could capture together how neoantigen landscapes evolve, and how does their distribution affect T cell recognition. To understand the first, and provided that cancer cells generally escape immune surveillance [Hanahan and Weinberg, 2011, Sharma et al., 2017], we used models of neutral evolution [Kimura, 1983, Williams et al., 2016] and identified a fast decay of neoantigen clonality [Aguadé-Gorgorió and Solé, 2020]. For the second, we used previous research on HIV diversity limits [Nowak et al., 1991] and found the existence of a potential threshold beyond which neoantigens are so heterogenous that T cells fail at providing a specialized response [Aguadé-Gorgorió and Solé, 2020]. Furthermore, this threshold was found to be consistent with clinical data of immunotherapy prognosis in melanoma [Aguadé-Gorgorió and Solé, 2020]. To solve the problem of neoantigen heterogeneity, we propose a combination therapy scheduling where a targeted agent is used in order to induce selection for a resistant subclone, thus enforcing neoantigen homogeneization.

Several open questions arise from our study of immune surveillance in unstable tumors. In particular, current knowledge of the cancer microenvironment tells us that the immune system does not always play an anti-tumor response. What is the role of macrophages here? Evidence indicates that tumors secrete molecules that polarize them towards a pro-tumor (type-2) phenotype (Fig. 3d). How does the whole cancer-lymphocyte picture that we have studied change? A key element here could be that of ecological complex networks [Sole and Montoya, 2001]. As in complex ecosystems (Fig. 3a), immune interactions are so diverse and multi-layered that a network approach is necessary. Could such framework help us compare the possible network states with pro- and anti-tumor immune responses, and detect the therapeutic interventions able to modulate such response?

4.2 Ecological interactions in cancer development

As discussed along the introductory chapter, experimental evidence regarding epigenetic phenomena confronts several aspects of the clonal selection model in cancer [Meacham and Morrison, 2013]. Two main relevant points have been risen here: the existence of Cancer Stem Cells (CSC) at the top of tumor hierarchies [Clarke and Hass, 2006] and the experimental evidence that these hierarchies are in turn highly plastic [Flavahan et al.,

2017]. After the initial research projects on genetic instability and immunotherapy, we have focused on studying how mathematical models could help in understanding aberrant cancer development (Fig. 13.

Interestingly enough, when targeting both issues, we have found ourselves dealing with the mathematical models natural to ecological population dynamics. In this context, it appears that the dynamics of both CSC-fueled cellular hierarchies (Fig. 11) and more plastic architectures (Fig. 12) are grounded on ecological interactions. A consistent body of mathematical models has already been developed for the study of CSCs (see e.g. [Dingli and Michor, 2006]). In them, the main approach recalls using ODEs or PDEs to capture universal patterns of CSC-driven tumors, such as their response to therapy (as compared to that of a clonal evolution tumor) [Dingli and Michor, 2006] or their propensity to reach predictable equilibria between cell compartments [Werner et al., 2016].

As discussed along the introduction, the possibility that many cancers are structured in a maturation hierarchy has reopened the door to differentiation therapy (DTH) [de Thé et al., 2018]. DTH success in APL, a genetically-simple and hierarchically (but homogeneously) organized leukemia [Meng-Er et al., 1988], was not followed by a similar trend in other malignancies, and solid tumors remain a possible candidate for DTH [de Thé et al., 2018]. Within this context, we asked ourselves if there are spatial constraints that play previously unadverted roles in DTH-treated tumors.

The starting point of our DTH model is relatively straightforward: in a spatially-constrained tumor, cells that are not in a replicating compartment still occupy space [Solé and Aguadé-Gorgorió, 2021]. As learned in *1.3.1 Space and resources in tumor growth*, this alone is sufficient to change the overall dynamics of a system, and there is extense literature on the nonlinear effects of habitat loss and fragmentation in complex ecosystems [Huxel and Hastings, 1999]. Within this picture, we outline a minimal hierarchical model, with normal and differentiated cells, and define DTH as affecting the cancer compartment both by differentiating it and affecting the amount of invadable habitat [Bascompte and Solé, 1996]. A key result of the model, consistent with clinical evidence on APL, is that molecular agents that only induce differentiation without cytotoxicity cannot totally eradicate a tumor [de Thé et al., 2018]. What happens when CSCs are introduced? An important element here is that CSC do not, in general, inhabit the same microenvironment as the rest of the tumor [Plaks et al., 2015]. As so, they do not perceive habitat loss in the same manner. This translates into CSC replication obeying softer density-dependent growth constraints, as opposed to those of simple logistic growing populations (Eq. 7). The resulting dynamics of the niche model imply that only extreme (and possibly unrealistic) doses of DTH agents would amount for CSC (and whole tumor) clearance [Solé and Aguadé-Gorgorió, 2021]. Our model introduces a novel framework to understand this effect, pointing towards possible combination treatment options that specifically target the CSC niche and its sensitivity to spatial constraints.

Phenotypic switching (PHS) is a more recent discovery in cancer. Following extensive studies on CSC characterization, it appeared clear that hierarchies, and therefore phenotypic maturation, where not as hard-wired as expected [Chaffer et al., 2011, Batlle and Clevers, 2017]. These initial findings have been followed by extensive literature that demonstrates the existence of epigenetic (non-mutational) phenotypic plasticity as a mechanism for drug tolerance [Sharma et al., 2010], EMT heterogeneity [Kalluri et al., 2009] or overall cancer evolution through the coexistence and cooperation of complex

switching phenotypes [Gupta et al., 2011, Neftel et al., 2019]. As so, phenotypic plasticity is a dynamical mechanism maintaining heterogeneity beyond the clonal selection model. Knowing the relevance of heterogeneity in fueling cancer therapy resistance, there is an urgent need to understand how PHS maintains cellular diversity and growth, and if it provides predictable patterns of tumor evolution different from those of Darwinian evolution.

For this purpose, we have built a simple mathematical framework inspired in mathematical modeling of PHS in drug-tolerant bacteria [Balaban et al., 2004]. During discussions with Stuart Kauffman at the Santa Fe Institute in December 2019, it appeared obvious that even the simplest models where able to capture how PHS maintains cellular heterogeneity. Even in the case where one of the phenotypes is targeted by therapy, the rest of the tumor is likely to hold it in place through stochastic switching [Aguadé-Gorgorió et al., 2020]. This means that any single-agent therapy will fail if PHS is in place.

Furthermore, we were able to extend previous 2-population systems to larger architectures by grouping self-similar populations. This results in a novel tool to understand high-dimensional PHS systems such as Glioblastomas [Neftel et al., 2019] or breast cancers [Gupta et al., 2011]. The $N > 2$ model indicates that there are, again, thresholds that limit the effect of therapies targeting replication. All in all, the model explores the feasibility of so-called *Transition Therapy*, a hypothetical treatment rationale based on combining drugs that increase stochastic switching to drain replicative phenotypes into drug-sensitive ones.

Many questions remain open in the field of cancer epigenetics. A basic point here would be to understand why do some tumors evolve through somatic mutations, others are fueled by CSCs, and others explore their landscape through PHS. Is this a tissue-specific constraint, or do they represent alternative adaptive pathways to choose from? As for the CIN vs MIN issue of genome instability, creating an evolutionary framework able to understand this pattern is a difficult task. However, predicting the evolutionary pathway –and the evolutionary footprints involved– taken by a given cancer would render a major tool for treatment design.

4.3 Merging ecology, evolution and development to understand cancer complexity

So far, we have presented results within the evolution of neoantigens as a key element in the cancer-immune ecology [Aguadé-Gorgorió and Solé, 2019, Aguadé-Gorgorió and Solé, 2020], or the role of ecological dynamics in the development of aberrant tissue hierarchies in tumors [Solé and Aguadé-Gorgorió, 2021]. Interestingly, all the dynamical approaches to cancer complexity and therapy studied along the present thesis involve theoretical aspects of ecology, evolution and development, and cancer progression appears to make sense only in the light of the three.

As discussed along the present work, a major issue in cancer drug resistance is the pervasive heterogeneity across tumor sites [Marusyk et al., 2012]. Cancers are diverse among all possible levels. As a final step in the research scope of the thesis (Fig. 13, we sought to find a conceptual framework to understand cancer heterogeneity at the tissue level. Why are some cancer types rapidly growing and aggressive, while others take decades to progress? Why do some evolve unstable phenotypes and others progress with

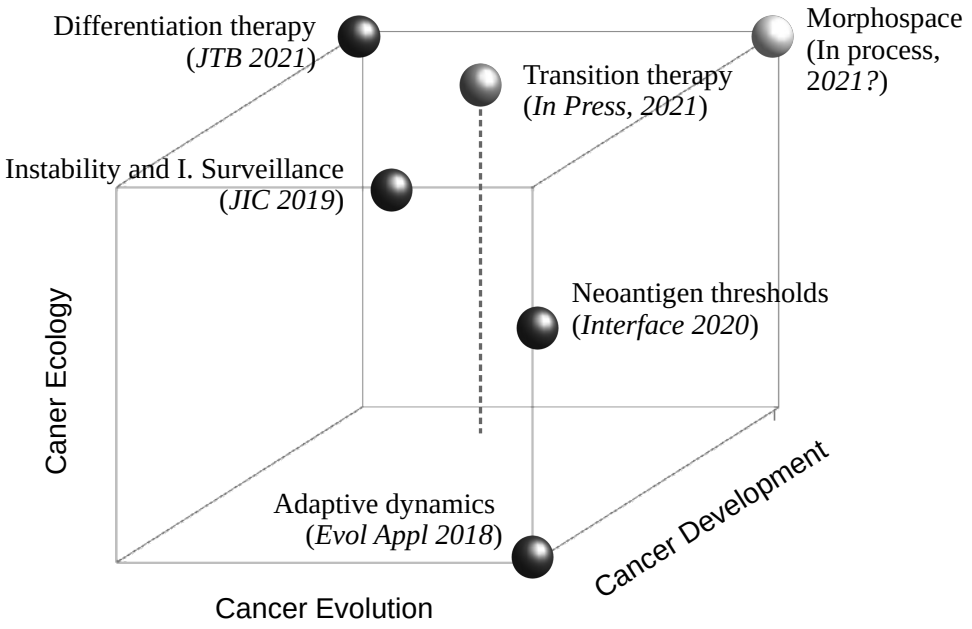

Figure 14: **A toy morphospace for the research projects involved in this thesis**. To understand the conceptual basis of the so-called *morphospaces* [Mitteroecker and Huttegger, 2009], we present a minimal classification scheme capturing which fields of research have been tackled for each Results article (black – published, grey – under revision, white – under development). As expected from the original objectives, ecology, evolution and development are correlated and needed to understand cancer complexity. Only [Aguadé-Gorgorió and Solé, 2018] has studied cancer phenomena through targeting one axis of complexity only (here evolution).

virtually untouched genomes? And more importantly, are there universal motifs across cancer types, implying limits to this heterogeneity? Could we provide a cartography of these?

