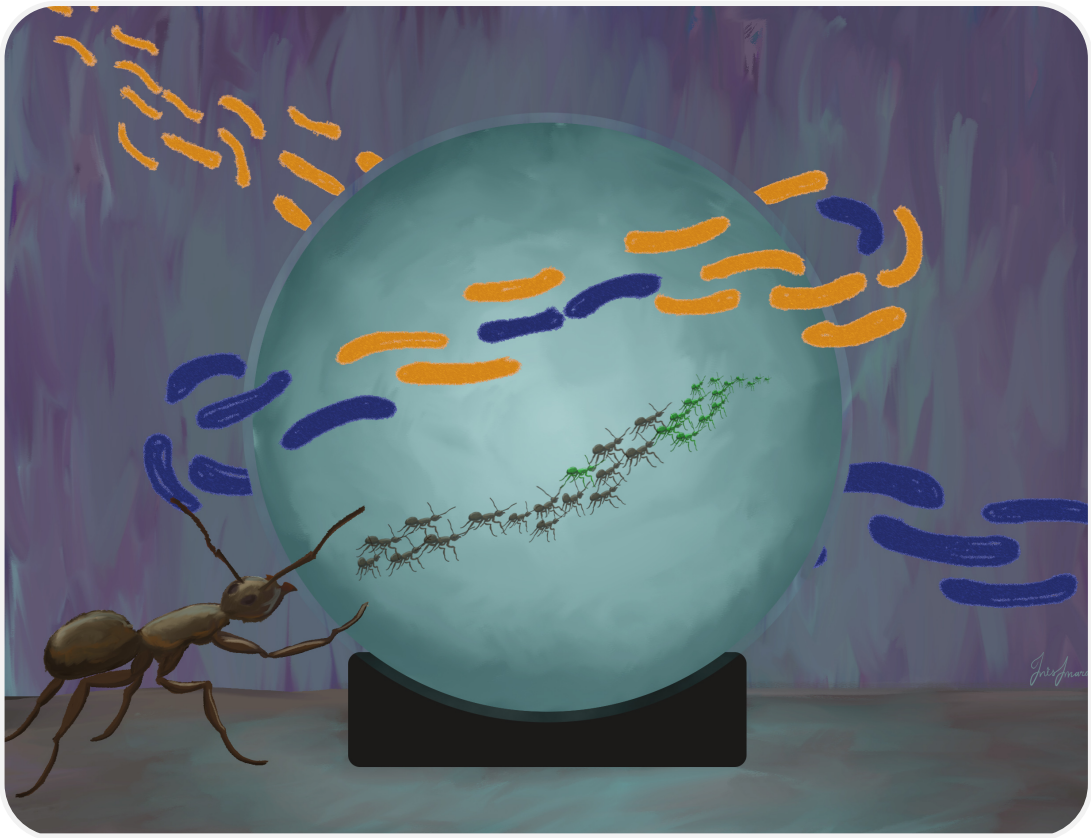
2023

Fitness landscapes for predicting evolution between environments

Ana-Hermina Ghenu



Thesis presented to obtain the **Ph.D degree in**

Integrative Biology & Biomedicine

Oeiras, November, 2023

INSTITUTO DE TECNOLOGIA QUÍMICA E BIOLÓGICA ANTÓNIO XAVIER | UNIVERSIDADE NOVA DE LISBOA

Fitness landscapes for predicting evolution between environments

Ana-Hermina Ghenu

Thesis presented to obtain the Ph.D degree in
Integrative Biology & Biomedicine

Instituto de Tecnologia Química e Biológica António Xavier | Universidade NOVA de Lisboa

Research work coordinated by:

FUNDAÇÃO CALOUSTE GULBENKIAN
Instituto Gulbenkian de Ciência

Oeiras, November, 2023

ITQB NOVA

DOCTORAL THESIS

Fitness landscapes for predicting evolution between environments

Author:
Ana-Hermina GHENU

Supervisors:
Dr. Claudia BANK
Dr. Isabel GORDO

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

in the

Evolutionary Dynamics & Evolutionary Biology Groups
Integrative Biology and Biomedicine PhD Programme
Instituto Gulbenkian de Ciência

Declaration

I, Ana-Hermina GHENU, declare that this thesis titled, "Fitness landscapes for predicting evolution between environments" and the work presented in it are my own. In particular, I declare that:

- This work was done while in candidature for a research degree at this University.
- Where I have consulted or quoted from the published work of others, this is always clearly attributed.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.
- I have followed the rules of scientific good practice.

The Portuguese language abstract was translated from English by Dr. Leila Shirai. The cover art was created by Inês Amaro.

This work was funded by the Fundação Calouste Gulbenkian and the Fundação para a Ciência e a Tecnologia (FCT) PhD funding grant PD/BD/138215/2018 awarded to the student. Additional funding was provided by the FCT PREPARE project JPIAMR/0001/2016-ERA-NET awarded to Isabel Gordo and Claudia Bank; European Research Council (ERC) Starting Grant 804569 (FIT2GO), Human Frontiers Science Program Young Investigator grant no. RGY0081/2020, and Swiss National Science Foundation Project grant no. 315230_204838/1 (MiCo4Sys) awarded to Claudia Bank.

This work was supported by my thesis committee, Dr. Lounès Chikhi and Dr. Erida Gjini, and was successfully defended on the 23rd of November, 2023 in front of the following examiners: Dr. J. Arjan G. M. de Visser, Dr. C. Brandon Ogbunu, Dr. Giulia Ghedini, and Dr. Élio Sucena.

ITQB NOVA

Abstract

Instituto Gulbenkian de Ciência

Integrative Biology and Biomedicine PhD Programme

Doctor of Philosophy

Fitness landscapes for predicting evolution between environments

by Ana-Hermina GHENU

Prediction of evolution is an ambitious undertaking that would consolidate knowledge from all fields of biology for the benefit of global health and biodiversity. Although prediction has been a foundational goal of population genetics theory, this goal is obstructed by the common simplifying assumptions of absent or weak genetic interactions (G×G), gene-by-environment interactions (G×E), and higher-order epistasis-by-environment interactions (G×G×E). This thesis examines the challenges posed by genetic and environmental interactions to the goal of predicting evolution. Fitness landscapes models are brought to bear on data from both wild populations and laboratory conditions in order to investigate the predictability of two pressing issues: species-level biodiversity and antibiotic resistance evolution.

For predicting speciation, we quantify the likely long-term outcomes of a wild, hybrid population of Finnish wood ants. Using a fitness landscape model with conflict between speciation-promoting and speciation-prohibiting forces, we show that there are more possible evolutionary outcomes for haplodiploids as compared to diploids. Our modeling suggests that the hybrid status of the studied Finnish wood ant population will be maintained in the long-term.

For predicting antibiotic resistance evolution, we quantify an

empirical fitness landscape combining antibiotic resistance amino-acid substitutions with whole-gene deletion mutations, in *Escherichia coli*, under two gradients of abiotic environments. The genotype-to-fitness landscapes showed decreasing epistatic effects as antibiotic concentration increased. Therefore, our empirical study shows that evolution may be more predictable in challenging environments as compared to benign ones.

This work is complemented by a methodological study about the common practice of using growth curves to estimate microbial growth rates. Accurate fitness estimates are needed for predicting evolution but we find that there are many opportunities for improvement of growth rate estimates.

In summary, this thesis highlights the sensitivity of organisms to their genetic and environmental conditions, and the theoretical, empirical, and methodological challenges faced in the study of the predictability of evolution. Although these challenges currently impede our efforts towards prediction, integration of knowledge from genetics, integrative biology, and other fields, especially focused on estimating fitness effects, brings us closer to predicting evolution across a broad range of biological systems. Given the inevitability of evolutionary predictions and control, a cross-disciplinary ethical discourse is needed. Therefore, the thesis begins with a history of predicting evolution, from its eugenic origins, and closes with a morally relativistic ethical discussion that invites us to grant agency to the more-than-human world.

ITQB NOVA

Resumo

Instituto Gulbenkian de Ciência

Biologia Integrativa e Biomedicina

Grau de Doutor

Paisagens adaptativas para prever a evolução entre ambientes

por Ana-Hermina GHENU

Prever a evolução biológica é uma tarefa ambiciosa que consolidaria o conhecimento de todos os campos da biologia em benefício da saúde global e da biodiversidade. Embora esta previsão tenha sido um objetivo presente desde a origem da teoria da genética de populações, ele é obstruído por suposições comumente simplificadas de interações genéticas ausentes ou fracas (G×G), interações gene-ambiente (G×E) e interações epistáticas de ordem superior com o ambiente (G×G×E). Esta tese examina os desafios impostos pelas interações genéticas e ambientais ao objetivo de prever a evolução. Os modelos de paisagens adaptativas são baseados em dados de populações selvagens e sob condições de laboratório para investigar a previsibilidade de duas questões urgentes: biodiversidade a nível de espécie e evolução da resistência a antibióticos.

Para prever especiação, quantificamos os prováveis resultados a longo prazo de uma população híbrida selvagem de formigas finlandesas. Usando um modelo de paisagem adaptativa com forças conflitantes que promovem e que não permitem especiação, mostramos que há mais resultados evolutivos possíveis para haplodiploides em comparação a diploides. Nossa modelagem sugere que o estatuto híbrido da população de formigas finlandesas estudada será mantido a longo prazo.

Para prever a evolução da resistência a antibióticos, quantificamos uma paisagem adaptativa empírica combinando mutações

em genes de resistência a dois tipos de antibióticos, substituição de aminoácidos e deleção de genes inteiros, em *Escherichia coli*, sob dois gradientes ambientais abióticos. As paisagens adaptativas genotípicas mostraram efeitos epistáticos decrescentes à medida que a concentração de antibióticos aumentava. Portanto, nosso estudo empírico mostra que a evolução pode ser mais previsível em ambientes desafiadores do que em ambientes mais benignos.

Este trabalho é complementado por um estudo metodológico sobre a prática comum de usar curvas de crescimento para estimar as taxas de crescimento microbiano. Estimativas precisas de aptidão são necessárias para prever a evolução, mas descobrimos que existem muitas oportunidades para melhorar as estimativas da taxa de crescimento.

Em resumo, esta tese destaca a sensibilidade dos organismos às suas condições genéticas e ambientais, além dos desafios teóricos, empíricos, e metodológicos enfrentados no estudo da previsibilidade evolutiva. Embora esses desafios impeçam atualmente nossos esforços de previsão, a integração do conhecimento da genética, fisiologia e outros campos da biologia, especialmente aqueles focados na estimativa dos efeitos adaptativos, nos aproxima de prever evolução para uma ampla gama de sistemas biológicos. Dada a inevitável extrapolação de previsões a controle evolutivo, é necessário um discurso ético interdisciplinar. Portanto, a tese começa com uma história da previsão da evolução, desde suas origens eugênicas, e termina com uma discussão ética moralmente relativista que nos convida à diligência do mundo sobre-humano.

Contents

* References are included at the end of each chapter.

List of Abbreviations

ABR	**A**nti**b**iotic **r**esistance
ANOVA	**An**alysis **o**f **va**riance
AUC	**A**rea **u**nder the **c**urve
BDMI	**B**ateson-**D**obzhansky-**M**uller **i**ncompatibility
DNA	**D**eoxyribo**n**ucleic **a**cid
E. coli	*E**scherichia coli***
F. aquilonia	*F**ormica aquilonia***
F. polyctena	*F**ormica polyctena***
G×E	**G**enotype–by(×)–**E**nvironment interactions
GFP	**G**reen **f**luorescent **p**rotein
G×G	**G**ene–by(×)–**G**ene interactions (epistasis)
G×G×E	**G**ene–by(×)–**G**ene–by(×)–**E**nvironment
KO	**K**nock-**o**ut mutation (whole-gene deletion)
KYA	Thousands of **y**ears **a**go
marR	**M**ultiple **A**B**R** DNA-binding transcriptional repressor
mCherry	A type of red fluorescent protein
nuoC	**N**ADH:q**u**inone **o**xidoreductase subunit **C**D
OD	**O**ptical **d**ensity
RNA	**R**ibo**n**ucleic **a**cid
rpoB	**R**NA **po**lymerase subunit **b**eta
SNP	**S**ingle-**n**ucleotide **p**olymorphism
waaP	Lipopolysaccharide core heptose (I) kinase
WT	**W**ild-**t**ype genotype
ybfG	Pseudogene of PF10964 domain-containing protein
yidK	Putative transporter (Solute:Sodium Symporter)

For my dad, who emigrated before he had a chance to defend his PhD.

And for everyone who supported me.

But most of all:
For myself, who didn't think I would make it.

Chapter 1

Introduction

There are no new ideas. There are only new ways of making them felt.
– Audre Lorde

Contents of Introduction Chapter

1 Introduction

When I was a teenager there was a docu-fiction TV show, *The Future is Wild*, that vividly documented the creatures that might inhabit the earth millions of years from now, animating in detail their future lifestyles and adaptations (de Lespinois and Coules, 2003). Fiction writers have been entertaining non-scientist audiences with such depictions of speculative evolution at least since Herbert George Wells' publication of *The Time Machine* in 1885 (Wikipedia, 2022). But even as Wells was writing his novella, biologists had already begun to work on predicting evolution. The ability to predict evolution has captured the imagination of non-scientists and scientists alike since the Victorian era because of the almost godly powers it would give us over biological systems. Predicting evolution would be useful for medical applications (like combating cancer and antibiotic resistance), conservation, responding to climate change, and many much more fanciful ideas that fiction writers are more qualified than I to dream up.