To address these questions we have made use of the so-called *morphospace*. Morphospaces are tools from evolutionary developmental morphology that provide a geometrical object to organize the heterogeneous forms of individual species or entities [McGhee, 2006]. This is often done through the construction of a 3-dimensional space, not necessarily Euclidean[7], where each axes refers to a given property or quantity of the studied agents. To present a toy example of a conceptual morphospace, it is possible to organize the research articles in the present thesis (Fig. 13) by considering (in a qualitative manner) the extent of ecology, evolution and development involved in their mathematical results (Fig. 14). Despite the present morphospace is obviously metaphoric, it does indicate, for example, that no research has been done considering modeling within the Evolution-Development (and no Ecology) axes.

With Josep Costa from Yale university we started to discuss what happens when can-

[7]Morphospaces can range from strictly euclidean mathematical objects to qualitative metaphors. Despite there is no constraint involved, it is relevant to understand the nature of a given space to understand the validity of the results it presents. A key example here is that of non-euclidean spaces that result from axes built on different metrics. This spaces do not have a proper definition for *distance*, but other objects such as lines connecting entities remain well-defined [Mitteroecker and Huttegger, 2009].

cers are organized following their degree of complex alterations at the three axes. A minimal measure is being studied able to quantify, to a certain extent, the degree of alteration at each complexity axis. It is obviously difficult to gather data that is consistently captured for different tumor types, but some cancer types can be used as a preliminary proof of concept [Guinney et al., 2015].

The resulting geometry of cancer subtype classification is likely to provide clues regarding their possible progression pathways and heterogeneity. Adenoma-carcinoma sequences will often organize as linear trends in space, meaning that one could possibly characterize a tumor only by its position within one or the other regions (and without the need of obtaining genome-specific information on mutational signatures). The most relevant aspect of the cancer morphospace approach regards the existence of empty volums, indicating the possibility that voids in the morphospace characterize non-attainable evolutionary regions. If so, treatment design could benefit from the morphospace cartography by designing therapies able to nudge tumors towards these voids, a result that is being studied at the time of the thesis presentation.

The conceptual approach of the cancer morphospace is likely to produce similar insight into many cancer types. Furthermore, spatial geometries and clustering in the morphospace carry relevant and straightforward meaning, as opposed to data-driven approaches built upon a vast set of genomic variables. Gathering simple Eco-Evo-Devo data for different cancer types to built a 3D classification will yield interesting tumor progression trajectories and, hopefully, universal patterns in the way cancers explore complexity that can help us in better understanding the disease.

4.4 The space of possible combination therapies

Another interesting outcome of the present thesis is that all the mathematical frameworks developed here result in the proposal of combination treatment approaches. Following evidence indicating that single-agent therapies are likely to fail in a variety of settings [Bozic et al., 2013, de Thé et al., 2018, Aguadé-Gorgorió et al., 2020], combining drugs in a rationale that considers tumor complexity is likely our best opportunity of finding durable remissions. To summarize the space of studied therapy designs, we can build a table for the pairwise combination of a set of basic available treatments (Fig. 15). These involve chemotherapy, targeted agents, mutagenesis, epigenetic approaches and immunotherapy.

One interesting result of the present table is the relevance of unfilled spaces. They indicate combination therapies that we have not discussed or have not appeared as feasible solutions for the studied models. In particular, they can be separated in two: options that comprise widely used or redundant strategies (chemotherapy or targeted agents) and options that point towards unknown regions of cancer therapy, mostly epigenetics and its possible combination with immunotherapy or DNA damaging agents (Fig. 15). Regarding the first, the possible role of epigenetic plasticity in immunotherapy remains an open question. However, evidence on immune edition in given CRC subtypes indicates that some tumors could maintain immune-evading phenotypes in place through PHS, something that could be targeted by transition therapy approaches [Guinney et al., 2015, Aguadé-Gorgorió et al., 2020]. On the other hand, the unknown possibility that the Waddington landscape involves limits to phenotypic plasticity (as fitness landscapes do

	Mutagenic	**Chemotherapy**	**Epigenetic**	**Immune**	**Targeted**
Mutagenic	Adaptive dynamics of unstable cancer populations (*Evol Appl 2018*)	Redundant	Futuristic?	Genetic instability as a driver of immune surveillance (*JITC 2019*)	Dangerous (escape pathways)
Chemotherapy		Classic, insufficient	The ecology of cancer differentiation therapy (*JTB 2021*)	Tumour neoantigen heterogeneity thresholds (*Interface 2020*)	Common
Epigenetic			The ecology of cancer differentiation therapy (*JTB 2021*)	Unknown effects	Transition therapy: tackling the ecology of phenotypic plasticity (*In Press, 2021*)
Immune				Genetic instability as a driver of immune surveillance (*JITC 2019*)	Tumour neoantigen heterogeneity thresholds (*Interface 2020*)
Targeted					Unspecific, many options

Figure 15: **Combination treatments studied along the present thesis.** The mathematical models developed along the thesis (and subsequent published articles) have provided a table of which combination treatment options are available, which are interesting to study and which refer to unfeasible or unknown therapeutic methods. In doing so, this table provides hints towards unstudied areas of research indicating potential treatment opportunities. Examples of this are epigenetic treatment in immunotherapy-targeted cancers and combining mutagenic and epigenetic agents to address the possibility of instability thresholds.

for genetic instability [Solé and Deisboeck, 2004]) highlights the opportunity of designing epigenetic-mutagenic approaches able to hamper cell viability in a consistent, multi-scale approach.

5 Conclusions

The conclusions that arise after the present research project can be presented as findings related to the original thesis questions (see Objectives):

Cancer evolution

If there are viability limits to genetic instability [Solé and Deisboeck, 2004], how can cancer cells survive them? In this context, what is the dynamical nature underlying mutation rate evolution? Do cancer cells live at the critical mutation rate, as viruses do [Solé and Elena, 2018]? If so, can lethal mutagenesis be applied as a therapeutic approach to unstable tumors?

- As an extension to the adaptive dynamics framework, we have built a mathematical model where genetic instability is itself a stochastic trait that can evolve. This provides a tool to understand how cancer cells navigate adaptive landscapes by *mutating their mutation rates* [Aguadé-Gorgorió and Solé, 2018].
- The model captures how cancer cells evolve their instability rates to rapidly explore the landscape they are in. It shows the conditions under which purifying selection maintains cancer populations within the critical instability region without trespassing the boundaries of the error catastrophe [Solé and Deisboeck, 2004]. This puts back in place the possibility that mutagenic agents could be used to push cells beyond criticality. Within this therapeutic hypothesis, the model establishes a limit to the viability of treatment related to the capacity of surviving cells to reattain optimal mutation rates [Aguadé-Gorgorió and Solé, 2018].

How does chromosomal instability (CIN) modulate the adaptive landscape of cancer cells? Why do many cancers organize around an average ploidy of $\sim$3.3? Is this an optimal evolutionary value, or a physiological constraint? Are there error thresholds to chromosomal instability?

- The genotype-to-phenotype-to-fitness map is relatively easy to visualize for genomes of fixed size, either through Sewall Wright's *fitness landscapes* [Wright, 1932] or with Kauffman's Boolean networks [Kauffman, 1969] and the underlying *cancer attractors* theory [Huang et al., 2009].

- We have tried, departing from mathematical models of gene duplication [Wagner, 1994], to understand how the evolutionary landscape of cancer cells unfolds when not only genes but chromosome copy numbers are altered. The complexity of both the genetic background and the involved mathematical shape of GRNs hampered us from obtaining simple and relevant results. However, we believe that intuition for possible error thresholds in chromosomal instability could maybe result from this (unfinished) mathematical approach.

- Neutral models of chromosome missegregation have also been developed during the PhD period. In them, optimal ploidy values only result from applying physiological limits in our model (mostly, that cells with more than n copies of a given chromosome are unlikely to survive). Since we do not have experimental data corroborating this, models remain incomplete. We have studied other possible approaches to understand if there exists indeed an evolutionary trade-off justifying optimal ploidy levels (see below).

What is the relation of Whole Genome Doubling (WGD) with the previous question? Why is WGD such a common event in cancer, despite being apparently detrimental to fast cellular proliferation?

- We are currently studying with UPF student Quim Martí the possibility that the pervasiveness of WGD, and the existence of optimal ploidy values could be understood by means of probabilistic models. Within this framework, chromosome copy numbers could have evolved to optimize the equilibria between mutating Tumor Suppressor genes but maintaining House-Keeping genes in place. This project will be part of his TFG as well as a research article.