While science fiction writers like those of *The Future is Wild* agree that we cannot predict what will evolve millions of years from now, the principles of evolution as described by population genetics theory coupled with the recent deluge of available genomic data suggest that it may be possible to quantitatively predict – and even control – where evolution is headed. Contemporary evolution is relatively unique among the natural sciences for having such a historical and predictive, backward and forward, Janus-facing characteristic. This uniformitarianist view of the evolutionary process dates back to the work of Charles Darwin (1859). More broadly, however, prediction and control of the natural world were stated goals of the Scientific Revolution that modern Western science often continues to perpetuate (Lloyd, 1993; Leiss, 1994; Keller, 1996), as the prevalent science-pessimism of the social sciences and humanities tells us. Predicting evolution has a fraught history that includes the eugenics movements of the 19th to mid-20th century. However, this history is often examined in a shallow way by non-scientists (e.g., eugenics is evil therefore any historical or methodological associations with eugenics are necessarily bad too) and either ignored or minimized by scientists. The recent scientific enthusiasm for predicting evolution (Lässig, Mustonen, and Walczak, 2017; Nghe et al., 2020; Burford Reiskind et al., 2021; Thompson et al., 2021; Wortel et al., 2022; Baltazar-Soares, Brans, and Eizaguirre, 2022; The Royal Society, 2023) and examples of evolutionary predictions already being applied (Agor and Özaltın, 2018; Nasr et al., 2019) indicate that evolutionary predictions are an inevitability. As such, it is critical that scientists substantively re-contextualize the role of the eugenics movements on their history and begin to engage in cross-disciplinary discussion about the ethics of predicting evolution.

My thesis work addresses the challenges posed by genetic and environmental interactions to the goal of predicting evolution under the classical population genetics paradigm. This work uses fitness landscapes theory and data to extend the population genetics paradigm in service of predicting evolution. In the first part of this introductory chapter, I trace the history of predicting evolution to show that evolutionary prediction has been a main motivator for the study of population genetics since the beginning of the discipline. Then, I highlight current challenges in predicting evolution, focusing on the challenge posed by genetic and environmental interactions. The final part of this chapter outlines the five questions of my work and introduces the remaining chapters of my thesis. At the time of writing, I feel more enthusiastic about the history / philosophy of predicting evolution and less enthusiastic about my contributions to the science of predicting evolution, probably because I feel the latter

contributions are small, lost within the deluge of scientific publications, while the former go against the flow, having the potential to shift academic paradigms both within and outside of science. I have tried my best to fulfill the criteria for completion of a PhD thesis in the sciences but I know that my enthusiasm is mismatched with my thesis title. I hope the reader can accommodate this.

2 Defining evolutionary predictions and the challenge of interactions

Evolution is Janus-faced (figure 1). The same forces that were acting in the past (selection, mutation, recombination, drift, and migration) will act in the future to accumulate changes that are recorded in the homologous phenotypes and genotypes between extant groups, as well as in fossil remains. At the geological timescale, environmental and genomic conditions change so substantially that, on the forward-facing side, evolutionary predictions are beyond our visual horizon and, on the backward-facing side, evolutionary explanations become uncertain (Ghenu et al., 2016). But at shorter evolutionary time-scales, one or a few generations away, environmental and genomic conditions are sufficiently similar to those of yesterday and today that it may be possible to glimpse into the future. Because of the nature of the world and our conceptual understanding of it, our ability to understand the evolutionary past informs our ability to predict the evolutionary future. Better understanding into past and present evolutionary processes will inevitably lead to an improved understanding of the evolutionary future. However, I argue that predicting evolution is a separate, but related, enterprise from studying/explaining evolution. The field of predicting evolution has a tacit goal of turning evolutionary understanding into evolutionary application. A more controversial way to phrase this is that predicting evolution seeks to turn evolution science into evolution technology.

Science is a knowledge-making discipline whose unique aim is to find explanations for observed phenomena (Godfrey-Smith, 2003). Technology, or applied science, on the other hand, seeks to solve problems (Bunge, 1966). In order to separate when predicting evolution is science and when it is technology, I follow Wortel et al., 2022 in distinguishing between three different aspects of the field of predicting evolution: evolutionary predictions in service of improving our scientific explanatory ability, evolutionary predictions of particular outcomes in service of human preparation (e.g., waiting time to evolution of antibiotic resistance in a particular patient), and

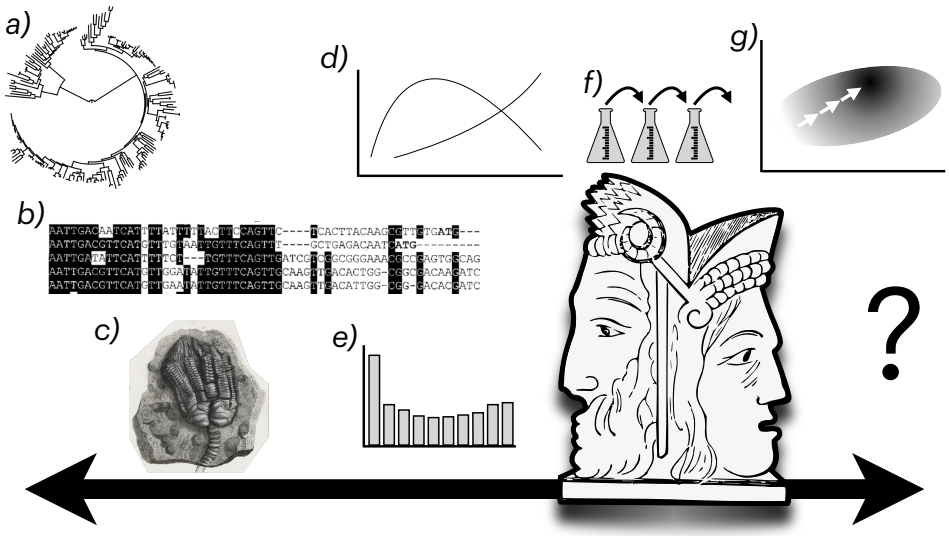


Figure 1: **Evolution is Janus-faced, affording a long view into the past (left-facing arrow) and a brief glimpse into the future (right-facing arrow).** The panels are meant to represent different data and disciplines of evolution, with their arrangement roughly indicating the time-frame that each focuses on. *a)* Phylogenetic methods allow us to see into the past, from recent population diversification events to the last universal common ancestor of all life on Earth. These methods can be informed by homologous phenotypes from *b)* genetic/genomic data as well as *c)* fossil data. Population genetics looks into the recent past to understand how *d)* allele frequencies have changed over time and *e)* how genetic variation arose then was maintained or lost. Contemporary population genetics can also inform us about how evolution is shaping populations in the present moment. *f)* Experimental evolution observes populations evolving in real-time to the imposed experimental conditions. *g)* Shows a phenotype-to-fitness landscape where the x- and y-axes are two (weakly) correlated traits and the darker shading indicates higher fitness. This is an example of integrative biology data that can inform us about the future. The question mark indicates all the ways that the aforementioned disciplines and data can help inform our understanding of the evolutionary future.

evolutionary predictions in service of control (e.g., of populations that cause human disease). The first one is science and the latter two apply evolutionary science as a technology. In practice, this strict distinction between science and technology is difficult to delineate, both because predictions of particular outcomes are used for hypothesis-testing in science and because the resultant technology of evolutionary control would be neither commercial nor industrial.[1] In conclusion, it is easier to

[1] As optimistic as I may be about our ability to predict evolution, I don't see commercial or industrial

distinguish between science and technology in their intent (Bunge, 1966; Valentine, 2021). Science seeks to explain while technology seeks to impact the world for the betterment of human life.

Let us take a closer look at the relationship between explanation (i.e., the purpose of science) and prediction (i.e., something common to both science and technology). A common argument among evolutionary biologists in favour of predicting evolution is that predictions can be used to test the validity of our explanations (Rull, 2022; Wortel et al., 2022), in other words to provide falsifiable hypotheses for evolutionary theories (Popper, 1962, but see Elgin and Sober, 2017). While this may be true, prediction alone is not a sufficient criteria for a good scientific explanation (Scriven, 1962; Salmon, 2006). Successful prediction of an event could be due to correlation, not causation. Moreover, not all predictions are scientific explanations: if I empirically try most of the figs from a fig tree and they all taste bad, then I can predict with high probability that the next fig I try will also taste bad. But this prediction is not an explanation for why the tree makes bad figs, it is only an inference about a specific tree and/or season (Salmon, 2006). The field of statistics distinguishes clearly between explanation and forecasting (Shmueli, 2010). Because explanation and forecasting are different tasks, different models or even different types of data may be needed (Levins, 1966; Huneman, 2014; Shaw, 2019). In particular, when predicting evolution is intended to be used as a technology to prepare us for the most likely future outcomes, accurate predictions are needed. Accurate predictions require high precision and high realism, consequently such predictions would likely come with a cost of low generality. On the other hand, when the technology of predicting evolution is intended to be used for evolutionary control, a detailed and general explanatory / mechanistic understanding is needed such that appropriate measures are taken to push the system into a different behaviour. In summary, evolutionary prediction is a separate but related endeavour from evolutionary explanation as part of science. At least, if we want to focus on evolutionary science, evolutionary predictions can be the new horizon to aim for if we believe that we have a sufficiently robust and comprehensive understanding of evolutionary processes. Many works on predicting evolution, however, focus on the application of evolutionary theory for solving problems out in the world (Lässig, Mustonen, and Walczak, 2017; Loga and Gerlinger, 2017; Russell and Jong, 2017; Shaw, 2019; Nghe et al., 2020; Lässig and Mustonen, 2020; Burford Reiskind et al., 2021; Thompson et al., 2021). This means that

applications possible within my lifetime. I am used to thinking about technology as commercial and industrial but there are other types of technologies, too. The study of what is technology constitutes its own academic discipline; see: Ihde, 2004.

much of predicting evolution is technological development.

I focus on population genetic forecasts of evolution, in other words descriptions of the future state of a population. Following Wortel et al., 2022, this includes the future frequency of alleles or strains in the population; the direction, rate, and limit of trait evolution; the future mean fitness of the population; the expected waiting time until the evolution of a trait; etc. Sometimes I refer to these types of predictions as **micro-evolutionary** to emphasize the relatively short time scale of the forecast. This is to distinguish them from macro-evolutionary predictions like those found in paleontology (Gould, 1977), which are outside the scope of this work.

The presence of genetic and environmental interactions, meaning interactions among genes (G×G), between genotype and environment (G×E), and higher order interactions (like G×G×E), make it difficult to predict evolution via the contemporary paradigm of population genetics. Although the study of G×G and G×E interactions predates the union of Darwinian natural selection with Mendelian genetics, population genetics was established and flourished under the simplifying assumption that these types of interactions are negligible or weak. Therefore, when strong G×G, G×E, and/or G×G×E interactions are present, the phenotype and/or the fitness of the organism will depend not just on the focal alleles, but also on the genetic background and the environment, contradicting the assumptions of population genetics theory.

3 The history of predicting evolution and the challenge of interactions

I tell the history of predicting evolution in part because I think the history of science is interesting. Given the role of historical contingency in scientific assumptions and discoveries, it is useful for scientists to be familiar with the history of their fields. More importantly, however, I tell the history of predicting evolution because I could not find any previous publications about this subject. What did Darwin, the biometricians, the early mendelians, and the fathers of the modern synthesis have to say about evolutionary predictions?, I wondered. In trying to understand if the idea is new, I did a Google scholar search for the term ‘predicting evolution’ between 1850-1960 and I came up with only one hit, a paper titled “Perspectives in population eugenics”

(Allen, 1955). In the back of my mind[2] I suspected that the history of predicting evolution may include the eugenics movement, but I had hoped to be able to avoid the controversial and complex topic of eugenics altogether. However, after reading this 1955 paper, it became clear that the two are inseparable. More importantly, it became clear that eugenics and predicting evolution share more than a history. I argue that predicting evolution and eugenics share five key similarities:

1. a common goal of investigating the future-facing side of evolution;
2. a need to quantify populations;
3. a rallying call for a deeper functional understanding of heredity and consolidation of knowledge across biological disciplines;
4. technological application of evolutionary theory; and
5. these technological solutions are proposed in response to a perceived (genetic or environmental) deterioration that poses an existential threat to human populations.

[NB: I'm not certain that these five similarities are exactly spot-on. For example, the first is trivial, likewise the second and third can be seen as descriptions of the general positivist program of contemporary science/biology and so are common with all population genetics theory and genetics. But I think it may be okay as a first-order approximation.]

I trace the history of predicting evolution in a roughly chronological order. I start with Darwin's resistance to the possibility of predicting evolution and the subsequent enthusiasm for the topic by the early eugenic scientists of the biometric school. Then there is an interlude where I summarize the history of environmental and genetic interactions. This is important because, ostensibly, the purpose of this introduction is to present the challenges posed by interactions to predicting evolution. Finally, I tie these threads together in the history of the modern synthesis to understand what role eugenics played and to understand why interactions still pose a challenge today. This history is very brief and heavily biased in favour of the English-language scientific literature, particularly towards England.

[2]This is a figure of speech. My emotional reactions and persistent thought intrusions told me that the history of predicting evolution includes the history of eugenics. But my scientific training has taught me that 'good science is apolitical' and I was trying to defend a good, respectable scientific PhD thesis. I will return to these philosophical ideas in the Discussion Chapter.

3.1 Predicting evolution before Mendel

Theories of evolution are common throughout human histories and cultures but the most authoritative scientific theory of evolution arose in mid-19th century Victorian England. This moment in history offered the right combination of: a socially dominant creation story (i.e., Abrahamic) that separates humans from other animals and living from non-living matter; uniformitarianist ideas of universal laws of nature as developed during the Scientific Revolution; increasing regard for materialistic, as opposed to theistic, explanations (and decreasing ecclesiastical power) due to Renaissance philosophical movements; rising idealization of meritocracy as a result of greater wealth mobility under recent industrial capitalism and the rising power of the professional middle-class; and more opportunity to travel the world as a result of colonial exploration (MacKenzie, 1979; Levins and Lewontin, 1985; Brown, 1986; Gould, 1993; Hubbard, 2003; Gross and Averill, 2003). I don't have time to unpack that list but the main point of this paragraph is to acknowledge that science happens within particular material and ideological historical conditions. Evolution, like all scientific knowledge, is both shaped by various social factors **and** describes real phenomena that exist out in the world.