Why do CIN and micro-satellite instability pathways appear as mutually-exclusive evolutionary mechanisms in cancer [Guinney et al., 2015]

- I believe that CIN is a deleterious pathway for diploid genomes, meaning that extensive copy number aberrations might lead to unbearable genetic loss. However, after WGD events protecting the genome from these events, chromosomal instability is likely to provide a much faster and complex way to explore fitness landscapes. We have not yet been able to fully translate this perspective into a mathematical model.

- However, this buffering effect would not totally explain why cancers that have undergone WGD events do not show mutator phenotypes at the single-nucleotide level (Microsatellite Instability, MIN) [Bielski et al., 2018], but only at the CIN level. Building a mathematical framework able to capture why CIN and MIN appear as mutually exclusive remains an open question that needs further extensive research both at the genetic and mathematical levels.

Cancer ecology

*What is the role of genetic instability in the cancer-immune (predator-prey) interaction? If T cells recognize cancer neoantigens, and these result from accumulated mutations, is there a trade-off balancing cancer evolution and immune recognition? Can mutagenic therapies be combined to checkpoint blockade inhibitors to enhance immune surveillance? (an **Eco-Evo** question)*

- Mutation rates govern a trade-off between cancer adaptation and immune recognition that could explain why some cancers progress and others remit after immunotherapy.

- Departing from previous predator-prey systems [Kuznetsov et al., 1994], genetic instability has been introduced as a variable affecting both cancer exploration of oncogenic mutations and production of neoantigens eliciting immune recognition. The model shows that there is indeed a critical transition between (roughly put) cancer-wins and immune-wins stable states, and that this trade-off is consistent with experimental values for genetic instability levels after Mismatch Repair (MMR) knockout [Aguadé-Gorgorió and Solé, 2019], explaining observations on the role of MMR in immunotherapy efficiency in mice [Germano et al., 2017].

- Furthermore, the model postulates that only tumor dormancy (and not total eradication states) can be achieved by MMR knockout, putting forward the need to combine mutagenic agents with adoptive cell transfer strategies for complete therapeutic success.

*How does heterogeneity in neoantigen landscapes affect the previous question? Can a mathematical model capture why (and how) increased neoantigen diversity correlates with immunotherapy failure? Are there specific limits to neoantigen heterogeneity, and can therapy take advantage of them? (another **Eco-Evo** question)*

- To study the role of neoantigen heterogeneity in immune surveillance, a multilayered mathematical model has been developed that is able to capture how T cell attack on cancer cells depends on both the immunogenic effect of neoantigens and their distribution [Aguadé-Gorgorió and Solé, 2020]. Models and computational simulations have been build to study each.

- A probabilistic approach proves that, in early tumor growth, production of more immunogenic neoantigens grows linear with mutation accumulation.

- In turn, modeling based on foraging strategies in animal movement [Krummel et al., 2016] predicts that immune search efficiency scales linearly with neoantigen clonality, *i.e.* the fraction of the tumor harboring the given antigen.

- As tumors grow, clonal immunogeneicity increases with time, while neoantigen clonality decreases. Combination of both effects in a model of cancer growth under T cell surveillance indicates the presence of a threshold level to neoantigen heterogeneity, similar to that governing HIV progression [Nowak and May, 1991], beyond

which tumor epitopes are so diverse that T cells fail at specializing and targeting them. This is consistent with evidence on neoantigen heterogeneity as a marker for immunotherapy prognosis.

- Inclusion of the model's results on the immunogeneicity–clonality threshold adds predictive value to the prognosis of anti-CTLA-4 therapy in a cohort of melanoma patients [Aguadé-Gorgorió and Solé, 2020]. This reinforces a possible novel biomarker to predict how cancer heterogeneity will affect immunotherapy.

- To avoid the likelihood of tumor escape, a combination therapy approach is designed, where a targeted agent is given prior to immunotherapy to homogeneize the neoantigen landscape. Analytical modeling captures the dynamical tempos of this process, thus providing a framework for therapeutic scheduling.

Is the Warburg effect (the use of inefficient glycolytic metabolism even in the presence of oxygen) an example of a critical transition with hysteresis?

- Apparently not. We have developed mathematical models of population dynamics and game theory to understand the conditions under which glycolytic phenotypes would remain even in the presence of oxygen. The Warburg effect (WE) did not appear as a transition with hysteresis, and our models only justified WE resilience under two possible scenarios that have already been studied: either due to environmental fluctuations in oxygen levels or else because of the acidification of the environment that results in a fitness advantage to surrounding cancer subclones.

Can mathematical modeling help us understand the ecosystem coengineering nature of cancer and type-2 macrophages [Myers et al., 2020]? Provided that M1-M2 equilibria can be restored by therapy, can we predict the conditions under which the immune system will transit from a pro-tumor to an anti-tumor response?

- A mathematical model was developed (see Fig. 7d) to study Th-1 and Th-2 immune responses. The model did show how macrophage polarization could change under certain therapeutic effects, such as inhibition of efferocytosis[8].

- However, the complexity of the population dynamics involved (cancer cells, M1 and M2, and cytotoxic or effector cells) made the model unapproachable by analytical means. This means that the conditions for treatment success are only obtained by computational simulations, and no simple bifurcation landscape can be drawn. Consistent with the overall scope of the thesis, we have not valued complex computational modeling if it does not provide simple and comprehensive results.

What is the ecological nature of Concomitant Resistance? Why do tumors apparently secrete molecular compounds that do not allow secondary tumors to grow? Can a mathematical model capture this apparently indirect competition process? If surgery is successful for small tumors, but awakens large metastatic burdens in large tumor excisions, can we predict the presence of a threshold tumor size with potential therapeutic value?

[8]Efferocytosis is the process by which (type-2) macrophages phagocytize the remnants of apoptotic cancer cells. It has been proposed that blocking this process would led to the damping of cancer compounds to the bloodstream, thus activating a Th-1 response [Myers et al., 2020]

- This project started at the SFI in 2019 with Ricard Solé and is under development with UPF student June Monge-Lorenzo. We believe that a simple mean-field approach of a tumor-tumor interaction is able to capture Concomitant Resistance (CR) dynamics. This follows from understanding how tumors secrete growth-inhibiting molecules to the bloodstream, while produce other compounds that counteract them. Complex diffusion dynamics might justify the local-distant interactions between compounds.

- Once a minimal tumor-tumor model has been used to consistently capture the functional shape of CR, a more complex model will include metastatic seeding and growth. A key question here will be to understand how metastatic dynamics are affected by CR, and how their outgrowth after surgery depends on primary tumor size. This will be part of June's TFG as well as a research article.

Cancer development

*What is the dynamical nature of Differentiation Therapy (DTH) in solid tumors, and why does it usually fail? What role does the CSC niche play here? Can we use spatial ecological models to understand why DTH affects differently each cellular compartment in a hierarchy? (an **Eco-Devo** question)*

- Following early models of habitat loss and fragmentation, we have developed a mathematical model that captures how differentiated cells affect cancer progression by occupying otherwise invadable habitat in the tumor ecosystem [Solé and Aguadé-Gorgorió, 2021]. The model is able to reproduce clinical findings of DTH in APL and other leukemias, while providing interesting insight into the minimal signaling conditions underlying replication–differentiation decisions.

- The model explicitly captures the role of the stem cell niche in DTH, by analyzing how different density-dependent signaling dynamics in cancer stem cells alter the decay of tumor growth under maturation agents. In doing so, we point out towards the relevance of not only using DTH+chemotherapy approaches, but actually targeting the CSC niche itself by disrupting resource consumption or angiogenesis.

*If tumors adapt through stochastic phenotypic switching (PHS) instead of somatic mutation accumulation, what is the nature of therapeutic resistance? Can a simple analytical framework inform about the adaptive potential of PHS versus the clonal evolution model? Would it provide clues for the treatment of multi-phenotype structures such as those of Glioblastomas? (another **Eco-Devo** question)*

- A mathematical framework has been developed to capture the ecological dynamics of phenotypic switching populations (see Fig. 12d) [Aguadé-Gorgorió et al., 2020]. The model demonstrates how therapy resistance obeys different (and much more resilient) pathways than in clonal evolution or CSC-driven tumors.

- In particular, we have been able to demonstrate why single-drug therapies fail in the presence of PHS. This in turn points towards the need to take into account switching rates as a relevant anti-cancer target, opening the possibility of draining replicative phenotypes into drug-sensitive ones (*transition* therapy).

- The analytical approach is extended to larger PHS architectures ($N > 2$). This is relevant to study growth conditions in breast cancer or glioblastoma where PHS expands beyond sensitive-resistant structures. In particular, we derive analytical expressions for treatment efficiency thresholds in complex PHS systems, with direct implications for therapy design. Model results provide clues on which combination treatment (targeting replication or PHS rates) will yield the highest benefit in terms of tumor arrest rates.

Is it possible to develop a mathematical framework that unites Waddington's landscape and Kauffman's cancer attractors to capture the dynamical limits of epigenetic plasticity? Can it account for why (and how) PHS strategies evolve in the first place? Can we understand if PHS architectures have evolved de novo or from an aberrant –but already existing– tissue hierarchy?

- This is an ambitious field of research. By the time of the present thesis, and consistent with our previous research on the role of CIN on unfolding of evolutionary landscapes, it appears clear that the mathematical framework that would capture the nature of epigenetic instability is not a simple object.

- On the other hand, understanding if PHS architectures evolve by creating novel canals in Waddington landscape, or by taking advantage of previously existing ones is probably a question more related to the field of evolutionary genetics, as the models developed so far do not show any dependency on the pathways towards PHS. The fact that many cancers evolve PHS strategies points out to an evolutionary motif, but the patterns of PHS seem unique to each cancer type, meaning they probably respond to tissue-of-origin developmental possibilities and not to landscape-based universal laws.