The history of predicting evolution as a scientific undertaking can be traced at least as far back as the work that directly precipitated the theory of evolution by natural selection. Thomas Robert Malthus' book, *An Essay on the Principle of Population*, led both Charles Robert Darwin and Alfred Russel Wallace to the realization that competition between individuals coupled with limited population size would result in natural selection as the perfecting force of evolution (Mayr, 1982). Malthus' *Essay* is remembered in evolution for its influential hypothesis that long-term human population sizes are limited by food production. Malthus laid out his thesis to dispute contemporary arguments in favour of "systems of equality" (i.e., precursors to communism). He also made an evolutionary prediction (along with very many societal predictions): that there exists a limit to trait evolution by artificial selection in animals, plants, and humans (pg. 268-9). Malthus' writing style and intended audience was not a scientific one but his work is relevant as it presents an early example of predicting evolution that was read by both Darwin and Wallace.

Despite the influence of Malthus and his example of making predictions, both Darwin and Wallace limited the scope of their work to providing arguments in favour of the process of natural selection and a historical – not a predictive – theory of evolution. Darwin was a cautious scientist who perfected the arguments in favour of his theories over many years and emphasized natural selection as *not* goal directed

(Mayr, 1982; van Wyhe, 2007; Buchanan and Bradley, 2017; Partridge, 2018). Darwin had categorically rejected the Judeo-Christian creation story and, by extension, all theistic evolutionary explanations prior to his earliest sketch of *On the Origin of Species* (in 1842) because he saw it as lacking "predictive" and explanatory power (Barrett, 1987; Bowler, 1983; Brown, 1986). The publication and critical reception in 1844 of the theistic evolution book *Vestiges of the Natural History of Creation* further determined Darwin to amass a convincing body of naturalistic evidence before publishing his theory of evolution by natural selection (Bowler, 1983; Buchanan and Bradley, 2017; Partridge, 2018). Although he occasionally used theologically ambiguous language in his works (sometimes strategically, see Brown, 1986), Darwin spent significant effort in his scientific writing and personal correspondence debunking the theistic concepts of order, design, and directed variation. Darwin maintained a complex personal relationship with theism throughout his life (Brown, 1986), but scientifically he was a materialist who drew a lot of inspiration from the success of Lyell's application of uniformitarianism to geology (Gould, 1977). In contrast, Wallace was a declared (although heterodox and materialist) theist who eventually claimed that some "spiritual influx" had directed natural selection resulting in the evolution of life, animals, and humans (Wallace, 1889, pg. 476).

One reason why Darwin wrote more about predicting evolution than Wallace can be gathered from the role that each saw for artificial selection in the theory of evolution. Artificial selection and predicting evolution are closely linked and, as we shall see, led many scientists after Darwin to hypothesize about the direction, rate, and limit of trait evolution. Darwin drew many insights about the evolution of natural populations from studying artificial selection – unlike Wallace, who saw little parallelism between the study of domestic varieties and the study of natural species (Mayr, 1982). Therefore Darwin's interest in artificial selection makes it more likely for him, rather than Wallace, to have written about predicting evolution.

Darwin discussed the possibility of predicting evolution on several occasions in *Origin* and, in all but one, concluded that evolution cannot be predicted. While championing the importance of natural selection, Darwin wanted his audience to understand evolution as a complex process governed by a combination of deterministic and stochastic (i.e., *unpredictable*, from the perspective of someone working before the establishment of frequentist statistics and *in silico* simulations) factors. The word "predict" appears four times in the first edition of *Origin*[3]: once to state that biolo-

[3]It appears five times in total in the sixth edition. The extra occurrence (pg. 197) is in a new chapter defending natural selection, where Darwin discusses a cladistic prediction about plant stem growth.

gists are unable to predict which species can survive in foreign climates (pg. 104) and three times (pg. 126) to surmise that genera with more species are more likely to contribute to future biodiversity, "[b]ut which groups will ultimately prevail, no man can predict". This macro-evolutionary prediction was important enough that Darwin maintained it in all future editions, referred to it multiple times, and restated it near the conclusion of *Origin* – this time omitting the caution of unpredictability (pg. 489). Among all instances of the word "future" in the main text of *Origin* (27 in the original 1859 edition), only three putatively refer to predictions (in comparison 11 refer to Darwin's future chapters and works, eight refer to predictions about the future of science). The first edition of *Origin* has an ambiguously phrased hypothesis, similar to Sewall Wright's shifting balance hypothesis 70 years later, that could be read as Darwin making a prediction about the future (pg. 107). By the last edition that Darwin worked on in his lifetime, the sixth edition, he had removed the phrase "looking into the future" and changed the future tense to past tense (Darwin, 1876, pg. 84) to clarify that he was referring to the process of diversification, not forecasting future evolutionary outcomes. The next instance of "future" is indeed a prediction: the macro-evolutionary prediction explained above. His third putative prediction is not found in the first edition, appears by the fifth edition, but is modified in the sixth edition. Here Darwin distinguishes between rudimentary organs (which are vestigial and unimportant for future fitness) and "nascent" organs (which are new adaptations). In the fifth edition he predicts that nascent organs would become larger and more important for the physiology of the organism but stresses that "[i]t is difficult to know what organs are nascent; looking to the future, we cannot of course tell how any part will be developed, and whether it is now in nascent condition" (Darwin, 1871b, pg. 406); in the sixth edition all mention of the future is removed, "[i]t is, however, often difficult to distinguish between rudimentary and nascent organs; for we can judge only by analogy whether a part is capable of further development" (Darwin, 1876, pg. 398). Overall, this shows that Darwin realized the predictive implications of his theory of natural selection but was careful to conclude that the future cannot be known.

Other than his macro-evolutionary prediction, Darwin did propose a thought experiment on predicting evolution. He describes a possible experiment of placing a plant or animal in a new country among new competitors. He predicts that even if the climate is the same, the biotic environment would lead to different evolutionary changes than in its native home. A modern reader might hope that Darwin would herein propose an evolutionary experiment to study and predict what the different

outcomes could be. Alas, he does not. Instead, he makes the opposite conclusion: that we are ignorant of the mutual relationships that exist between organisms. “Probably in no single instance should we know what to do, so as to succeed”, i.e., “to increase its [the transplanted species’] average numbers in its new home” (Darwin, 1859, pg. 78). In summary, based on his purely qualitative theory of natural selection and the state of the biological sciences in his day, Darwin likely believed it would be too complicated to predict evolution. (A more controversial conclusion could be that, as a conflicted atheist/agnostic, he was ideologically opposed to the arrogance of the implication that man could predict god’s work, see pg. 78-9.)

Malthus had made evolutionary predictions about the limit of trait change by artificial selection. Darwin wrote extensively on artificial selection in *Origin* and expanded this study in *The Variation of Plants and Animals Under Domestication* (Darwin, 1868). I thought that Darwin would follow up on Malthus’ predictions but I could not find any clear examples in either of these publications where Darwin postulated on the direction or rate of trait evolution. Instead, Darwin emphasizes that artificial selection demonstrates that there is no limit to trait evolution. We may consider this to be a prediction but I think it is better read as Darwin highlighting the power of natural selection.

Finally, to understand why Darwin thought predicting evolution was too complicated and trait variation limitless, it is important to consider the competing evolutionary processes of his day. While evolution was broadly accepted, natural selection was widely disputed among scientists during Darwin’s life (and until the modern synthesis; Provine, 2001). Theistic, Lamarckian, and orthogenetic processes were the main contenders to natural selection at that time (Mayr, 1982; Bowler, 1983). Orthogenesis is the concept that internal forces compel species or higher taxonomic groups to follow a directional evolutionary path of birth, growth/development, and death/extinction (Gould, 1977; Popov, 2018). Orthogenesis shares its “perfecting principle” with Jean-Baptiste Lamarck’s original theory and some theistic evolution theories, but evolution under strict orthogenesis requires no external input from the environment or god (but see Ulett, 2014). I focus here on orthogenetic evolution because the challenges it posed to natural selection are similar to those posed by theistic evolution and pre-*Origin* Lamarckism. From our historical vantage point we distinguish several, distinct evolutionary hypotheses but it is important to keep in mind that there was substantial overlap and influence among theories. For example, post-*Origin* Lamarckism integrated natural selection and focused on environmentally acquired heritable traits (and, as explained in the next section, Darwin’s own theory

of heredity included environmentally acquired traits; Bowler, 1983). Darwin died before the major orthogenetic theories of evolution were espoused (Ulett, 2014; Popov, 2018) but he was aware of at least some early orthogenetic theories as evidenced by his rebuttal to Carl Wilhelm von Nägeli's work in the sixth edition of *Origin*. A 21st century reader can best understand the main argument of orthogenesis as relying on claims of strong pleiotropy caused by developmental (and environmental) constraints common to all organisms, such that the variation available to natural selection is constrained and therefore evolution follows a limited set of trajectories (Ulett, 2014). Darwin addressed these objections to his theory by arguing against internal factors as the cause of heritable variation (see next section) and by using examples of animal and plant breeds to show that the possible variation was virtually unlimited when artificially selected (Darwin, 1859; Darwin, 1868; Bowler, 1983; Popov, 2018). Darwin's steadfast persistence that evolution has no direction and cannot be predicted was expedient in the context of competing evolutionary theories.

In conclusion, Darwin perceived evolutionary theory as Janus-faced, affording both a historical explanation of the past and a glimpse into the future. This is highlighted both on the first pages of *Origin*, where he states that biotic interactions "determine [...] the future success and modification of every inhabitant of this world" (Darwin, 1859, pg. 6) presumably through natural selection, and on the last pages, where he restates his macro-evolutionary prediction. During Darwin's lifetime, theistic and orthogenetic evolution were major challenges to his theory of natural selection that would have shaped how any suggestion of predicting micro-evolution would have been understood by his audience. Darwin's hesitancy on the matter of predicting evolution was due to both his conviction that evolution is neither overseen by god nor limited by heritable variation, and the shortcomings of his purely qualitative theory of natural selection.

3.2 Early eugenics as the start of quantitative evolutionary predictions

The political movements that called themselves eugenic provide an ominous historical case study of applying evolutionary predictions. More importantly, I argue that eugenics was not a political perversion of overly simplistic genetic theories, as casually dismissed by some evolutionary biologists, but that eugenic applications were a main motivator for the advancement of population genetics. Therefore eugenics is central in my argument that predicting evolution has been one of the main motiva-

tions for the study of population genetics since the origin of the discipline.

The term eugenics was coined by Francis Galton to refer to "the science of improving [the human] stock" (Galton, 1883), either by increasing the reproductive output of some individuals who were deemed as valuable to the population (positive eugenics) or by preventing the reproduction of individuals deemed as worthless and sometimes even dangerous to the population (negative eugenics; Levine and Bashford, 2010). I focus specifically on the self-identified eugenic political movements, from the mid-19th until the mid-20th century, whose main goal was state-sanctioned interference into human reproductive practices through the application of ostensibly scientific theories of heredity (Paul and Moore, 2010; Turda, 2010). Many evolutionary biologists and geneticists supported eugenic political movements in their countries until the 1930's, either directly as part of eugenic organizations or indirectly by providing forums for discussion in scientific journals and conferences (Kevles, 1985; Roll-Hansen, 2010). Although eugenics is most closely associated with the forced sterilization of racialized, Black, and Indigenous people in the USA and the genocides of Jewish, Romani, disabled people, and people with psychosis perpetuated by the Nazis during World War II, eugenic policies were enacted around the world during the 20th century (see Bashford and Levine, 2010). The political and scientific ideas of the eugenics movements were not abandoned *en masse* after World War II (Allen, 1955; Mazumdar, 2002). The use of genetic technologies to facilitate or prevent the inheritance of particular genes, haplotypes, or chromosomal ploidies continues today as "reprogenetics" (a portmanteau of reproduction and genetics) through medical genetics, genetic counselling, assisted reproductive technologies, etc. It can be argued that the eugenics movements disappeared due to the global rise of neoliberal politics, which replaced collective social support by the state for marginalized groups (e.g., welfare) with an internalized individual preoccupation for oneself and one's children to contribute in the state economy (Valentine, 2021). Reproductive control of human populations as an individual or cultural reproductive practice/technology has multiple independent, ancient origins and continues today (Kevles, 1985; Bashford, 2010). However, as my goal is to tell the history of predicting evolution with little distraction from current bio-ethical debates about reprogenetics/"new eugenics"/etc., I focus on eugenics specifically as a historical state project (i.e., to be enacted by state institutions for the net benefit of the normative population, in other words, those who were judged as its most valuable citizens) that ended in the mid-20th century.

The future-facing side of evolutionary theory was proposed to be applied to hu-

mans less than a decade after the publication of *Origin*. Galton combined Darwin's theories with contemporary Victorian ideas about the innate nature of intelligence and morality in order to argue that human populations could be improved by selective breeding (Galton, 1865; Cowan, 1972; Paul and Moore, 2010). Galton was more mathematically-minded than Darwin and had been introduced to statistical forecasting methods from meteorology (Cowan, 1972; Fancher, 2009). Following Darwin's recommendation, from *Origin* and repeated in *Variation*, for further studies on the hereditary basis of traits, Galton began investigating the hereditary basis of human intelligence from a quantitative perspective using the pedigrees of several thousand eminent men. The resultant book, *Hereditary Genius: An Inquiry Into Its Laws and Consequences* (Galton, 1869), was admired by both Wallace and Darwin and was a great influence for *The Descent of Man*, although it was not widely popular (Paul and Moore, 2010; Stigler, 2010). As well as repeating similar concerns as Malthus, that the "weak", "incompetent" classes are out-breeding the "prudent", "vigorous" classes, Galton wrote about gifted men making a smaller contribution to the breeding pool by marrying late and he ranked the human races based on their hereditary propensity for talent. Not yet having coined the term (positive) eugenics, he stated that society should incentivize gifted people to breed and so steer the future evolution of humans. Finally, Galton appealed for mathematicians to use Darwin's hypothesis of inheritance (pangenesis, see below) in order to derive a formula that predicts the average distribution of traits for the offspring of a given breeding pair. In summary, within a decade after the publication of *Origin* and with almost no quantitative developments in evolutionary theory, Galton was already proposing the application of artificial selection to human populations and heralding quantitative evolutionary predictions.