Overall

Is it possible to build a conceptual framework to understand how ecology, evolution and development are intertwined in cancer? Can we map how different cancers explore complexity along each of these axes? Does this provide a cartography of alternative tumorigenic pathways?

- We have used the notion of *morphospaces* to create a conceptual tool where ecology, evolution and development are merged into a single framework. The final research project is still under development at the time of the publication of this PhD thesis. As a proof of concept, we aim at characterizing the degree of Evo, Eco and Devo across colorectal cancer subtypes. This is likely to result in a 3-dimensional morphospace that captures the complex nature of each cancer subtype together with a map of mutually-exclusive oncogenic pathways that could allow for simple disease classification.

- The so-called *Cancer Morphospace* is likely to provide consistent hints on the possibilities of therapeutic design. In some cancers, a comprehensive review on treatment options for each subtype indicates that therapy could focus on pushing tumors towards the empty regions of the morphospace not explored by other subtypes. These, in turn, probably indicate unviable carcinogenic configurations.

- Developing morphospace classifications for other cancer types, or one that is able to concentrate all possible cancers, is likely to result in novel insight on the topology of disease progression and the extent and boundaries of its heterogeneity.

Bibliography

[Adami, 2002] Adami, C. (2002). What is complexity? *BioEssays*, 24(12):1085–1094.

[Adami, 2004] Adami, C. (2004). Information theory in molecular biology. *Physics of Life Reviews*, 1(1):3–22.

[Aguadé-Gorgorió et al., 2020] Aguadé-Gorgorió, G., Kauffman, S., and Sole, R. (2020). Transition therapy: Tackling the ecology of tumour phenotypic plasticity. *Preprints: https://www.preprints.org/manuscript/202007.0547/v1*.

[Aguadé-Gorgorió and Solé, 2018] Aguadé-Gorgorió, G. and Solé, R. (2018). Adaptive dynamics of unstable cancer populations: The canonical equation. *Evolutionary applications*, 11(8):1283–1292.

[Aguadé-Gorgorió and Solé, 2019] Aguadé-Gorgorió, G. and Solé, R. (2019). Genetic instability as a driver for immune surveillance. *Journal for immunotherapy of cancer*, 7(1):1–13.

[Aguadé-Gorgorió and Solé, 2020] Aguadé-Gorgorió, G. and Solé, R. (2020). Tumour neoantigen heterogeneity thresholds provide a time window for combination immunotherapy. *Journal of the Royal Society Interface*, 17(171):20200736.

[Aguirre-Ghiso, 2007] Aguirre-Ghiso, J. A. (2007). Models, mechanisms and clinical evidence for cancer dormancy. *Nature Reviews Cancer*, 7(11):834–846.

[Al-Hajj et al., 2003] Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J., and Clarke, M. F. (2003). Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences*, 100(7):3983–3988.

[Alberts, 2008] Alberts, B. (2008). Molecular biology of the cell.

[Bailar and Gornik, 1997] Bailar, J. C. and Gornik, H. L. (1997). Cancer undefeated. *New England Journal of Medicine*, 336(22):1569–1574.

[Bak et al., 1987] Bak, P., Tang, C., and Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*, 59(4):381.

[Balaban et al., 2004] Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science*, 305(5690):1622–1625.

[Balkwill et al., 2012] Balkwill, F. R., Capasso, M., and Hagemann, T. (2012). *The tumor microenvironment at a glance*. The Company of Biologists Ltd.

[Barbosa and Castellanos, 2005] Barbosa, P. and Castellanos, I. (2005). *Ecology of predator-prey interactions*. Oxford University Press.

[Bascompte and Solé, 1996] Bascompte, J. and Solé, R. V. (1996). Habitat fragmentation and extinction thresholds in spatially explicit models. *Journal of Animal ecology*, pages 465–473.

[Batlle and Clevers, 2017] Batlle, E. and Clevers, H. (2017). Cancer stem cells revisited. *Nature medicine*, 23(10):1124.

[Baylin et al., 2001] Baylin, S. B., Esteller, M., Rountree, M. R., Bachman, K. E., Schuebel, K., and Herman, J. G. (2001). Aberrant patterns of dna methylation, chromatin formation and gene expression in cancer. *Human molecular genetics*, 10(7):687–692.

[Baylin and Herman, 2000] Baylin, S. B. and Herman, J. G. (2000). Dna hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends in genetics*, 16(4):168–174.

[Bayliss et al., 2016] Bayliss, J., Mukherjee, P., Lu, C., Jain, S. U., Chung, C., Martinez, D., Sabari, B., Margol, A. S., Panwalkar, P., Parolia, A., et al. (2016). Lowered h3k27me3 and dna hypomethylation define poorly prognostic pediatric posterior fossa ependymomas. *Science translational medicine*, 8(366):366ra161–366ra161.

[Becker et al., 1963] Becker, A. J., McCulloch, E. A., and Till, J. E. (1963). Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells.

[Benzekry et al., 2017] Benzekry, S., Lamont, C., Barbolosi, D., Hlatky, L., and Hahnfeldt, P. (2017). Mathematical modeling of tumor–tumor distant interactions supports a systemic control of tumor growth. *Cancer research*, 77(18):5183–5193.

[Benzekry et al., 2014] Benzekry, S., Lamont, C., Beheshti, A., Tracz, A., Ebos, J. M., Hlatky, L., and Hahnfeldt, P. (2014). Classical mathematical models for description and prediction of experimental tumor growth. *PLoS Comput Biol*, 10(8):e1003800.

[Bergers and Benjamin, 2003] Bergers, G. and Benjamin, L. E. (2003). Tumorigenesis and the angiogenic switch. *Nature reviews cancer*, 3(6):401–410.

[Bielski et al., 2018] Bielski, C. M., Zehir, A., Penson, A. V., Donoghue, M. T., Chatila, W., Armenia, J., Chang, M. T., Schram, A. M., Jonsson, P., Bandlamudi, C., et al. (2018). Genome doubling shapes the evolution and prognosis of advanced cancers. *Nature genetics*, 50(8):1189–1195.

[Bonnet and Dick, 1997] Bonnet, D. and Dick, J. E. (1997). Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature medicine*, 3(7):730–737.

[Borges, 1941] Borges, J. L. (1941). The library of babel. *Collected fictions*.

[Borovski et al., 2011] Borovski, T., Felipe De Sousa, E. M., Vermeulen, L., and Medema, J. P. (2011). Cancer stem cell niche: the place to be. *Cancer research*, 71(3):634–639.

[Boucher, 1985] Boucher, D. H. (1985). *The biology of mutualism: ecology and evolution*. Oxford University Press on Demand.

[Bozic et al., 2013] Bozic, I., Reiter, J. G., Allen, B., Antal, T., Chatterjee, K., Shah, P., Moon, Y. S., Yaqubie, A., Kelly, N., Le, D. T., et al. (2013). Evolutionary dynamics of cancer in response to targeted combination therapy. *elife*, 2:e00747.

[Bray et al., 2018] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 68(6):394–424.

[Bruno et al., 2003] Bruno, J. F., Stachowicz, J. J., and Bertness, M. D. (2003). Inclusion of facilitation into ecological theory. *Trends in ecology & evolution*, 18(3):119–125.

[Burnet, 1957] Burnet, M. (1957). Cancer-a biological approach: I. the processes of control. ii. the significance of somatic mutation. *British medical journal*, 1(5022):779.

[Burrell et al., 2013] Burrell, R. A., McGranahan, N., Bartek, J., and Swanton, C. (2013). The causes and consequences of genetic heterogeneity in cancer evolution. *Nature*, 501(7467):338–345.

[Carmeliet, 2005] Carmeliet, P. (2005). Vegf as a key mediator of angiogenesis in cancer. *Oncology*, 69(Suppl. 3):4–10.

[Carroll et al., 2013] Carroll, S. B., Grenier, J. K., and Weatherbee, S. D. (2013). *From DNA to diversity: molecular genetics and the evolution of animal design*. John Wiley & Sons.

[Chabner and Roberts, 2005] Chabner, B. A. and Roberts, T. G. (2005). Chemotherapy and the war on cancer. *Nature Reviews Cancer*, 5(1):65–72.

[Chaffer et al., 2011] Chaffer, C. L., Brueckmann, I., Scheel, C., Kaestli, A. J., Wiggins, P. A., Rodrigues, L. O., Brooks, M., Reinhardt, F., Su, Y., Polyak, K., et al. (2011). Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Sciences*, 108(19):7950–7955.

[Chiarella et al., 2012] Chiarella, P., Bruzzo, J., Meiss, R. P., and Ruggiero, R. A. (2012). Concomitant tumor resistance. *Cancer letters*, 324(2):133–141.

[Clarke and Hass, 2006] Clarke, M. F. and Hass, A. T. (2006). Cancer stem cells. *Reviews in Cell Biology and Molecular Medicine*.

[Coffey et al., 2003] Coffey, J. C., Wang, J., Smith, M., Bouchier-Hayes, D., Cotter, T., and Redmond, H. (2003). Excisional surgery for cancer cure: therapy at a cost. *The lancet oncology*, 4(12):760–768.

[Cohnheim, 1914] Cohnheim, J. (1914). *Ueber entzündung und eiterung*, volume 23. JA Barth.

[Coleman and Tsongalis, 2009] Coleman, W. B. and Tsongalis, G. J. (2009). *Molecular pathology: the molecular basis of human disease*. academic Press.

[Coley, 1893] Coley, W. B. (1893). The treatment of malignant tumors by repeated inoculations of erysipelas: With a report of ten original cases. 1. *The American Journal of the Medical Sciences (1827-1924)*, 105(6):487.