One important difference that distinguished Darwin's view on evolution (and its unpredictability) from Galton's view was the role of the environment on heritable traits. Darwin's hypothesis for the mechanism of heredity, which he considered "a provisional hypothesis of pangenesis," was among his most awaited but least enduring contributions to biology (Darwin, 1868; Provine, 2001). *Variation* served to supply the missing piece in his theory of natural selection: the source of heritable variation. Darwin reconciled blending inheritance with 'reversion' (i.e., when traits skip a generation) by positing that all organs threw off minuscule, seed-like particles called gemmules that could sit dormant for multiple generations. More importantly, any environmentally-induced or habit-derived phenotypes could potentially be inherited when the gemmules from those organs aggregated in the reproductive organs

to form the gametes. Darwin's provisional hypothesis of pangenesis explained how large amounts of heritable, random variation could be generated for natural selection to act upon without the influence of orthogenetic internal factors. However, Darwin solved the problem of genetic variation by attributing it principally to changed environmental conditions and habits. Under this combined type of 'soft' and 'hard' inheritance, species traits could be evolved by subjecting the population to conducive environmental conditions and/or enforcing the intended habit(s) so that the desired trait(s) would eventually become hard-sense heritable. With consultation from Darwin, Galton tested Darwin's pangenesis hypothesis by transfusing blood between rabbits with different coat colours then inbreeding the treated rabbits (Galton, 1871; Mayr, 1982; Fancher, 2009; Paul and Moore, 2010). When the offspring showed the unaltered coat colours of their parents, Galton abandoned soft inheritance and turned his attention to studying 'hard' inheritance (meanwhile Darwin maintained his conviction for pangenesis, describing Galton's experiments as inconclusive in the second edition of *Variation*). Strict, hard inheritance was a fringe view among biologists until the 1890's (Mayr, 1982).

Under Darwin's pangenetic inheritance, eugenics would mean giving people good environmental conditions so that their heritable traits could gradually be improved (Paul and Moore, 2010). This is why some modern advances that are currently regarded as socialist, like public health and environmental medicine, in fact have their roots in eugenic ideas. Moreover, it is important to understand that eugenics was combined with ideologies across the political spectrum, from Marxist radicalism to liberal democracy to *laissez faire* conservatism (Kevles, 1985). From our 21st century vantage point, we tend to associate eugenics with supremacist ideologies that enforced sterilization and extermination of specific human groups. But historical personages do not share our privileged vantage point and so did not understand their biases as supremacist. Darwin's belief that good habits become hereditary, and his shared preoccupation with Galton that the most able men should reproduce, led Darwin to favour moral education and development of reasoning as a eugenic strategy (Darwin, 1871a). On the other hand, Galton's belief in hard inheritance led his eugenics towards class-based segregation and forcible control of reproduction (Galton, 1909). Karl Pearson, Galton's successor in 1907 as the director of the Galton Eugenics Laboratory at University College London (which became the first Galton Professor of Eugenics several years later), believed that between-group competition through colonization and war led to evolutionary advancement (Semmel, 1958; Magnello, 1999). Pearson's nationalist, imperialist politics combined with his (now

scientifically well accepted) belief in hard heredity and led eugenics towards the antisemitism and racism now most associated with the eugenics movements (Norton, 1983; Delzell and Poliak, 2013).

Turning our focus away from eugenics and back to the history of predicting the evolution: The idea of predicting evolution emerged in its current form as soon as quantitative methods were brought to bear on natural populations. Galton developed regression because he wanted to predict the traits of offspring based on the traits of their parents, and so to know how to select the most eugenically favourable human breedings (Stigler, 2010; Paul and Moore, 2010). Walter Frank Raphael Weldon turned Galton's statistical methods towards the study of crab morphology, looking at the variation both within populations and between species (Provine, 2001; Davis, 2022). Weldon and his wife, Florence Joy Weldon (née Tebb), measured continuous traits with the aim of understanding the variation of and correlation among traits. His 1893 paper ends (Weldon, 1893):

> It cannot be too strongly urged that the problem of animal evolution is essentially a statistical problem: that before we can properly estimate the changes at present going on in a race or species we must know accurately (a) the percentage of animals which exhibit a given amount of abnormality with regard to a particular character; (b) the degree of abnormality of other organs which accompanies a given abnormality of one; (c) the difference between the death rate per cent, in animals of different degrees of abnormality with respect to any organ; (d) the abnormality of offspring in terms of the abnormality of parents, and vice versa. These are all questions of arithmetic; and when we know the numerical answers to these questions for a number of species we shall know the direction and the rate of change in these species at the present day —- a knowledge which is the only legitimate basis for speculations as to their past history and future fate.

Pearson, who had just been appointed professor in geometry at University College London and was not yet acquainted with Galton, was brought onto the crab biometry project around 1892 to develop further statistical methods for analyzing the wealth of population data produced by Raphael and Florence Weldon (Davis, 2022). The crab population data of the Biometric Laboratory led Pearson to improve on Galton's method of regression and develop the concepts of standard deviation, covariance, and the coefficient of variation. Weldon himself was not a eugenicist and he warned Pearson to consider environmental contributions to traits (Kevles, 1985). However

Weldon had died by the time that Pearson was put in charge of the Galton Eugenics Laboratory at the behest of Galton (Magnello, 1999). The contributions of the gradualist biometric school of Galton, Weldon, and Pearson to the field of quantitative genetics and their later disputes with the saltationist Mendelian school of William Bateson and others has been well documented (Provine, 2001; Davis, 2022). It is important to note that while the biometricians disagreed with the Mendelians about how evolution proceeded, both camps tended to agree about the importance of eugenics (Kevles, 1985; Spencer and Paul, 1998). Although the early eugenicists, Galton and Pearson, lacked a Mendelian model for genes, they pursued statistical genetic methods with the goal of predicting and directing human evolution.

3.3 Studies of environmental and genetic interactions predate the modern synthesis

Prior to the advent of population genetics, early evolutionary biologists observed instances of G$\times$E but they could not study these interaction terms because they lacked theories to distinguish between genotype versus phenotype and to compartmentalize the phenotypic variation into the binary categories of genotype versus environment. As alluded to above, Darwin documented many G$\times$E interactions in both wild and domestic species. However, he could not distinguish between the role of selection versus environment: "in many cases it is most difficult to distinguish between the definite result of changed conditions, and the accumulation through natural selection of indefinite variations which have prove[n] serviceable" (Darwin, 1868, Vol II, pg. 281). Other naturalists were similarly stymied. For example, before working on breeding experiments, William Bateson searched for evidence of parallel adaptive selection among cockles and two other mollusc species in the Aral sea and Egypt (Bateson, 1889). He observed little within-population variation and concluded that the contribution of the environment on the quantified traits was so great that it could not be shown that natural selection was operating. Lacking a gene concept made it difficult for pre-Mendelian evolutionary biologists to distinguish between genetic and environmental contributions, and their interactions.

The history of G$\times$G interactions in evolution could not begin outright until the rediscovery of Mendel's laws of inheritance in 1900 and the development of the model that genes are discrete, potentially interacting, units. The first evidence of gene interactions were discovered shortly after Gregor Johann Mendel's work gained popular attention. Upon reading one of Mendel's papers in 1900, Bateson became a cham-

pion of Mendel's laws of inheritance (Provine, 2001) and, as a result, his lab was first to document examples of G×G (Miko, 2008). In order to build undeniable evidence supporting Mendel's laws, Bateson expanded existing breeding experiments in his lab (Provine, 2001). Contrary to the parsimony principle, when members of Bateson's group encountered apparent exceptions to Mendel's laws, they proposed increasingly more complex hypotheses that could be consistent with Mendel's theory (Mayr, 1982; Richmond, 2001). It was in this way that they were able to document at least four unique examples of epistasis by 1908: a two-locus interaction impacting poultry comb shape was found by Reginald Crundall Punnett, a four-locus interaction impacting flower colour and leaf hairiness in *Matthiola* was found by Edith Rebecca Saunders (Bateson, Saunders, and Punnett, 1904; Bateson, Saunders, and Punnett, 1906), a four-locus interaction impacting mouse coat colour was found by Florence Margaret Durham (Bateson, Saunders, and Punnett, 1908), and a four-locus interaction impacting snapdragon flower colour was found by Muriel Wheldale Onslow (Wheldale, 1907). Bateson coined the term "epistasis" to refer to Wheldale's findings (Richmond, 2001). The discovery of sex chromosomes was another important example of the prevalence of gene interactions. With the popularization of Mendel's laws of inheritance after 1900, experimental evidence was beginning to accumulate on the importance of G×G.

While early geneticists working in controlled laboratory conditions carefully cataloged instances of G×G, the first unambiguous observations of G×E were being documented. In 1909 Richard Woltereck posited the environmental reaction norm (Mayr, 1982). In a surprisingly modern experimental design, Woltereck measured helmet height in clonal populations of *Daphnia* taken from different lakes (Keilhack, 1909). He found that helmet height varied as a function of nourishment quantity and that this norm of reaction differed between genotypes depending on their provenance. Woltereck even conducted experimental evolution on the helmet height reaction norm: it could be shifted after about two years of selection. As Woltereck examined other environmental variables and traits, he discovered other reaction norms, few of which were predictable between genotypes (Falk, 2000). The concept of a genotype-dependent environmental reaction norm was exceedingly complex for biologists of his day, who were still skeptical about fundamental genetic principles (Mayr, 1982; Sturtevant, 2001). Most evolutionary biologists eventually shifted their focus away from the environmental contributions on phenotypes and towards the role of genes. Although the first publication on G×E was contemporaneous with the discovery of the first novel mutation by Thomas Hunt Morgan, G×E was outcast from

dominant evolutionary theory until well after the establishment of the modern synthesis.

The establishment of Mendelian genetics as the dominant paradigm of inheritance required biologists to, for almost a century, ignore or stifle observations and ideas about the role of the environment on phenotype and evolution. For example, the mounting orthodoxy that environment plays no role on heritable traits led the early population geneticist Raymond Pearl to argue that biological evidence was being purposefully ignored in favour of genetic theory (1917). Seminal books on evolutionary theory, from John Burdon Sanderson Haldane's in 1932 to Ernst Walter Mayr's in 1982, argued vehemently against the possibility that heritable traits could be environmentally-induced. The role of the environment on mutation and heritable variation is still debated today (e.g., Fitzgerald, Hastings, and Rosenberg, 2017; Ashe, Colot, and Oldroyd, 2021). For early geneticists, like Morgan, any neo-Lamarckian exceptions to Mendelian inheritance had to be ignored in order to promote the importance of genetics. As genetics began to establish itself between 1910-1920, it eclipsed natural selection: many leading evolutionary biologists of the time championed a predominately mutation-driven view of evolution (Provine, 2001; Mayr, 1982). The development of the field of population genetics finally reconciled genetic inheritance with natural selection. But, in order to keep the mathematical framework manageable, the effects of genetic and environmental interactions were sidelined.

3.4 Eugenics and the modern synthesis

Darwin's ideas, and biometry, were eventually united with Mendelian genetics in the 'modern synthesis' (Fisher, 1930; Wright, 1931; Haldane, 1932; Huxley, 1942). Working without computer simulations, Haldane, Ronald Aylmer Fisher, Sewall Green Wright, and others had to make simplifying assumptions to arrive at a tenable mathematics of population genetics. Fisher and Wright, in particular, were aware that G$\times$G, G$\times$E, and higher order interactions could exist but chose to assume in their math that the contribution of these interaction terms was negligible as compared to the main effects of evolutionary forces acting on additive alleles in a single, focal environment (Crow and Kimura, 1970; Mayr, 1982). Much attention has been given to Fisher and Wright's conflicting views on evolution (e.g., Provine, 1986; Provine, 2001; Crow, 2010), but there are also many similarities. At the time when they published the first works on population genetics, Fisher and Wright were both primarily concerned with artificial selection and management of agricultural populations but

less familiar with natural populations (Mayr, 1982; Sarkar, 2017). Moreover, both Fisher and Wright were concerned with predicting evolution.

Fisher often considered selection to be the predominant evolutionary force acting on populations (Crow and Kimura, 1970; Mayr, 1982; Provine, 1986). Although he understood the important roles that mutation and drift could play (and, contrary to some claims (e.g., Sarkar, 2017), Fisher *was* aware of contemporary genetic findings like epistasis), Fisher's math focused on large, intermixing populations operating under a weak mutation regime (1930; but see pg. 102-4). Picking up from the observation of differential reproductive values among individuals in human populations, where Pearson and Galton left off (Pearson, 1897; Galton, 1901), Fisher's fundamental theorem showed that the partial rate of change in fitness is equal to the additive genetic variance in fitness of the population. Although so general that it is sometimes considered trivial (Price, 1972) or too abstract to be biologically relevant (Crow and Kimura, 1970), Fisher's fundamental theorem conclusively bound Mendelian genetics together with Darwinian natural selection and eliminated the need for biased mutations, Lamarckian environmentally induced traits, or orthogenetic effects to drive rapid evolution (Grafen, 2003; Plutynski, 2006). Moreover, it provides a direction for evolution: towards increasing population fitness (Fisher, 1930, pg. 36; Okasha, 2008). Unlike the population geneticists contemporary with Fisher and who extended his work after him (Kimura, 1958; Crow and Kimura, 1970; Hamilton, 1964), Fisher considered his fundamental theorem of central importance to evolution and made little attempt to relax the strict assumptions that he made (e.g., large population size, random mating, fixed fitness values, additive genetic variance, etc.) in later publications of his theorem (Crow and Kimura, 1970, pg. 210). Fisher's work was a continuation of the predictive aspects of the biometric research program: the final sentence of his chapter on the fundamental theorem (pg. 47) recalls Weldon's ideas of predicting evolution, "direct observational methods may yet determine the numerical values which condition the survival and progress of particular species." This research program of predicting evolution is continued, for example, by current work on estimating the distribution of fitness effects of mutations, especially for the beneficial tail of the distribution (e.g., Eyre-Walker and Keightley, 2007; Gordo, Perfeito, and Sousa, 2011; Frenkel, Good, and Desai, 2014). More importantly, after confusing population geneticists half a century ago, Fisher's fundamental theorem is now recognized as truly fundamental for predicting evolution (Shaw, 2019; Burford Reiskind et al., 2021). Although Fisher did not mention predicting evolution as directly as the biometricians before him, his dogged focus on the most predictable

evolutionary regime (strong selection weak mutation), his fundamental theorem of evolution, and his work on uncovering the genetic-basis of human disease are all best understood in the context of his eugenic politics.