[Collins, 1956] Collins, V. P. (1956). Observation on growth rates of human tumors. *Am J Roentgenol*, 76:988–1000.

[Comen et al., 2011] Comen, E., Norton, L., and Massague, J. (2011). Clinical implications of cancer self-seeding. *Nature reviews Clinical oncology*, 8(6):369–377.

[Coombs et al., 2015] Coombs, C., Tavakkoli, M., and Tallman, M. (2015). Acute promyelocytic leukemia: where did we start, where are we now, and the future. *Blood cancer journal*, 5(4):e304–e304.

[Coussens and Werb, 2002] Coussens, L. M. and Werb, Z. (2002). Inflammation and cancer. *Nature*, 420(6917):860–867.

[Dagogo-Jack and Shaw, 2018] Dagogo-Jack, I. and Shaw, A. T. (2018). Tumour heterogeneity and resistance to cancer therapies. *Nature reviews Clinical oncology*, 15(2):81.

[Darch et al., 2012] Darch, S. E., West, S. A., Winzer, K., and Diggle, S. P. (2012). Density-dependent fitness benefits in quorum-sensing bacterial populations. *Proceedings of the National Academy of Sciences*, 109(21):8259–8263.

[Darwin, 1859] Darwin, C. (1859). On the origin of species by means of natural selection. 1859. *London: Murray Google Scholar*.

[Davidson and Levin, 2005] Davidson, E. and Levin, M. (2005). Gene regulatory networks. *Proceedings of the National Academy of Sciences*, 102(14):4935–4935.

[de Thé et al., 2018] de Thé, H. et al. (2018). Differentiation therapy revisited. *Nature Reviews Cancer*, 18(2):117.

[Delves and Roitt, 2000] Delves, P. J. and Roitt, I. M. (2000). The immune system. *New England journal of medicine*, 343(1):37–49.

[Derényi and Szöllősi, 2017] Derényi, I. and Szöllősi, G. J. (2017). Hierarchical tissue organization as a general mechanism to limit the accumulation of somatic mutations. *Nature communications*, 8(1):1–8.

[Desmazières, 1998] Desmazières, É. (1998). *Onze estampes inspirées de la nouvelle Jorge Luis Borges "La biblioteca de babel" / inventées et gravées park Érik Desmazières*. Paris: Aux dépens de l'artiste.

[Dewhurst et al., 2014] Dewhurst, S. M., McGranahan, N., Burrell, R. A., Rowan, A. J., Grönroos, E., Endesfelder, D., Joshi, T., Mouradov, D., Gibbs, P., Ward, R. L., et al. (2014). Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer discovery*, 4(2):175–185.

[Diaz Jr et al., 2012] Diaz Jr, L. A., Williams, R. T., Wu, J., Kinde, I., Hecht, J. R., Berlin, J., Allen, B., Bozic, I., Reiter, J. G., Nowak, M. A., et al. (2012). The molecular evolution of acquired resistance to targeted egfr blockade in colorectal cancers. *Nature*, 486(7404):537–540.

[Diekmann, 2002] Diekmann, O. (2002). A beginners guide to adaptive dynamics. *Summer school on mathematical biology*, pages 63–100.

[Dingli and Michor, 2006] Dingli, D. and Michor, F. (2006). Successful therapy must eradicate cancer stem cells. *Stem cells*, 24(12):2603–2610.

[Dobzhansky, 2013] Dobzhansky, T. (2013). Nothing in biology makes sense except in the light of evolution. *The american biology teacher*, 75(2):87–91.

[Doolittle, 2013] Doolittle, W. F. (2013). Is junk dna bunk? a critique of encode. *Proceedings of the National Academy of Sciences*, 110(14):5294–5300.

[Duelli and Lazebnik, 2003] Duelli, D. and Lazebnik, Y. (2003). Cell fusion: a hidden enemy? *Cancer cell*, 3(5):445–448.

[Duesberg, 2007] Duesberg, P. (2007). Chromosomal chaos and cancer. *Scientific American*, 296(5):52–59.

[Duffy et al., 2007] Duffy, J. E., Cardinale, B. J., France, K. E., McIntyre, P. B., Thébault, E., and Loreau, M. (2007). The functional role of biodiversity in ecosystems: incorporating trophic complexity. *Ecology letters*, 10(6):522–538.

[Dvorak, 1986] Dvorak, H. F. (1986). Tumors: wounds that do not heal. *New England Journal of Medicine*, 315(26):1650–1659.

[Eftimie et al., 2011] Eftimie, R., Bramson, J. L., and Earn, D. J. (2011). Interactions between the immune system and cancer: a brief review of non-spatial mathematical models. *Bulletin of mathematical biology*, 73(1):2–32.

[Ehrlich, 1909] Ehrlich, P. R. (1909). *The Collected Papers of Paul Ehrlich*, volume 3. Pergamon Press.

[Eigen and Schuster, 1977] Eigen, M. and Schuster, P. (1977). A principle of natural self-organization. *Naturwissenschaften*, 64(11):541–565.

[Eldredge, 2014] Eldredge, N. (2014). Punctuated equilibria: an alternative to phyletic gradualism. In *Time frames*, pages 193–224. Princeton University Press.

[Ezkurdia et al., 2014] Ezkurdia, I., Juan, D., Rodriguez, J. M., Frankish, A., Diekhans, M., Harrow, J., Vazquez, J., Valencia, A., and Tress, M. L. (2014). Multiple evidence strands suggest that there may be as few as 19 000 human protein-coding genes. *Human molecular genetics*, 23(22):5866–5878.

[Fidler, 2003] Fidler, I. J. (2003). The pathogenesis of cancer metastasis: the'seed and soil'hypothesis revisited. *Nature reviews cancer*, 3(6):453–458.

[Flavahan et al., 2017] Flavahan, W. A., Gaskell, E., and Bernstein, B. E. (2017). Epigenetic plasticity and the hallmarks of cancer. *Science*, 357(6348).

[Fox and Loeb, 2010] Fox, E. J. and Loeb, L. A. (2010). Lethal mutagenesis: targeting the mutator phenotype in cancer. In *Seminars in cancer biology*, volume 20, pages 353–359. Elsevier.

[Fulawka et al., 2014] Fulawka, L., Donizy, P., and Halon, A. (2014). Cancer stem cells–the current status of an old concept: literature review and clinical approaches. *Biological research*, 47(1):1–9.

[Gardner, 2002] Gardner, R. (2002). Stem cells: potency, plasticity and public perception. *Journal of anatomy*, 200(3):277–282.

[Gatenby and Brown, 2018] Gatenby, R. and Brown, J. (2018). The evolution and ecology of resistance in cancer therapy. *Cold Spring Harbor perspectives in medicine*, 8(3):a033415.

[Gatenby, 1991] Gatenby, R. A. (1991). Population ecology issues in tumor growth. *Cancer Research*, 51(10):2542–2547.

[Gatenby et al., 2009] Gatenby, R. A., Silva, A. S., Gillies, R. J., and Frieden, B. R. (2009). Adaptive therapy. *Cancer research*, 69(11):4894–4903.

[Gatenby and Vincent, 2003] Gatenby, R. A. and Vincent, T. L. (2003). An evolutionary model of carcinogenesis. *Cancer research*, 63(19):6212–6220.

[Gerlee, 2013] Gerlee, P. (2013). The model muddle: in search of tumor growth laws. *Cancer research*, 73(8):2407–2411.

[Gerlee and Anderson, 2015] Gerlee, P. and Anderson, A. R. (2015). The evolution of carrying capacity in constrained and expanding tumour cell populations. *Physical biology*, 12(5):056001.

[Germano et al., 2017] Germano, G., Lamba, S., Rospo, G., Barault, L., Magrì, A., Maione, F., Russo, M., Crisafulli, G., Bartolini, A., Lerda, G., et al. (2017). Inactivation of dna repair triggers neoantigen generation and impairs tumour growth. *Nature*, 552(7683):116–120.

[Gillies et al., 2008] Gillies, R. J., Robey, I., and Gatenby, R. A. (2008). Causes and consequences of increased glucose metabolism of cancers. *Journal of Nuclear Medicine*, 49(Suppl 2):24S–42S.

[Goldberg et al., 2007] Goldberg, A. D., Allis, C. D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell*, 128(4):635–638.

[González-García et al., 2002] González-García, I., Solé, R. V., and Costa, J. (2002). Metapopulation dynamics and spatial heterogeneity in cancer. *Proceedings of the National Academy of Sciences*, 99(20):13085–13089.

[Gorelik et al., 1981] Gorelik, E., Segal, S., and Feldman, M. (1981). On the mechanism of tumor 'concomitant immunity'. *International journal of cancer*, 27(6):847–856.

[Greaves and Maley, 2012] Greaves, M. and Maley, C. C. (2012). Clonal evolution in cancer. *Nature*, 481(7381):306–313.

[Guinney et al., 2015] Guinney, J., Dienstmann, R., Wang, X., De Reynies, A., Schlicker, A., Soneson, C., Marisa, L., Roepman, P., Nyamundanda, G., Angelino, P., et al. (2015). The consensus molecular subtypes of colorectal cancer. *Nature medicine*, 21(11):1350.

[Gupta and Massagué, 2006] Gupta, G. P. and Massagué, J. (2006). Cancer metastasis: building a framework. *Cell*, 127(4):679–695.

[Gupta et al., 2011] Gupta, P. B., Fillmore, C. M., Jiang, G., Shapira, S. D., Tao, K., Kuperwasser, C., and Lander, E. S. (2011). Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell*, 146(4):633–644.

[Gusev et al., 2000] Gusev, Y., Kagansky, V., and Dooley, W. (2000). A stochastic model of chromosome segregation errors with reference to cancer cells. *Mathematical and computer modelling*, 32(1-2):97–111.

[Hajdu, 2011] Hajdu, S. I. (2011). A note from history: landmarks in history of cancer, part 1. *Cancer*, 117(5):1097–1102.

[Hanahan and Weinberg, 2011] Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *cell*, 144(5):646–674.

[Hanski et al., 1999] Hanski, I. et al. (1999). *Metapopulation ecology*. Oxford University Press.

[Hellström and Hellström, 1969] Hellström, K. E. and Hellström, I. (1969). Cellular immunity against tumor antigens. *Advances in cancer research*, 12:167–223.

[Heyde et al., 2019] Heyde, A., Reiter, J. G., Naxerova, K., and Nowak, M. A. (2019). Consecutive seeding and transfer of genetic diversity in metastasis. *Proceedings of the National Academy of Sciences*, 116(28):14129–14137.

[Hodi et al., 2010] Hodi, F. S., O'Day, S. J., McDermott, D. F., Weber, R. W., Sosman, J. A., Haanen, J. B., Gonzalez, R., Robert, C., Schadendorf, D., Hassel, J. C., et al. (2010). Improved survival with ipilimumab in patients with metastatic melanoma. *New England Journal of Medicine*, 363(8):711–723.

[Holland and Cleveland, 2009] Holland, A. J. and Cleveland, D. W. (2009). Boveri revisited: chromosomal instability, aneuploidy and tumorigenesis. *Nature reviews Molecular cell biology*, 10(7):478–487.

[Horning, 2017] Horning, S. J. (2017). A new cancer ecosystem.

[Huang, 2012] Huang, S. (2012). The molecular and mathematical basis of waddington's epigenetic landscape: A framework for post-darwinian biology? *Bioessays*, 34(2):149–157.

[Huang et al., 2009] Huang, S., Ernberg, I., and Kauffman, S. (2009). Cancer attractors: a systems view of tumors from a gene network dynamics and developmental perspective. In *Seminars in cell & developmental biology*, volume 20, pages 869–876. Elsevier.

[Huang and Ingber, 2007] Huang, S. and Ingber, D. E. (2007). A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks. *Breast disease*, 26(1):27–54.

[Huxel and Hastings, 1999] Huxel, G. R. and Hastings, A. (1999). Habitat loss, fragmentation, and restoration. *Restoration ecology*, 7(3):309–315.

[Huxley et al., 1942] Huxley, J. et al. (1942). Evolution. the modern synthesis. *Evolution. The Modern Synthesis.*

[Jackson and Brues, 1941] Jackson, E. B. and Brues, A. M. (1941). Studies on a transplantable embryoma of the mouse. *Cancer Research*, 1(6):494–498.

[Jain, 2005] Jain, R. K. (2005). Antiangiogenic therapy for cancer: current and emerging concepts. *Oncology (Williston Park, NY)*, 19(4 Suppl 3):7–16.

[Jemal et al., 2011] Jemal, A., Bray, F., Center, M. M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA: a cancer journal for clinicians*, 61(2):69–90.

[Kalluri et al., 2009] Kalluri, R., Weinberg, R. A., et al. (2009). The basics of epithelial-mesenchymal transition. *The Journal of clinical investigation*, 119(6):1420–1428.

[Kauffman, 1969] Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of theoretical biology*, 22(3):437–467.

[Kennedy and Dewhirst, 2010] Kennedy, K. M. and Dewhirst, M. W. (2010). Tumor metabolism of lactate: the influence and therapeutic potential for mct and cd147 regulation. *Future Oncology*, 6(1):127–148.

[Kersey, 1997] Kersey, J. H. (1997). Fifty years of studies of the biology and therapy of childhood leukemia. *Blood, The Journal of the American Society of Hematology*, 90(11):4243–4251.

[Kimura, 1983] Kimura, M. (1983). The neutral theory of molecular evolution. Cambridge Univ Press.

[Krummel et al., 2016] Krummel, M. F., Bartumeus, F., and Gérard, A. (2016). T cell migration, search strategies and mechanisms. *Nature Reviews Immunology*, 16(3):193.

[Kurzrock et al., 1988] Kurzrock, R., Gutterman, J. U., and Talpaz, M. (1988). The molecular genetics of philadelphia chromosome–positive leukemias. *New England Journal of Medicine*, 319(15):990–998.

[Kuznetsov et al., 1994] Kuznetsov, V. A., Makalkin, I. A., Taylor, M. A., and Perelson, A. S. (1994). Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. *Bulletin of mathematical biology*, 56(2):295–321.

[Lakatos et al., 2020] Lakatos, E., Williams, M. J., Schenck, R. O., Cross, W. C., Household, J., Zapata, L., Werner, B., Gatenbee, C., Robertson-Tessi, M., Barnes, C. P., et al. (2020). Evolutionary dynamics of neoantigens in growing tumors. *Nature Genetics*, 52(10):1057–1066.

[Lawrence et al., 2013] Lawrence, M. S., Stojanov, P., Polak, P., Kryukov, G. V., Cibulskis, K., Sivachenko, A., Carter, S. L., Stewart, C., Mermel, C. H., Roberts, S. A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*, 499(7457):214–218.

[Lehrer et al., 1998] Lehrer, M. S., Sun, T.-T., and Lavker, R. M. (1998). Strategies of epithelial repair: modulation of stem cell and transit amplifying cell proliferation. *Journal of cell science*, 111(19):2867–2875.

[Levins, 1969] Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *American Entomologist*, 15(3):237–240.

[Li and Bhatia, 2011] Li, L. and Bhatia, R. (2011). Stem cell quiescence. *Clinical cancer research*, 17(15):4936–4941.

[Lin et al., 2007] Lin, W.-W., Karin, M., et al. (2007). A cytokine-mediated link between innate immunity, inflammation, and cancer. *The Journal of clinical investigation*, 117(5):1175–1183.

[Lloyd et al., 2017] Lloyd, M. C., Gatenby, R. A., and Brown, J. S. (2017). Ecology of the metastatic process. In *Ecology and evolution of cancer*, pages 153–165. Elsevier.

[Lockwood et al., 2009] Lockwood, J. L., Cassey, P., and Blackburn, T. M. (2009). The more you introduce the more you get: the role of colonization pressure and propagule pressure in invasion ecology. *Diversity and distributions*, 15(5):904–910.

[Loeb, 2001] Loeb, L. A. (2001). A mutator phenotype in cancer. *Cancer research*, 61(8):3230–3239.

[Loeb et al., 2019] Loeb, L. A., Kohrn, B. F., Loubet-Senear, K. J., Dunn, Y. J., Ahn, E. H., O'Sullivan, J. N., Salk, J. J., Bronner, M. P., and Beckman, R. A. (2019). Extensive subclonal mutational diversity in human colorectal cancer and its significance. *Proceedings of the National Academy of Sciences*, 116(52):26863–26872.

[López et al., 2020] López, S., Lim, E. L., Horswell, S., Haase, K., Huebner, A., Dietzen, M., Mourikis, T. P., Watkins, T. B., Rowan, A., Dewhurst, S. M., et al. (2020). Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. *Nature genetics*, 52(3):283–293.

[Luria and Delbrück, 1943] Luria, S. E. and Delbrück, M. (1943). Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28(6):491.

[Mack et al., 2014] Mack, S., Witt, H., Piro, R., Gu, L., Zuyderduyn, S., Stütz, A., Wang, X., Gallo, M., Garzia, L., Zayne, K., et al. (2014). Epigenomic alterations define lethal cimp-positive ependymomas of infancy. *Nature*, 506(7489):445–450.

[Maderspacher, 2008] Maderspacher, F. (2008). Theodor boveri and the natural experiment. *Current Biology*, 18(7):R279–R286.

[Margueron and Reinberg, 2010] Margueron, R. and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. *Nature Reviews Genetics*, 11(4):285–296.

[Martincorena and Campbell, 2015] Martincorena, I. and Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science*, 349(6255):1483–1489.

[Marusyk et al., 2012] Marusyk, A., Almendro, V., and Polyak, K. (2012). Intra-tumour heterogeneity: a looking glass for cancer? *Nature Reviews Cancer*, 12(5):323–334.

[Massagué and Obenauf, 2016] Massagué, J. and Obenauf, A. C. (2016). Metastatic colonization by circulating tumour cells. *Nature*, 529(7586):298–306.

[Mbeunkui and Johann, 2009] Mbeunkui, F. and Johann, D. J. (2009). Cancer and the tumor microenvironment: a review of an essential relationship. *Cancer chemotherapy and pharmacology*, 63(4):571–582.

[McComb et al., 2016] McComb, S., Aguadé-Gorgorió, J., Harder, L., Marovca, B., Cario, G., Eckert, C., Schrappe, M., Stanulla, M., von Stackelberg, A., Bourquin, J.-P., et al. (2016). Activation of concurrent apoptosis and necroptosis by smac mimetics for the treatment of refractory and relapsed all. *Science Translational Medicine*, 8(339):339ra70–339ra70.

[McDougall et al., 2006] McDougall, S. R., Anderson, A. R., and Chaplain, M. A. (2006). Mathematical modelling of dynamic adaptive tumour-induced angiogenesis: clinical implications and therapeutic targeting strategies. *Journal of theoretical biology*, 241(3):564–589.

[McGhee, 2006] McGhee, G. R. (2006). *The geometry of evolution: adaptive landscapes and theoretical morphospaces*. Cambridge University Press.

[McGranahan et al., 2016] McGranahan, N., Furness, A. J., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., Jamal-Hanjani, M., Wilson, G. A., Birkbak, N. J., Hiley, C. T., et al. (2016). Clonal neoantigens elicit t cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*, 351(6280):1463–1469.