Fisher himself clearly stated that his eugenic ideas cannot be separated from his population genetics work (1930, pg. *x*), and historians have tended to agree (Norton, 1983; Kevles, 1985; Gould, 1991; Moore, 2007). Fisher was actively involved with the Eugenics Society starting from 1909, founding the Cambridge chapter during his undergraduate studies, and finally left in 1941. The Eugenics Society provided Fisher with funding to work on and publish his seminal 1918 paper. The last third of *The Genetical Theory of Natural Selection* is dedicated to eugenic considerations. Representing the Eugenics Society in 1934 as part of a committee tasked by the British Parliament, Fisher recommended the legalization of voluntary sterilization for people impacted by or likely to transmit hereditary mental disorders and grave disabilities. Fisher remained steadfast in his politics, but stopped proselytizing, in the lead up to World War II (MacKenzie, 1976; Bodmer et al., 2021).

The recognition that Fisher was a eugenicist is not meant to diminish the enormity of Fisher's contributions to population genetics (or other fields of science; Bodmer et al., 2021); indeed it is precisely because his contributions are so foundational and durable to our field that an evaluation is warranted about how his eugenic and otherwise hegemonic beliefs may have featured in his work. Fisher's beliefs, like those of any scientist (Rouse, 2004; Branch et al., 2022), likely impacted his assumptions and the findings he considered important. For example, Fisher derived the inverse relationship between selection strength and population size, which is a key insight for neutral and nearly neutral theory, but he discounted its importance because he considered most populations to have effective sizes on the order of at least 10^6 (1930, pg. 10, 118). It is not merely that Fisher considered most populations as large, it reads to me that he discounted small populations from evolutionary importance as a value judgement; the small populations are inconsequential because they are not the species that dominate the earth (1930, pg. 96-99, 174-175). There seems to be a similar sentiment in his theory of dominance modifiers, that the wild-type must literally assert its dominance over the recessive mutants (1930, pg. 53, 65-9). Another oddity that has previously been remarked upon (Price, 1972) is that Fisher's fundamental theorem presupposes a constantly deteriorating environment, ostensibly due to between-species interactions (1930, pg. 41-2) as later independently summarized by the Red Queen hypothesis (van Valen, 1973; Bell, 1982). A preoccupation with the deterioration of populations and their environments is one of

the central themes of eugenics in general and Fisher's eugenics in particular (1930, pg. 222-227, 263; Levine and Bashford, 2010). Therefore, although a constantly deteriorating environment may seem a wrong assumption for some, it was an acceptable one in Fisher's worldview. Ultimately science is a social endeavor (Shapin and Schaffer, 2011), therefore unsupported ideas originating from the specific worldview of a single scientist are usually weeded out. This is why few learn about Fisher's theory on the evolution of dominance but many learn that Fisher was the first to consider the evolution of modifiers (Otto and Day, 2007).

Perhaps the most lasting contribution of Fisher (i.e., to population genetics specifically) is the gene's-eye view. This perspective is implicit in Fisher's fundamental theorem, even though he never addressed it directly and his successors who articulated it more clearly only later attributed it to Fisher (Richard Dawkins' *The Selfish Gene* is a prominent example; Edwards, 2014). The gene's-eye view refers to the modeling assumption that an evolving population should "best be thought of as a population of fundamental replicating units – of genes – rather than as a population of individual animals or of cells" (P. B. Medawar 1981 as cited by Ågren, 2021). Fisher employed the gene's-eye view in his fundamental theorem by focusing on the change in additive genetic variance caused by replacing one allele with its alternative allele (i.e., for a biallelic locus), therefore externalizing as part of the environment all contributions from other loci, epistatic genetic interactions, additional alleles on the same locus, varying selection on the two focal alleles, and other types of biotic factors that are usually *not* defined as part of the environment (Okasha, 2008; Ågren, 2021). Although most prominent in Fisher's fundamental theorem, the gene's-eye view appears in other parts of Fisher's work; for example, he articulates it quite clearly in a discussion on the effectiveness of selection: "each gene is constantly tending to create genetic situations favourable to its own survival" (Fisher, 1930, pg. 95). Most evolutionary biologists, including Fisher himself (e.g., pg. 54-6), understand the gene's-eye view to be a convenient simplifying model for organismal evolution. Nevertheless the gene's-eye view is a routine assumption of population genetics that has led to significant controversy over what exactly is the unit of selection that evolution acts upon (e.g., Lewontin, 1970; Wright, 1980; Gould and Lloyd, 1999).

Wright approached the question of predicting evolution from a different perspective than Fisher. Wright considered $G \times G \times E$ important ("The evolution of complex organisms rests on the attainment of gene combinations which determine a varied repertoire of adaptive cell responses in relation to external conditions." 1931,

pg. 147), but he did not incorporate this complexity into his equations (Mayr, 1982; Provine, 1986). In his 1931 paper, Wright derives the recursion equation for the change in allele frequency between successive generations for a 2 locus trait but does not develop this any further, merely concluding that additivity is more straightforward. Under the heading "control of evolution," Wright talks about how an adaptive process could be accelerated by tweaking selection, mutation, population size, and population structure. Implicitly assuming the prevalence of pleiotropy, he points out that strong selection on one trait at the expense of other characters will lead to quickly exhausting all variation available in the population. In opposition to Fisher's large populations with strong selection, Wright posits the shifting balance hypothesis (creating smaller demes and applying selection to those demes with occasional migration between them and, perhaps, increased mutation rate) as the way to best manage populations by optimizing the rate of evolution.

Wright was never actively involved in eugenic political movements, although he is listed as sitting on the Advisory Council of the American Eugenics Society from 1927-35 and renewing his membership to the society once after World War II (Provine, 1986; The American Eugenics Society, 2015). Already a prominent evolutionary biologist, his title and name undoubtedly added to the scientific legitimacy of the political movement, as did many other scientists' titles and names during those years. Provine describes Wright as sympathetic to eugenic ideas but not much impressed with the resultant work, and so opposed to actively advocating in favour of eugenics (1986, pg. 110). I would say that a similar general attitude towards eugenics persists among working evolutionary biologists and geneticists today.

The biometricians were quick to understand that a quantitative knowledge of the relationship between traits, their heritability, and their fitness effects would allow the prediction of evolution. Mendelian genetics provided a seemingly simple mechanism of heredity that could lead to actionable eugenic efforts for the "improvement" of human populations. However, it became increasingly obvious with further progress in genetics that the relationship between genotype, phenotype, and fitness is highly complex. This is why most evolutionary biologists who held eugenic convictions eventually divested from eugenic politics in the 1930's (Kevles, 1985), before the genocides perpetuated in World War II and before the public disclosure of the forced sterilizations in the USA. Even Fisher, who held strong eugenic convictions, lamented in 1930: "it is scarcely possible to imagine a problem more intricate, or requiring so inconceivably detailed a knowledge of the bionomic situation, as that of tracing the net gain in fitness of any particular genetic change" (pg. 53-54). Like

many other eugenic biologists, Fisher directed his future biological research efforts towards deepening our functional and genetic understanding of human traits.

This very brief history of the relationship between predicting evolution and eugenics has focused superficially on the work of three key figures, Galton, Pearson, and Fisher. This is due primarily to my own space and time limitations. A deeper treatment of this subject area would provide better context and a more comprehensive understanding. As scientists in general, but I think particularly in theoretical works where the perspective of one individual can radically shift our scientific paradigm, we often focus on a few key figures and forget that science is a social endeavour that we, as certified experts, construct together (Kuhn, 1996; Latour and Woolgar, 1987). Although the history of predicting evolution is entangled with the history of eugenics, the present of predicting evolution owes much to scientists who did not advocate for eugenics, like Wright and others, and to vocal eugenics opponents, like Stephen Jay Gould (1996; 1991).

I am told by my lab-mates who reviewed the draft of this Chapter that the history of eugenics ends abruptly and that this can be very disquieting to the reader. I have not managed to think of an improved structure for this thesis that would center my scientific contributions to the field of predicting evolution without side-lining my ethical concerns about evolutionary predictive technologies. If you require some attempt at a resolution to these issues before you can engage with the rest of this work, section 6 of the Discussion Chapter (pg. 148) picks up the historical and ethical questions of eugenics / predicting evolution where we leave off here.

4 The current state of predicting evolution

Prediction of evolution is an ambitious undertaking that would consolidate knowledge from all fields of biology for the benefit of human health and global biodiversity. There are many current and potential examples of predicting evolution and applying those evolutionary predictions to manage or solve problems of human and ecosystem health (reviewed in Lässig, Mustonen, and Walczak, 2017; Burford Reiskind et al., 2021; Thompson et al., 2021; Wortel et al., 2022). Evolutionary forecasting methods are used to predict the next year's dominant strain of seasonal influenza in order to maximize the efficiency of annual vaccines (Agor and Özaltın, 2018); molecular markers of cancer progression are used to predict the likely time-frames and outcomes for human cancer patients (Nasr et al., 2019); functional models of microbial community composition could be used to treat human diseases caused

by dysbiosis (Barraclough, 2019); predicting the evolution of antimicrobial and pesticide resistance would allow targeted deployment and extend the viable lifetimes of these anthropologically useful compounds (Liu et al., 2017; Furusawa, Horinouchi, and Maeda, 2018); incorporating the effects of life-history trait evolution would improve eco-evolutionary predictions of species' responses to climate change (Shefferson, Mizuta, and Hutchings, 2017); incorporating evolutionary predictions could improve the outcomes of management and conservation efforts for endangered species (Shefferson et al., 2018). For these and many other applications, predicting evolution would be an advantageous technology.

My thesis investigates evolutionary predictions about two urgent issues that pertain to very different biological systems: predicting the outcome of a species hybrid zone and determining the fitness landscape(s) that underpins antibiotic resistance evolution.

The process that generates species-level biodiversity ("speciation") through the evolution of reproductive isolation in sexually reproducing organisms is well described (Coyne and Orr, 2004). One mechanism for the reproductive isolation of populations acts post-zygotically when hybrid individuals have a lower fitness than individuals from the parental populations. The fixation of these incompatible alleles between speciating populations was described independently by Bateson, Theodosius Grigorievich Dobzhansky, and Hermann Joseph Muller and, so, these types of epistatic interactions are now referred to as Bateson-Dobzhansky-Muller incompatibilities (BDMIs; Bateson, 1909; Dobzhansky, 1936; Muller, 1942). Therefore, G$\times$G is a key player in speciation.

One application of predicting evolution could be to predict the future outcome of incipient speciation and hybridization. Hybrid zones can either lead to further reproductive isolation of the parental populations that produced the hybrid zone or to the long-term maintenance of genetic diversity. Under selective conditions that promote speciation, like the frequent production of hybrid individuals with fitness costs, post-zygotic BDMIs can be reinforced by pre-zygotic mechanisms, like mate choice, leading to the reinforcement of speciation (Servedio and Noor, 2003). On the other hand, when hybrid individuals have a fitness benefit, hybrid zones can be maintained for long periods of time and, rarely, can even lead to hybrid speciation (reviewed in Ottenburghs, 2018; Taylor and Larson, 2019). Predictions about the likely evolutionary outcomes of incipient hybrid zones can guide conservation efforts, for example to increase species-level biodiversity by preserving hybrid populations that are more likely to result in speciation. Conversely, for sexually reproducing species

deemed parasites/pests, speciation predictions can be used, for example, to prevent the increase of genetic-level biodiversity (and give us an edge in evolutionary arms races) by targeted destruction of hybrid zones that are more likely to result in ongoing genetic introgression. In light of species range shifts and anthropogenic loss of species-level biodiversity, which will both accelerate under climate change, predicting the outcome of hybrid zones could benefit our ability to steward human health and global biodiversity.

Another application of predicting evolution is the prediction of antibiotic resistance (ABR) evolution (Hall, 2004; Palmer and Kishony, 2013; Santos-Lopez et al., 2021). Here, ostensibly asexual bacterial populations evolve a new trait. Therefore, the field of ABR evolution is related to the evolution of novelty. Although many of the changes in allele frequencies of ABR mutations across different bacterial populations are probably due to horizontal gene transfer (Ochman and Moran, 2001; Sommer et al., 2017), many studies of ABR evolution are gene-centric and focus on the role of *de novo* mutations (Knöppel, Näsvall, and Andersson, 2017; but see Baquero, Tedim, and Coque, 2013; Gil-Gil et al., 2021). Nevertheless, the change in frequency of ABR alleles due to both segregating variation and novel mutations is impacted by epistatic and environmental interactions. The efficacy of an ABR mutation can depend on its genetic background (e.g., Rodríguez-Verdugo, Gaut, and Tenaillon, 2013), or not (e.g., Knopp and Andersson, 2018). Many ABR mutations exhibit a cost of resistance (Melnyk, Wong, and Kassen, 2015), meaning that they have a fitness benefit as compared to the wild-type (WT) in the presence of sufficiently large concentrations of antibiotic but they have a fitness cost in the absence of antibiotic. These so-called costly ABR mutations by definition exhibit G$\times$E. ABR mutations can also exhibit G$\times$E with non-antibiotic environments. For example, ABR can evolve in the absence of antibiotic treatment, either in response to non-antibiotic environmental stressors or to benign lab environments (Tenaillon et al., 2012; Knöppel, Näsvall, and Andersson, 2017). Moreover, costly ABR mutations may evolve compensatory mutations such that this novel combination of mutants together exhibits a similar fitness as the WT in the absence of antibiotics (Durão, Balbontín, and Gordo, 2018). This compensatory evolution is an example of G$\times$G leading to the loss of G$\times$E, thus implying that G$\times$G$\times$E interactions are afoot.