[Meacham and Morrison, 2013] Meacham, C. E. and Morrison, S. J. (2013). Tumour heterogeneity and cancer cell plasticity. *Nature*, 501(7467):328–337.

[Medzhitov and Janeway, 2002] Medzhitov, R. and Janeway, C. A. (2002). Decoding the patterns of self and nonself by the innate immune system. *Science*, 296(5566):298–300.

[Meng-Er et al., 1988] Meng-Er, H., Yu-Chen, Y., Shu-Rong, C., Jin-Ren, C., Jia-Xiang, L., Lin, Z., Long-Jun, G., and Zhen-Yi, W. (1988). Use of all-trans retinoic acid in the treatment of acute promyelocytic leukemia. *Blood*, 72(2):567–572.

[Merlo et al., 2006] Merlo, L. M., Pepper, J. W., Reid, B. J., and Maley, C. C. (2006). Cancer as an evolutionary and ecological process. *Nature reviews cancer*, 6(12):924–935.

[Miller and Sadelain, 2015] Miller, J. F. and Sadelain, M. (2015). The journey from discoveries in fundamental immunology to cancer immunotherapy. *Cancer cell*, 27(4):439–449.

[Misale et al., 2012] Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Valtorta, E., Schiavo, R., Buscarino, M., Siravegna, G., et al. (2012). Emergence of kras mutations and acquired resistance to anti-egfr therapy in colorectal cancer. *Nature*, 486(7404):532–536.

[Mitteroecker and Huttegger, 2009] Mitteroecker, P. and Huttegger, S. M. (2009). The concept of morphospaces in evolutionary and developmental biology: mathematics and metaphors. *Biological Theory*, 4(1):54–67.

[Morrison et al., 1995] Morrison, S. J., Uchida, N., and Weissman, I. L. (1995). The biology of hematopoietic stem cells. *Annual review of cell and developmental biology*, 11(1):35–71.

[Mosser and Edwards, 2008] Mosser, D. M. and Edwards, J. P. (2008). Exploring the full spectrum of macrophage activation. *Nature reviews immunology*, 8(12):958–969.

[Mukherjee, 2010] Mukherjee, S. (2010). *The emperor of all maladies: a biography of cancer*. Simon and Schuster.

[Myers et al., 2020] Myers, K. V., Pienta, K. J., and Amend, S. R. (2020). Cancer cells and m2 macrophages: cooperative invasive ecosystem engineers. *Cancer Control*, 27(1):1073274820911058.

[Nassar and Blanpain, 2016] Nassar, D. and Blanpain, C. (2016). Cancer stem cells: basic concepts and therapeutic implications. *Annual Review of Pathology: Mechanisms of Disease*, 11:47–76.

[Neftel et al., 2019] Neftel, C., Laffy, J., Filbin, M. G., Hara, T., Shore, M. E., Rahme, G. J., Richman, A. R., Silverbush, D., Shaw, M. L., Hebert, C. M., et al. (2019). An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell*, 178(4):835–849.

[Nguyen et al., 2009] Nguyen, D. X., Bos, P. D., and Massagué, J. (2009). Metastasis: from dissemination to organ-specific colonization. *Nature Reviews Cancer*, 9(4):274–284.

[Norton, 1988] Norton, L. (1988). A gompertzian model of human breast cancer growth. *Cancer research*, 48(24 Part 1):7067–7071.

[Nowak et al., 1991] Nowak, M. A., Anderson, R. M., McLean, A. R., Wolfs, T., Goudsmit, J., and May, R. M. (1991). Antigenic diversity thresholds and the development of aids. *Science*, 254(5034):963–969.

[Nowak and May, 1991] Nowak, M. A. and May, R. M. (1991). Mathematical biology of hiv infections: antigenic variation and diversity threshold. *Mathematical Biosciences*, 106(1):1–21.

[Nowell, 1976] Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, 194(4260):23–28.

[Ohta and Gillespie, 1996] Ohta, T. and Gillespie, J. H. (1996). Development of neutral and nearly neutral theories. *Theoretical population biology*, 49(2):128–142.

[Ollé-Vila et al., 2016] Ollé-Vila, A., Duran-Nebreda, S., Conde-Pueyo, N., Montañez, R., and Solé, R. (2016). A morphospace for synthetic organs and organoids: the possible and the actual. *Integrative Biology*, 8(4):485–503.

[O'Reilly et al., 1999] O'Reilly, M. S., Wiederschain, D., Stetler-Stevenson, W. G., Folkman, J., and Moses, M. A. (1999). Regulation of angiostatin production by matrix metalloproteinase-2 in a model of concomitant resistance. *Journal of Biological Chemistry*, 274(41):29568–29571.

[Paget, 1889] Paget, S. (1889). The distribution of secondary growths in cancer of the breast. *The Lancet*, 133(3421):571–573.

[Pavlova and Thompson, 2016] Pavlova, N. N. and Thompson, C. B. (2016). The emerging hallmarks of cancer metabolism. *Cell metabolism*, 23(1):27–47.

[Pepper et al., 2009] Pepper, J. W., Scott Findlay, C., Kassen, R., Spencer, S. L., and Maley, C. C. (2009). Synthesis: cancer research meets evolutionary biology. *Evolutionary applications*, 2(1):62–70.

[Perelson and Oster, 1979] Perelson, A. S. and Oster, G. F. (1979). Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *Journal of theoretical biology*, 81(4):645–670.

[Perelson and Weisbuch, 1997] Perelson, A. S. and Weisbuch, G. (1997). Immunology for physicists. *Reviews of modern physics*, 69(4):1219.

[Pienta et al., 2008] Pienta, K. J., McGregor, N., Axelrod, R., and Axelrod, D. E. (2008). Ecological therapy for cancer: defining tumors using an ecosystem paradigm suggests new opportunities for novel cancer treatments. *Translational oncology*, 1(4):158–164.

[Pierce and Speers, 1988] Pierce, G. B. and Speers, W. C. (1988). Tumors as caricatures of the process of tissue renewal: prospects for therapy by directing differentiation. *Cancer research*, 48(8):1996–2004.

[Pilosof et al., 2017] Pilosof, S., Porter, M. A., Pascual, M., and Kéfi, S. (2017). The multilayer nature of ecological networks. *Nature Ecology & Evolution*, 1(4):1–9.

[Pinero and Solé, 2019] Pinero, J. and Solé, R. (2019). Statistical physics of liquid brains. *Philosophical Transactions of the Royal Society B*, 374(1774):20180376.

[Pisco et al., 2013] Pisco, A. O., Brock, A., Zhou, J., Moor, A., Mojtahedi, M., Jackson, D., and Huang, S. (2013). Non-darwinian dynamics in therapy-induced cancer drug resistance. *Nature communications*, 4(1):1–11.

[Plaks et al., 2015] Plaks, V., Kong, N., and Werb, Z. (2015). The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? *Cell stem cell*, 16(3):225–238.

[Poulsom et al., 2002] Poulsom, R., Alison, M. R., Forbes, S. J., and Wright, N. A. (2002). Adult stem cell plasticity. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, 197(4):441–456.

[Prehn, 1993] Prehn, R. T. (1993). Two competing influences that may explain concomitant tumor resistance. *Cancer research*, 53(14):3266–3269.

[Quintanal-Villalonga et al., 2020] Quintanal-Villalonga, Á., Chan, J. M., Helena, A. Y., Pe'er, D., Sawyers, C. L., Sen, T., and Rudin, C. M. (2020). Lineage plasticity in cancer: a shared pathway of therapeutic resistance. *Nature Reviews Clinical Oncology*, 17(6):360–371.

[Quintas-Cardama et al., 2009] Quintas-Cardama, A., Kantarjian, H., and Cortes, J. (2009). Imatinib and beyond–exploring the full potential of targeted therapy for cml. *Nature reviews Clinical oncology*, 6(9):535.

[Ramalho-Santos and Willenbring, 2007] Ramalho-Santos, M. and Willenbring, H. (2007). On the origin of the term "stem cell". *Cell stem cell*, 1(1):35–38.

[Read and Strachan, 1999] Read, A. and Strachan, T. (1999). Chapter 18: Cancer genetics. *Human molecular genetics*, 2.

[Recamier, 1829] Recamier, J. (1829). *Recherches sur le Traitement du Cancer*. Paris.

[Retsky et al., 2010] Retsky, M., Demicheli, R., Hrushesky, W., Baum, M., and Gukas, I. (2010). Surgery triggers outgrowth of latent distant disease in breast cancer: an inconvenient truth? *Cancers*, 2(2):305–337.

[Rettig, 2005] Rettig, R. A. (2005). *Cancer crusade: the story of the National Cancer Act of 1971*. iUniverse.

[Reya et al., 2001] Reya, T., Morrison, S. J., Clarke, M. F., and Weissman, I. L. (2001). Stem cells, cancer, and cancer stem cells. *nature*, 414(6859):105–111.

[Riedel, 2005] Riedel, S. (2005). Edward jenner and the history of smallpox and vaccination. In *Baylor University Medical Center Proceedings*, volume 18, pages 21–25. Taylor & Francis.

[Rowley and Corces, 2018] Rowley, M. J. and Corces, V. G. (2018). Organizational principles of 3d genome architecture. *Nature Reviews Genetics*, 19(12):789–800.

[Samuelsson and Troein, 2003] Samuelsson, B. and Troein, C. (2003). Superpolynomial growth in the number of attractors in kauffman networks. *Physical Review Letters*, 90(9):098701.

[Sansregret and Swanton, 2017] Sansregret, L. and Swanton, C. (2017). The role of aneuploidy in cancer evolution. *Cold Spring Harbor perspectives in medicine*, 7(1):a028373.