ABR has existed for millions of years in natural populations probably because microbes secrete low concentrations of antibiotic compounds to regulate the population sizes of their neighbours (D'Costa et al., 2011; Andersson and Hughes, 2014; Paun et al., 2021). The discovery of antibiotic compounds paved the way for mod-

ern medicine, however, it was accompanied by the concurrent discovery of ABR (Aminov, 2010). The widespread use of antibiotics has led to a rise in ABR (Goldstone and Smith, 2017). Today, ABR exerts a large toll on human health and productivity (Murray et al., 2022). Some warn that we are headed towards a post-antibiotics world, where antibiotics are largely ineffective due to the ubiquitous presence of ABR (Gross, 2019). Therefore, technologies for the prediction and control of ABR evolution are needed urgently.

4.1 The main challenges for evolutionary predictions

I see three main challenges that stand in the way of predicting evolution: *i)* organisms are very sensitive to their environmental and genotypic conditions, *ii)* future environments are difficult to predict, and *iii)* a deep knowledge of the functional/physiological effects of current (as well as potential!) genotypes and the impact of those genotypes on organismal fitness effects is needed. Over long evolutionary time-scales, the details of specific evolutionary outcomes are impossible to predict (but see Lamsdell et al., 2017; Barnes, Sclafani, and Zaffos, 2021). For short time-scales, however, we may have a shot.

I list 'organisms are very sensitive to their environmental and genotypic conditions' as the first challenge to predicting evolution. This is intended as a less technical way of summarizing that G$\times$G, G$\times$E, G$\times$G$\times$E, and higher-order interactions pose a challenge to predicting evolution (Sailer and Harms, 2017). As overly simplified above, early evolutionary theory laid down the math of single genes evolving in single, constant environments and single, panmictic populations. Contemporary evolutionary theory has progressed beyond those assumptions, often using changes in total effective population size, or different rates of birth and death, as a way to understand the evolutionary process under fluctuating environments, changing genotypic backgrounds, and other more complex circumstances (Ewens, 1967; Karlin and Levikson, 1974; Tachida and Iizuka, 1991; Otto and Whitlock, 1997; Wahl and Zhu, 2015). These approaches are important but may be limited (Kern and Hahn, 2018; Rocha, 2018). Our organismal knowledge, however, continues to be limited by assumptions of weak environmental and genetic interactions. The bulk of our physiological and genetic knowledge comes from work with laboratory model organisms, where a limited set of (usually inbred) genotypes are grown and assayed under controlled, low-stress environmental conditions. Systematically quantifying the phenotypes and physiological responses of organisms across a vast array of possible

genotypes, environmental conditions, and the interaction of different genotypes with different environmental conditions is a sisyphean undertaking that has recently begun in earnest (e.g., Nichols et al., 2011; Marjoram, Zubair, and Nuzhdin, 2014; Scott et al., 2017; Brochado et al., 2018; but see more on this below). Evolution research also involves the study of natural populations, which are composed of genetically diverse organisms encountering sundry environmental conditions. Nevertheless, many of the tools we have for identifying past examples of adaptive evolution in the genome actually identify selective sweeps, with the strongest signals occurring under additive genetic effects and constant environments (Vatsiou, Bazin, and Gaggiotti, 2016; Pavlidis and Alachiotis, 2017). These methods often lack the sensitivity to detect adaptive evolution driven by multiple loci (but see Li, Schumer, and Bank, 2022) or restricted to a subpopulation living in a specific environment. While it is clear when we look back on evolutionary outcomes that gene combinations are important for adaptation and the evolution of novelty (Lunzer, Golding, and Dean, 2010; Starr et al., 2018; Park, Metzger, and Thornton, 2022), the role of epistasis in adaptive dyanmics is still debated when looking forward, towards evolutionary predictions (Whitlock et al., 1995; Bank, 2022). Similarly, signatures of selection are more difficult to identify under fluctuating environmental conditions (Gillespie, 1972; Mustonen and Lässig, 2009; Cvijović et al., 2015). Evolutionary biologists are working to construct modeling frameworks that will reconcile classical population genetics, which usually assume that organisms are in a single, constant environment, with reaction norms that describe trait responses to an environmental gradient (e.g., Ogbunugafor, 2022).

If we cannot predict the environment, then we cannot predict where evolution is headed. This is because the direction of evolutionary change is always determined by the environments organisms experience. Both biotic and abiotic environmental categories can change due to stochastic events (e.g., biotic: immigration of a new species; abiotic: asteroid impact or volcanic eruption) and by feed-backs within the same environmental category (e.g., biotic: changes in the primary producer composition impacts higher trophic levels; abiotic: a massive increase in atmospheric oxygen impacted the concentrations of other atmosopheric gases; Avice et al., 2018) or by feed-backs with the opposing environmental category (e.g., biotic: acidification of a lake due to acid rain leads to changes in the microbial community; abiotic: the evolutionary novelty of photosynthesis by cyanobacteria led to a massive increase in atmospheric oxygen levels). Better knowledge of Earth sciences coupled with a better understanding of functional ecology and real-time monitoring of relevant envi-

ronmental variables will lead to improved environmental predictions. Although some examples on the geological time-scale are given above, we can have better predictions for events at shorter time-scales. The challenge to predicting evolution posed by environmental predictions is outside the scope of my thesis and will not be discussed further.

The final scientific challenge I listed comes from a lack of deep functional knowledge (Thompson et al., 2021; Burford Reiskind et al., 2021). As explained in section 3 above, this is a major problem to predicting evolution that was already identified by Darwin, Galton, Pearson, and Fisher. We need to understand not just the current mapping of existing genotypes to their physiological and fitness effects, but also the physiological and functional effects of new mutational combinations. This can be achieved by mutational experiments where all the amino-acids or nucleotides of a protein, for example, are systematically changed (Fowler et al., 2010; Hietpas et al., 2012; Matuszewski et al., 2016; Karageorgi et al., 2019). Especially for understanding the mechanisms underlying G$\times$G and G$\times$E interactions, even the most high through-put screening approaches will likely prove sisyphean. The more promising approach is improved integration between all the disciplines of biology, for example to mechanistically understand how changes in the amino-acids of a protein impact its function and the entire organismal physiology, including its effects on fitness (Burford Reiskind et al., 2021).

Readers already knowledgeable in the field of predicting evolution will notice that I have not listed either historical contingency or mutational stochasticity as specific challenges to predicting evolution. The challenge from historical contingency is summarized by Gould's thought experiment of replaying the tape of life (Gould, 1989). If we were to wind back and replay the history of life over again, how likely would it be that evolution would proceed the same way? From his paleontological work on the Cambrian explosion, Gould concluded that evolution over very long time-scales is an unpredictable building-up process that depends on chance past events. Subsequent experimental evolution studies have shown that – at short time-scales – Gould was both right and wrong. The occurrence of specific random mutations and their order is important, especially when epistasis is present, but the same selection pressures often do result in the same evolutionary outcomes, especially for traits that are more closely associated with fitness (Travisano et al., 1995; Joshi, Castillo, and Mueller, 2003; Simões et al., 2008; Tenaillon et al., 2012). Therefore, I see the problem of historical contingency as a property that manifests due to a combination of the above three challenges (G$\times$G$\times$E acting in unpredictable environments, and the

need for deep functional knowledge).

Turning to the challenge of mutational stochasticity, experimental evolution studies as well as work on parallel and convergent evolution have found many instances where novel mutations consistently hit the same genomic targets (Colosimo et al., 2005; Tenaillon et al., 2012; Renaut, Owens, and Rieseberg, 2014; Ramiro, Costa, and Gordo, 2016; Monroe et al., 2016). Populations of larger size that have a larger input of mutations per generation, sexual populations that are less impacted by clonal interference (also called Hill-Robertson effect; reviewed in Otto, 2021), and adaptive phenotypes underlain by larger mutational targets are all less likely to be impacted by mutational stochasticity. Therefore we can expect the role of mutational stochasticity to vary depending on the system of interest (some examples of mutation-limited systems include: Tegze et al., 2012; Fitzpatrick et al., 2014; Langerhans, 2018). Indeed, the systems of interest studied in this thesis all focus on the fate of segregating genetic diversity (Chapter 2) and measuring the fitnesses of existing mutations (Chapters 3-5). As research progresses we may find some evolving systems may be resistant to prediction because their dynamics are in fact unpredictable even at short time-scales, for example due to low segregating diversity and intermediate mean waiting times for new mutations (i.e., high mutational stochasticity), clonal interference (Sniegowski and Gerrish, 2010; but see Good et al., 2017), or chaotic environmental/ecological dynamics (Barraclough, 2015). These systems do not pose an existential challenge to the endeavour of predicting evolution because their inherent unpredictability itself can be known. Therefore, one aspect of research on predicting evolution involves identifying populations and biological systems that are amenable to prediction (Burford Reiskind et al., 2021).

In addition to the three scientific challenges of predicting evolution discussed above, there is a final, practical challenge: How useful would it actually be to predict evolution? In other words, if predicting evolution is a technology, what are the usage situations for this technology. As explained by Burford Reiskind et al., 2021, evolutionary prediction technologies become more useful as their (putative) predictions become more precise. It is trivial to know that evolution will happen; it is very useful to know the probability of evolving different phenotypic traits, their likely underlying genetic targets, and their rate of / waiting time to evolution. More concretely using the example of ABR, health practitioners already know that frequent application of antibiotics leads to higher carriages of ABR's in the population. This explanatory (i.e., non-predictive) understanding of ABR alone has led to better antibiotic stewardship and a decreased prevalence of ABR for humans and animals (Bell et al.,

2014; Tang et al., 2017; Shen et al., 2020; but see Enne et al., 2001; Sundqvist et al., 2010). However, precise evolutionary predictions promise bespoke solutions to the problem of ABR. For example, matching of the antibiotic to the genotype of the parasite, patient, and their microbiota, as well as specified order and timing of antibiotic treatments that maximize patient outcomes at the same time as minimizing ABR risk to the whole community. Scientific evolution research with the goal of explanation is often sufficient to provide us with general insights, but evolutionary prediction technologies could capitalize on this explanatory power to provide precise solutions. This will depend on the situation, of course. Maybe sometimes prediction can help us fine-tune between several good choices (i.e., a low risk decision). If the margin of error for implementing the control measure is narrow (i.e., it can have either good outcomes or bad outcomes) then it's even more important to either find lower risk solutions or have a spot-on prediction.

This thesis focuses on the challenge to predicting evolution posed by genetic and environmental interactions. In section 3, I traced the history of gene and environment interactions in population genetics to highlight that our paradigm is founded on ignoring these types of interactions. For evolutionary biologists working before *in silico* methods were readily available, these simplifying assumptions were made primarily to facilitate the math of population genetics. Fitness landscapes theory was built to study epistatic interactions and has blossomed in the last 20 years, not only with simulations but, more importantly, with empirical data to inform these models.

4.2 Describing genetic and environmental interactions with fitness landscapes theory

The fitness landscape was first proposed as an analogy, by Wright (Provine, 1986), to illustrate the role of epistatic interactions in adaptation. Under Wright's shifting balance hypothesis, demes of small population size on a local fitness peak could more easily cross a valley to reach a nearby, taller peak as compared to large populations. Fitness landscapes have become a quantitative model for investigating the role of G$\times$G on evolutionary outcomes (Fragata et al., 2019; Bank, 2022). Fitness landscapes are a map of the relationship between genotype or phenotype and fitness. Since evolution can generally be construed as a hill-climbing algorithm, a detailed fitness landscape map could be used to predict where, and how fast, a population is evolutionarily headed.

Most fitness landscapes work has been theoretical and will likely continue to

require a strong modeling or *in silico* component. This is because the size of the possible genotypic space is immense (e.g., all the possible amino-acid combinations for a 70 residue length polypeptide (20^{70}) sum to more atoms than estimated in the universe; Fisher, 1930; Szendro et al., 2013) and therefore can never be completely assayed empirically.

While studies on the distribution of fitness effects of new, single mutations anywhere in the genome can provide information about the shape of a fitness landscape (e.g., Trindade, Sousa, and Gordo, 2012; Harmand et al., 2017), recent technical advancements allow the estimation of sizable fitness landscapes of targeted genomic regions, either one (e.g., Hietpas et al., 2012) or more (e.g., Khan et al., 2011; Bank et al., 2015) mutation-steps away from a 'WT' genotype. The shapes of empirical fitness landscapes may be different than the theoretical fitness landscapes that have been considered (Blanquart and Bataillon, 2016; Das et al., 2020). Moreover, different types of mutational classes and adaptive regimes may result in fitness landscapes with different characteristics. Therefore development of fitness landscape models that are informed by biology is an active area of research (reviewed in De Visser and Krug, 2014; Fragata et al., 2019; Bank, 2022).

Extensions of the fitness landscapes toolkit attempt to describe how fitness landscapes change across environments in order to study G$\times$E and G$\times$G$\times$E and possible higher-order interactions. The most commonly used model, Fisher's geometric model, is a phenotype-to-fitness landscape that assumes a single, multivariate-Gaussian distributed optimum among the phenotypic dimensions that contribute to fitness. The shape, slope, and distance of the optimum to the fixed 'WT' genotype changes with the environment, allowing prediction of the distribution of fitness effects and rate of evolutionary change for different environments (Blanquart et al., 2014; Martin and Lenormand, 2015). Many other theoretical fitness landscape models also exist (reviewed in Fragata et al., 2019) but it is either unclear how to model environmental changes using these landscapes or the effect of environmental changes on these landscapes has not yet been examined as extensively (Bank, 2022). In contrast to fitness landscape models that attempt to capture some generic features of biological systems, mechanistic models that incorporate functional genetic information specific to the system of interest, as well as environmental effects, can better explain the mapping of genotypes/phenotypes to fitness between environments (e.g., Levin et al., 2017; Guerrero et al., 2019; Das et al., 2020; Pinheiro et al., 2021). This theoretical work is complemented by empirical fitness landscapes were the fitness of the same genotypes are investigated in different environments (e.g., Hall et al.,

2019). In summary, there are multiple avenues of research possible for incorporating G×E interactions with fitness landscape models in order to study higher-order genetic and environmental interactions.

5 Thesis questions

My thesis focuses on predicting evolution across different environments. Broadly, I am interested in predicting how the genotypes present in a population are likely to change over time. Specifically, I focus on changes like nucleotide substitutions and gene deletions, but I ignore larger chromosomal changes, like the evolution of chromosome structure or ploidy. As a result of my focus on the genotype, evolution is here defined as the change in allele frequencies over time (Lewontin, 1974). Corresponding with predicting genotypic changes over time, the fitness landscapes used in my thesis are predominantly genotype-to-fitness landscapes.

My operational definition of 'environment' is more general than typically used in biology. Expanding on the gene's-eye view first introduced by Fisher (see section 3.4 above; Fisher, 1930; Edwards, 2014), I consider any condition external to the focal genotypic fitness landscape as being part of the environment. For example, under this definition a change in ploidy constitutes a change in the environment. This is a departure from the definition of environment that has typically been used in biology: any biotic or abiotic condition that is external to the organism (e.g., Losos, 2011; but see Lewontin, 1985). Nevertheless, this modeling assumption is predominant throughout population genetics. My thesis expands on the (usually biallelic) single locus gene's-eye view because I consider the (two-locus) fitness landscape as focal. I ask whether it is possible to make evolutionary predictions on a focal fitness landscape when different environments are considered. In Chapter 2 the different environments are different ploidy levels, while in Chapters 3 and 5 the different environments are abiotic conditions that are external to the organism (i.e., temperature and antibiotic).

The big-picture question that emerges given the focus of my thesis is:
Is possible to predict evolution on fitness landscapes in different environments? This is a very broad question. Some other, perhaps smaller, questions that emerge in attempting to answer the general question are:

1. Can we build fitness landscape models that capture the complex fitness effects observed in natural populations?
It is difficult to predict evolution in natural systems because there are often a diversity

of evolutionary forces acting on a population at any given time. In Chapter 2 we study a hybrid population of Finnish wood ants, *Formica polyctena* $\times$ *F. aquilonia*. These ants are known from other studies to exhibit ancestral hybridization, ongoing gene flow, changes in population size (Portinha et al., 2022), environmentally-dependent selection (Kulmuni et al., 2020; Martin-Roy et al., 2021), age-specific selection, sexual selection, and mate choice (Kulmuni and Pamilo, 2014). Focusing only on the latter three evolutionary factors, we used fitness landscape models to predict what will happen to the hybrid status of this population over time. We found that a rugged fitness landscape could explain the observed data but had little power to distinguish between future evolutionary outcomes.

2. What do empirical fitness landscapes of untested genetic combinations look like?

As reviewed above, models of theoretical fitness landscapes need to be informed by empirically observed fitness landscapes. Further, predicting evolution requires different kinds of data than explaining past evolutionary trajectories (Shaw, 2019). Fitness landscapes investigating combinations of mutations along an adaptive walk have different properties than those combining random mutations (De Visser and Krug, 2014; Bank, 2022). As the goal of predicting evolution is to look forward in time, it is important for us to better understand the properties of fitness landscapes that combine evolutionarily untested combinations, especially combinations of segregating variation (i.e., as opposed to *de novo* mutations).

In Chapters 3 and 5 I move away from the complexity of natural populations in order to investigate selection using a model organism, *Escherichia coli*, in controlled laboratory conditions. We describe an intergenic fitness landscape of gene knockouts (KOs) and single nucleotide ABR mutations. To our knowledge, this is the first empirical fitness landscapes study to look at between-gene epistasis among evolutionarily untested combinations of KOs and single nucleotide polymorphisms (SNPs). Moreover, we selected these mutations because they are segregating variants in wild *E. coli* populations that. We found that, in benign environments, the ruggedness of this empirical fitness landscape was similar to what has previously been observed for interacting single nucleotide mutations.

3. How do empirical fitness landscapes of untested genetic combinations change across environments?

Previous empirical fitness landscape studies have usually assayed the same set of genotypes in different, unrelated environments (e.g., Hall et al., 2019; Flynn et al., 2020; Hinz et al., in prep.). Theoretical fitness landscape models hypothesize that

fitness landscape features may change in a predictable way across related environments, for example along a gradient of increasing stress. In Chapters 3 and 5 we study the same set of genotypes across all possible combinations of two environmental gradients: antibiotic and high temperature. Since an environmental gradient provides a naive/logical ranking of selective pressures, this experimental design allowed us to assess the predictability of fitness landscapes across different but related environments and environmental combinations. We found that the fitness landscapes become smoother at higher antibiotic concentrations.

4. How do we measure fitness?

In order to characterize fitness landscapes and understand where a population is evolutionarily headed, we need to be able to obtain accurate estimates of fitnesses. Furthermore, due to the large size of the genome and the abundance of different environments, fitness measurements should be gathered in a high-throughput way. In Chapter 4 we look at a common method for estimating fitness in microbial populations: growth rate estimates from growth curve data. I intend to expand on this question with the data from Chapter 5, where we compare the competitive fitness estimates from Chapter 3 with population growth phenotypes.

5. Should we predict evolution?

From Darwin to Gould, many scientific arguments against predicting evolution focus on the difficulty of the task. The recent flurry of opinion pieces, review papers, and special issues touting the need for evolutionary predictions (Lässig, Mustonen, and Walczak, 2017; Loga and Gerlinger, 2017; Russell and Jong, 2017; Shaw, 2019; Nghe et al., 2020; Lässig and Mustonen, 2020; Burford Reiskind et al., 2021; Thompson et al., 2021; Iram et al., 2021; Wortel et al., 2022; Baltazar-Soares, Brans, and Eizaguirre, 2022; The Royal Society, 2023) suggest that we are closer than ever to predicting evolution in select systems and time-scales, with some predictions already being applied (Andersson et al., 2020; Feder et al., 2021; Rieder et al., in press). In the final chapter, I turn to the philosophical implications of predicting evolution.

References

Agor, J. K. and O. Y. Özaltın (2018). "Models for predicting the evolution of influenza to inform vaccine strain selection". In: *Hum. Vaccines Immunother.* 14.3, pp. 678–683. URL: https://doi.org/10.1080/21645515.2017.1423152.

Ågren, J. A. (2021). "Sewall Wright's criticism of the gene's-eye view of evolution". In: *Evolution* 75.10, pp. 2326–2334. URL: https://doi.org/10.1111/evo.14334.

Allen, G. (1955). "Perspectives in population eugenics". In: *Eugenics Quarterly* 2.2, pp. 90–97. URL: https://doi.org/10.1080/19485565.1955.9987230.

Aminov, R. I. (2010). "A brief history of the antibiotic era: lessons learned and challenges for the future". In: *Front. Microbiol.* 1, p. 134. URL: https://doi.org/10.3389/fmicb.2010.00134.

Andersson, D. I. and D. Hughes (2014). "Microbiological effects of sublethal levels of antibiotics". In: *Nat. Rev. Microbiol.* 12.7, pp. 465–478. URL: https://doi.org/10.1038/nrmicro3270.

Andersson, D. I. et al. (2020). "Antibiotic resistance: turning evolutionary principles into clinical reality". In: *FEMS Microbiol. Rev.* 44.2, pp. 171–188. URL: https://doi.org/10.1093/femsre/fuaa001.

Ashe, A., V. Colot, and B. P. Oldroyd (2021). "How does epigenetics influence the course of evolution?" In: *Philos. Trans. R. Soc. B* 376.1826, p. 20200111. URL: https://doi.org/10.1098/rstb.2020.0111.

Avice, G. et al. (2018). "Evolution of atmospheric xenon and other noble gases inferred from Archean to Paleoproterozoic rocks". In: *Geochim. Cosmochim. Acta* 232, pp. 82–100. URL: https://doi.org/10.1016/j.gca.2018.04.018.

Baltazar-Soares, M., K. I. Brans, and C. Eizaguirre, eds. (2022). *Evol. Appl.* 14.10: *Special issue: human induced evolution.* URL: https://onlinelibrary.wiley.com/toc/17524571/2021/14/10.

Bank, C. (2022). "Epistasis and Adaptation on Fitness Landscapes". In: *Annu. Rev. Ecol. Evol. Syst.* 53, pp. 457–479. URL: https://doi.org/10.1146/annurev-ecolsys-102320-112153.

Bank, C. et al. (2015). "A systematic survey of an intragenic epistatic landscape". In: *Mol. Biol. Evol.* 32.1, pp. 229–238. URL: https://doi.org/10.1093/molbev/msu301.

Baquero, F., A. P. Tedim, and T. M. Coque (2013). "Antibiotic resistance shaping multilevel population biology of bacteria". In: *Front. Microbiol.* 4, p. 15. URL: https://doi.org/10.3389/fmicb.2013.00015.

Barnes, B. D., J. A. Sclafani, and A. Zaffos (2021). "Dead clades walking are a pervasive macroevolutionary pattern". In: *PNAS* 118.15, e2019208118. URL: https://doi.org/10.1073/pnas.2019208118.

Barraclough, T. G. (2015). "How do species interactions affect evolutionary dynamics across whole communities?" In: *Annu. Rev. Ecol. Evol. Syst.* 46, pp. 25–48. URL: https://doi.org/10.1146/annurev-ecolsys-112414-054030.

— (2019). "Species matter for predicting the functioning of evolving microbial communities – An eco-evolutionary model". In: *PLOS One* 14.8, e0218692. URL: https://doi.org/10.1371/journal.pone.0218692.

Barrett, P. H. (1987). "Charles Darwin's notebooks". In: British Museum (Natural History). Chap. Essay on theology and natural selection. URL: http://darwin-online.org.uk/content/frameset?pageseq=1&itemID=F1582&viewtype=text.

Bashford, A (2010). "The Oxford Handbook of the History of Eugenics". In: ed. by A Bashford and P. Levine. Oxford University Press. Chap. Epilogue: where did eugenics go?

Bashford, A. and P. Levine, eds. (2010). *The Oxford Handbook of the History of Eugenics*. 1st. Oxford Academic. URL: https://doi.org/10.1093/oxfordhb/9780195373141.001.0001.

Bateson, W. (1889). "VII. On some variations of *Cardium edule* apparently correlated to the conditions of life". In: *Phil. Trans. R. Soc. B* 180, pp. 297–330. URL: https://doi.org/10.1098/rstb.1889.0007.

— (1909). "Heredity and variation in modern lights". In: *Darwin and modern science*, pp. 85–101.

Bateson, W., E. R. Saunders, and R. C. Punnett (1904). *Report II. Experimental Studies in the Physiology of Heredity*. Tech. rep. Reports to the Evolution Committee of the Royal Society. URL: https://archive.org/details/RoyalSociety.ReportsToTheEvolutionCommittee.ReportIi.Experimental.

— (1906). *Report III. Experimental Studies in the Physiology of Heredity*. Tech. rep. Reports to the Evolution Committee of the Royal Society.

— (1908). *Report IV. Experimental Studies in the Physiology of Heredity*. Tech. rep. Reports to the Evolution Committee of the Royal Society.

Bell, B. G. et al. (2014). "A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance". In: *BMC Infect. Dis.* 14.1, pp. 1–25. URL: https://doi.org/10.1186/1471-2334-14-13.

Bell, G. (1982). *The Masterpiece Of Nature: The Evolution and Genetics of Sexuality*. Berkeley, California: University of California Press.

Blanquart, F. and T. Bataillon (2016). "Epistasis and the structure of fitness landscapes: are experimental fitness landscapes compatible with Fisher's geometric model?" In: *Genetics* 203.2, pp. 847–862. URL: https://doi.org/10.1534/genetics.115.182691.

Blanquart, F. et al. (2014). "Properties of selected mutations and genotypic landscapes under Fisher's geometric model". In: *Evolution* 68, pp. 3537–3554. URL: https://doi.org/10.1111/evo.12545.

Bodmer, W. et al. (2021). "The outstanding scientist, R. A. Fisher: his views on eugenics and race". In: *Heredity* 126.4, pp. 565–576. URL: https://doi.org/10.1038/s41437-020-00394-6.

Bowler, P. J. (1983). *The Eclipse of Darwinism: Anti-Darwinian Evolution Theories in the Decades Around 1900*. 1st. JHU Press.

Branch, H. A. et al. (2022). "Discussions of the "not so fit": how ableism limits diverse thought and investigative potential in evolutionary biology". In: *Am. Nat.* 200.1, pp. 101–113. URL: https://doi.org/10.1086/720003.

Brochado, A. R. et al. (2018). "Species-specific activity of antibacterial drug combinations". In: *Nature* 559.7713, pp. 259–263. URL: https://doi.org/10.1038/s41586-018-0278-9.

Brown, F. B. (1986). "The evolution of Darwin's theism". In: *J. Hist. Biol.*, pp. 1–45. URL: https://www.jstor.org/stable/4330957.

Buchanan, R. D. and J. Bradley (2017). ""Darwin's Delay": A Reassessment of the Evidence". In: *Isis* 108, pp. 529–552. URL: https://doi.org/10.1086/694183.

Bunge, M. (1966). "Contributions to a Philosophy of Technology". In: ed. by F. Rapp. Dordrecht: Springer. Chap. Technology as Applied Science. URL: https://doi.org/10.1007/978-94-010-2182-1_2.

Burford Reiskind, M. O. et al. (2021). "Nothing in evolution makes sense except in the light of biology". In: *Bioscience* 71.4, pp. 370–382. URL: https://doi.org/10.1093/biosci/biaa170.

Colosimo, P. F. et al. (2005). "Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles". In: *science* 307.5717, pp. 1928–1933. URL: https://www.science.org/doi/10.1126/sciadv.abg5285.

Cowan, R. S. (1972). "Francis Galton's statistical ideas: the influence of eugenics". In: *Isis* 63.4, pp. 509–528. URL: https://doi.org/10.1086/351000.

Coyne, J. A. and H. A. Orr (2004). *Speciation*. 1st. Sunderland, MA: Sinauer Associates, Inc.