[Schoener, 1987] Schoener, T. W. (1987). A brief history of optimal foraging ecology. In *Foraging behavior*, pages 5–67. Springer.

[Schumacher and Schreiber, 2015] Schumacher, T. N. and Schreiber, R. D. (2015). Neoantigens in cancer immunotherapy. *Science*, 348(6230):69–74.

[Sell, 2004] Sell, S. (2004). Stem cell origin of cancer and differentiation therapy. *Critical reviews in oncology/hematology*, 51(1):1–28.

[Shackleton et al., 2009] Shackleton, M., Quintana, E., Fearon, E. R., and Morrison, S. J. (2009). Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell*, 138(5):822–829.

[Sharma and Allison, 2015] Sharma, P. and Allison, J. P. (2015). The future of immune checkpoint therapy. *Science*, 348(6230):56–61.

[Sharma et al., 2017] Sharma, P., Hu-Lieskovan, S., Wargo, J. A., and Ribas, A. (2017). Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell*, 168(4):707–723.

[Sharma et al., 2010] Sharma, S. V., Lee, D. Y., Li, B., Quinlan, M. P., Takahashi, F., Maheswaran, S., McDermott, U., Azizian, N., Zou, L., Fischbach, M. A., et al. (2010). A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*, 141(1):69–80.

[Shawver et al., 2002] Shawver, L. K., Slamon, D., and Ullrich, A. (2002). Smart drugs: tyrosine kinase inhibitors in cancer therapy. *Cancer cell*, 1(2):117–123.

[Sica et al., 2008] Sica, A., Larghi, P., Mancino, A., Rubino, L., Porta, C., Totaro, M. G., Rimoldi, M., Biswas, S. K., Allavena, P., and Mantovani, A. (2008). Macrophage polarization in tumour progression. In *Seminars in cancer biology*, volume 18, pages 349–355. Elsevier.

[Sinha et al., 1990] Sinha, A. A., Lopez, M. T., and McDevitt, H. O. (1990). Autoimmune diseases: the failure of self tolerance. *Science*, 248(4961):1380–1388.

[Solé and Aguadé-Gorgorió, 2021] Solé, R. and Aguadé-Gorgorió, G. (2021). The ecology of cancer differentiation therapy. *Journal of Theoretical Biology*, 511:110552.

[Solé and Elena, 2018] Solé, R. and Elena, S. F. (2018). *Viruses as complex adaptive systems*, volume 15. Princeton University Press.

[Solé et al., 2019] Solé, R., Moses, M., and Forrest, S. (2019). Liquid brains, solid brains.

[Solé and Deisboeck, 2004] Solé, R. V. and Deisboeck, T. S. (2004). An error catastrophe in cancer? *Journal of Theoretical Biology*, 228(1):47–54.

[Solé and Goodwin, 2000] Solé, R. V. and Goodwin, B. C. (2000). *Signs of life: How complexity pervades biology*. Basic books New York.

[Sole and Montoya, 2001] Sole, R. V. and Montoya, M. (2001). Complexity and fragility in ecological networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1480):2039–2045.

[Sole et al., 2014] Sole, R. V., Valverde, S., Rodriguez-Caso, C., and Sardanyés, J. (2014). Can a minimal replicating construct be identified as the embodiment of cancer? *Bioessays*, 36(5):503–512.

[Sottoriva et al., 2015] Sottoriva, A., Kang, H., Ma, Z., Graham, T. A., Salomon, M. P., Zhao, J., Marjoram, P., Siegmund, K., Press, M. F., Shibata, D., et al. (2015). A big bang model of human colorectal tumor growth. *Nature genetics*, 47(3):209–216.

[Starr et al., 2003] Starr, T. K., Jameson, S. C., and Hogquist, K. A. (2003). Positive and negative selection of t cells. *Annual review of immunology*, 21(1):139–176.

[Steel, 1977] Steel, G. (1977). Growth kinetics of tumours. clarendon.

[Sterelny and Turney, 2007] Sterelny, K. and Turney, J. (2007). *Dawkins vs. Gould: survival of the fittest*. Icon Books Cambridge.

[Strebhardt and Ullrich, 2008] Strebhardt, K. and Ullrich, A. (2008). Paul ehrlich's magic bullet concept: 100 years of progress. *Nature Reviews Cancer*, 8(6):473–480.

[Strobl et al., 2021] Strobl, M. A., West, J., Viossat, Y., Damaghi, M., Robertson-Tessi, M., Brown, J. S., Gatenby, R. A., Maini, P. K., and Anderson, A. R. (2021). Turnover modulates the need for a cost of resistance in adaptive therapy. *Cancer Research*, 81(4):1135–1147.

[Subramanian et al., 2015] Subramanian, N., Torabi-Parizi, P., Gottschalk, R. A., Germain, R. N., and Dutta, B. (2015). Network representations of immune system complexity. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 7(1):13–38.

[Sudhakar, 2009] Sudhakar, A. (2009). History of cancer, ancient and modern treatment methods. *Journal of cancer science & therapy*, 1(2):1.

[Swan, 1990] Swan, G. W. (1990). Role of optimal control theory in cancer chemotherapy. *Mathematical biosciences*, 101(2):237–284.

[Swartz et al., 2012] Swartz, M. A., Iida, N., Roberts, E. W., Sangaletti, S., Wong, M. H., Yull, F. E., Coussens, L. M., and DeClerck, Y. A. (2012). Tumor microenvironment complexity: emerging roles in cancer therapy.

[Takahashi and Yamanaka, 2006] Takahashi, K. and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *cell*, 126(4):663–676.

[Thomas and Lawrence, 1959] Thomas, L. and Lawrence, H. (1959). Cellular and humoral aspects of the hypersensitive states. *New York: Hoeber-Harper*, pages 529–32.

[Tilman, 1982] Tilman, D. (1982). *Resource competition and community structure*. Princeton university press.

[Topalian et al., 2015] Topalian, S. L., Drake, C. G., and Pardoll, D. M. (2015). Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer cell*, 27(4):450–461.

[Valverde et al., 2018] Valverde, S., Piñero, J., Corominas-Murtra, B., Montoya, J., Joppa, L., and Solé, R. (2018). The architecture of mutualistic networks as an evolutionary spandrel. *Nature ecology & evolution*, 2(1):94–99.

[van der Kooy and Weiss, 2000] van der Kooy, D. and Weiss, S. (2000). Why stem cells? *Science*, 287(5457):1439–1441.

[Van Holde, 2012] Van Holde, K. E. (2012). *Chromatin*. Springer Science & Business Media.

[Visvader and Clevers, 2016] Visvader, J. E. and Clevers, H. (2016). Tissue-specific designs of stem cell hierarchies. *Nature cell biology*, 18(4):349–355.

[Vogelstein and Kinzler, 1993] Vogelstein, B. and Kinzler, K. W. (1993). The multistep nature of cancer. *Trends in genetics*, 9(4):138–141.

[Vogelstein and Kinzler, 2004] Vogelstein, B. and Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature medicine*, 10(8):789–799.

[Waddington, 1942] Waddington, C. H. (1942). The epigenotype. *Endeavour*, 1:18–20.

[Waddington, 1957] Waddington, C. H. (1957). *The strategy of the genes*. Routledge.

[Wagers and Weissman, 2004] Wagers, A. J. and Weissman, I. L. (2004). Plasticity of adult stem cells. *Cell*, 116(5):639–648.

[Wagner, 1994] Wagner, A. (1994). Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. *Proceedings of the National Academy of Sciences*, 91(10):4387–4391.

[Wagner and Zhang, 2011] Wagner, G. P. and Zhang, J. (2011). The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nature Reviews Genetics*, 12(3):204–213.

[Warburg et al., 1931] Warburg, O. H., Dickens, F., et al. (1931). Metabolism of tumours.

[Weidner et al., 1991] Weidner, N., Semple, J. P., Welch, W. R., and Folkman, J. (1991). Tumor angiogenesis and metastasis–correlation in invasive breast carcinoma. *New England Journal of Medicine*, 324(1):1–8.

[Weiss, 2000] Weiss, L. (2000). Metastasis of cancer: a conceptual history from antiquity to the 1990s. *Cancer metastasis reviews*, 19(3-4).

[Werner et al., 2016] Werner, B., Scott, J. G., Sottoriva, A., Anderson, A. R., Traulsen, A., and Altrock, P. M. (2016). The cancer stem cell fraction in hierarchically organized tumors can be estimated using mathematical modeling and patient-specific treatment trajectories. *Cancer research*, 76(7):1705–1713.

[West et al., 2020] West, J., You, L., Zhang, J., Gatenby, R. A., Brown, J. S., Newton, P. K., and Anderson, A. R. (2020). Towards multidrug adaptive therapy. *Cancer research*, 80(7):1578–1589.

[Williams et al., 2016] Williams, M. J., Werner, B., Barnes, C. P., Graham, T. A., and Sottoriva, A. (2016). Identification of neutral tumor evolution across cancer types. *Nature genetics*, 48(3):238–244.

[Wolf et al., 2000] Wolf, J. B., Brodie, E. D., Wade, M. J., et al. (2000). *Epistasis and the evolutionary process*. Oxford University Press, USA.

[Wright, 1932] Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution.

[Yang et al., 2015] Yang, Y. et al. (2015). Cancer immunotherapy: harnessing the immune system to battle cancer. *The Journal of clinical investigation*, 125(9):3335–3337.

[Zheng and Xie, 2019] Zheng, H. and Xie, W. (2019). The role of 3d genome organization in development and cell differentiation. *Nature Reviews Molecular Cell Biology*, 20(9):535–550.

[Zielenski and Tsui, 1995] Zielenski, J. and Tsui, L.-C. (1995). Cystic fibrosis: genotypic and phenotypic variations. *Annual review of genetics*, 29(1):777–807.