Crow, J. F. (2010). "Wright and Fisher on inbreeding and random drift". In: *Genetics* 184.3, pp. 609–611. URL: https://doi.org/10.1534/genetics.109.110023.

Crow, J. F. and M. Kimura, eds. (1970). *An Introduction to Population Genetics Theory*. Caldwell, New Jersey: The Blackburn Press.

Cvijović, I. et al. (2015). "Fate of a mutation in a fluctuating environment". In: *PNAS* 112.36, E5021–E5028. URL: https://doi.org/10.1073/pnas.1505406112.

Darwin, C. (1859). *The Origin of Species by means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 1st. Albemarle Street, London: John Murray. URL: https://archive.org/details/darwin-online_1859_Origin_F373/.

— (1868). *The Variation of Animals and Plants Under Domestication*. 1st. Albemarle Street, London: John Murray.

— (1871a). *The Descent of Man and Selection in Relation to Sex*. 1st. London: John Murray.

— (1871b). *The Origin of Species by means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 5th. 549 & 551 Broadway, New York: D. Appleton and company. URL: https://archive.org/details/onoriginofspeci00darw/.

— (1876). *The Origin of Species by means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. 6th. 549 & 551 Broadway, New York: D. Appleton and company. URL: https://archive.org/details/onoriginspecies00dallgoo

Das, S. G. et al. (2020). "Predictable properties of fitness landscapes induced by adaptational tradeoffs". In: *eLife* 9, e55155. URL: https://doi.org/10.7554/eLife.55155.

Davis, L. K. (2022). "Weldon, Bateson, and the origins of genetics: Reflections on the unraveling and rebuilding of a scientific community". In: *PLOS Genet.* 18.10, e1010379. URL: https://doi.org/10.1371/journal.pgen.1010379.

Dawkins, R. (1976). *The Selfish Gene*. London, UK: Oxford University Press.

de Lespinois, P. (Writer) and V. (Director) Coules (2003). *The Future Is Wild*. ZDF Enterprises. URL: https://www.thefutureiswild.com/.

De Visser, J. A. G. M. and J. Krug (2014). "Empirical fitness landscapes and the predictability of evolution". In: *Nat. Rev. Genet.* 15.7, pp. 480–490. URL: https://doi.org/10.1038/nrg3744.

Delzell, D. A. P. and C. D. Poliak (2013). "Karl Pearson and eugenics: Personal opinions and scientific rigor". In: *Sci. Eng. Ethics* 19, pp. 1057–1070. URL: https://doi.org/10.1007/s11948-012-9415-2.

Dobzhansky, T. H. (1936). "Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids". In: *Genetics* 21.2, p. 113. URL: https://doi.org/10.1093/genetics/21.2.113.

Durão, P., R. Balbontín, and I. Gordo (2018). "Evolutionary mechanisms shaping the maintenance of antibiotic resistance". In: *Trends Microbiol.* 26.8, pp. 677–691. URL: https://doi.org/10.1016/j.tim.2018.01.005.

D'Costa, V. M. et al. (2011). "Antibiotic resistance is ancient". In: *Nature* 477.7365, pp. 457–461. URL: https://doi.org/10.1038/nature10388.

Edwards, A. W. F. (2014). "R. A. Fisher's gene-centred view of evolution and the Fundamental Theorem of Natural Selection". In: *Biol. Rev.* 89.1, pp. 135–147. URL: https://doi.org/10.1111/brv.12047.

Elgin, M. and E. Sober (2017). "Popper's shifting appraisal of evolutionary theory". In: *HOPOS* 7.1, pp. 31–55. URL: https://doi.org/10.1086/691119.

Enne, V. I. et al. (2001). "Persistence of sulphonamide resistance in *Escherichia coli* in the UK despite national prescribing restriction". In: *Lancet* 357.9265, pp. 1325–1328. URL: https://doi.org/10.1016/S0140-6736(00)04519-0.

Ewens, W. J. (1967). "The probability of survival of a new mutant in a fluctuating environment". In: *Heredity* 22, 438–443. URL: https://doi.org/10.1038/hdy.1967.53.

Eyre-Walker, A. and P. D. Keightley (2007). "The distribution of fitness effects of new mutations". In: *Nat. Rev. Genet.* 8.8, pp. 610–618. URL: https://doi.org/10.1038/nrg2146.

Falk, R. (2000). "Thinking about Evolution — Historical, Philosophical, and Political Perspectives". In: ed. by R. S. Singh et al. Cambridge: Cambridge University Press. Chap. Can the norm of reaction save the gene concept?

Fancher, R. E. (2009). "Scientific cousins: the relationship between Charles Darwin and Francis Galton." In: *Am. Psychol.* 64.2, p. 84. URL: https://psycnet.apa.org/doi/10.1037/a0013339.

Feder, A. F. et al. (2021). "Understanding patterns of HIV multi-drug resistance through models of temporal and spatial drug heterogeneity". In: *eLife* 10, e69032. URL: https://doi.org/10.7554/eLife.69032.

Fisher, R. A. (1918). "The correlation between relatives on the supposition of mendelian inheritance". In: *Trans. R. Soc. Edinb.* 53, pp. 399–433.

— ed. (1930). *The Genetical Theory of Natural Selection*. 1st. Cambridge, UK: Clarendon Press.

— (1934). *Report of enquiry into the children of mental defectives*. Tech. rep. H.M.S.O, London, UK: Report of the departmental committee on sterilisation, pp. 60–74.

Fitzgerald, D. M., P. J. Hastings, and S. M. Rosenberg (2017). "Stress-induced mutagenesis: implications in cancer and drug resistance". In: *Annu. Rev. Cancer Biol.* 1, pp. 119–140. URL: https://doi.org/10.1146/annurev-cancerbio-050216-121919.

Fitzpatrick, S. W. et al. (2014). "Parallelism isn't perfect: could disease and flooding drive a life-history anomaly in Trinidadian guppies?" In: *Am. Nat.* 183.2, pp. 290–300. URL: https://doi.org/10.1086/674611.

Flynn, J. M. et al. (2020). "Comprehensive fitness maps of Hsp90 show widespread environmental dependence". In: *eLife* 9, e53810. URL: https://doi.org/10.7554/eLife.53810.

Fowler, D. M. et al. (2010). "High-resolution mapping of protein sequence-function relationships". In: *Nat. Methods* 7.9, pp. 741–746. URL: https://doi.org/10.1038/nmeth.1492.

Fragata, I. et al. (2019). "Evolution in the light of fitness landscape theory". In: *Trends Ecol. Evol.* 34.1, pp. 69–82. URL: https://doi.org/10.1016/j.tree.2018.10.009.

Frenkel, E. M., B. H. Good, and M. M. Desai (2014). "The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations". In: *Genetics* 196.4, pp. 1217–1226. URL: https://doi.org/10.1534/genetics.113.160069.

Furusawa, C., T. Horinouchi, and T. Maeda (2018). "Toward prediction and control of antibiotic-resistance evolution". In: *Curr. Opin. Biotechnol.* 54, pp. 45–49. URL: https://doi.org/10.1016/j.copbio.2018.01.026.

Galton, F. (1865). "Hereditary Talent and Character". In: *Macmillan's Magazine* 12, pp. 157–166, 318–327.

— (1869). *Hereditary Genius: An Inquiry Into Its Laws and Consequences*. 1st. England: Macmillan Publishers.

— (1871). "Experiments in pangenesis, by breeding from rabbits of a pure variety, into whose circulation blood taken from other varieties had previously been largely transfused". In: *Proc. Roy. Soc.* 19, 393–410.

— (1883). *Inquiries into Human Faculty and Its Development*. 1st. London: J. M. Dent & Co.

— (1901). "The Possible Improvement of the Human Breed Under the Existing Conditions of Law and Sentiment". In: *Man* 1, pp. 161–164. URL: https://doi.org/10.2307/2840363.

— (1909). "Essays in Eugenics". In: London: Eugenics Education Society. Chap. The Possible Improvement of the Human Breed, under the Existing Conditions of Law and Sentiment.

Ghenu, A. H. et al. (2016). "Multicopy gene family evolution on primate Y chromosomes". In: *BMC Genom.* 17.1, pp. 1–17. URL: https://doi.org/10.1186/s12864-015-2187-8.

Gil-Gil, T. et al. (2021). "Antibiotic resistance: time of synthesis in a post-genomic age". In: *Comput. Struct. Biotechnol. J.* 19, pp. 3110–3124. URL: https://doi.org/10.1016/j.csbj.2021.05.034.

Gillespie, J. H. (1972). "The effects of stochastic environments on allele frequencies in natural populations". In: *Theor. Popul. Biol.* 3.3, pp. 241–248. URL: https://doi.org/10.1016/0040-5809(72)90001-9.

Godfrey-Smith, P. (2003). *Theory and Reality: an introduction to the philosophy of science*. 1st. The University of Chicago Press.

Goldstone, R. J. and D. G. E. Smith (2017). "A population genomics approach to exploiting the accessory 'resistome' of *Escherichia coli*". In: *Microb. Genom.* 3.4. URL: https://doi.org/10.1099/mgen.0.000108.

Good, B. H. et al. (2017). "The dynamics of molecular evolution over 60,000 generations". In: *Nature* 551.7678, pp. 45–50. URL: https://doi.org/10.1038/nature24287.

Gordo, I., L. Perfeito, and A. Sousa (2011). "Fitness effects of mutations in bacteria". In: *J. Mol. Microbiol. Biotechnol.* 21.1-2, pp. 20–35. URL: https://doi.org/10.1159/000332747.

Gould, S. J. (1977). "Eternal metaphors of palaeontology". In: *Developments in palaeontology and stratigraphy*. Vol. 5. Elsevier, pp. 1–26. URL: https://doi.org/10.1016/S0920-5446(08)70321-2.

— (1989). *Wonderful life: the Burgess Shale and the nature of history*. New York, New York: W. W. Norton & Company.

— (1991). "The smoking gun of eugenics". In: *Nat. Hist.* 12, pp. 8–17.

— (1993). "Eight Little Piggies: reflections in natural history". In: New York, New York: W. W. Norton & company. Chap. Darwin and Paley meet the invisible hand.

— (1996). *The Mismeasure of Man*. 2nd.

Gould, S. J. and E. A. Lloyd (1999). "Individuality and adaptation across levels of selection: how shall we name and generalize the unit of Darwinism?" In: *PNAS* 96.21, pp. 11904–11909. URL: https://doi.org/10.1073/pnas.96.21.11904.

Grafen, A. (2003). "Fisher the evolutionary biologist". In: *J. R. Stat. Soc. D (The Statistician)* 52.3, pp. 319–329. URL: https://doi.org/10.1111/1467-9884.00362.

Gross, M. (2019). "The race against antibiotics resistance". In: *Curr. Biol.* 29 (18), R859–R861. URL: https://doi.org/10.1016/j.cub.2019.09.003.

Gross, M. and M. B. Averill (2003). "Discovering Reality". In: ed. by S. Harding and M. B. Hintikka. 2nd. Springer Science+Business Media. Chap. Evolution and Patriarchal Myths of Scarcity and Competition.

Guerrero, R. F. et al. (2019). "Proteostasis environment shapes higher-order epistasis operating on antibiotic resistance". In: *Genetics* 212, pp. 565–575. URL: https://doi.org/10.1534/genetics.119.302138.

Haldane, J. B. S. (1932). *The Causes of Evolution*. London: Longmans, Green, and Co. URL: https://archive.org/details/causesofevolutio00hald_0/.

Hall, A. E. et al. (2019). "Environment changes epistasis to alter trade-offs along alternative evolutionary paths". In: *Evolution* 73, pp. 2094–2105. URL: https://doi.org/10.1111/evo.13825.

Hall, B. G. (2004). "Predicting the evolution of antibiotic resistance genes". In: *Nat. Rev. Microbiol.* 2.5, pp. 430–435. URL: https://doi.org/10.1038/nrmicro888.

Hamilton, W. D. (1964). "The genetical evolution of social behaviour. I". In: *J. Theor. Biol.* 7.1, pp. 1–16. URL: https://doi.org/10.1016/0022-5193(64)90038-4.

Harmand, N. et al. (2017). "Fisher's geometrical model and the mutational patterns of antibiotic resistance across dose gradients". In: *Evolution* 71.1, pp. 23–37. URL: https://doi.org/10.1111/evo.13111.

Hietpas, R. et al. (2012). "Fitness analyses of all possible point mutations for regions of genes in yeast". In: *Nat. Protoc.* 7.7, pp. 1382–1396. URL: https://doi.org/10.1038/nprot.2012.069.

Hinz, A. et al. (in prep.). In.

Hubbard, R. (2003). "Discovering Reality". In: ed. by S. Harding and M. B. Hintikka. 2nd. Springer Science+Business Media. Chap. Have Only Men Evolved?

Huneman, P. (2014). "Mapping an expanding territory: computer simulations in evolutionary biology". In: *Hist. Philos. Life Sci.* 36, pp. 60–89. URL: https://doi.org/10.1007/s40656-014-0005-2.

Huxley, J. (1942). *Evolution The Modern Synthesis*. London: George Allen & Unwin Ltd. URL: https://archive.org/details/in.ernet.dli.2015.280031/.

Ihde, D. (2004). "Philosophical Problems Today". In: ed. by P. Kemp. Vol. 3. Dordrecht: Springer. Chap. Philosophy of Technology. URL: https://doi.org/10.1007/1-4020-3027-4_3.

Iram, S. et al. (2021). "Controlling the speed and trajectory of evolution with counter-diabatic driving". In: *Nat. Phys.* 17.1, pp. 135–142. URL: https://doi.org/10.1038/s41567-020-0989-3.

Joshi, A., R. B. Castillo, and L. D. Mueller (2003). "The contribution of ancestry, chance, and past and ongoing selection to adaptive evolution". In: *J. Genet.* 82, pp. 147–162. URL: https://doi.org/10.1007/BF02715815.

Karageorgi, M. et al. (2019). "Genome editing retraces the evolution of toxin resistance in the monarch butterfly". In: *Nature* 574.7778, pp. 409–412. URL: https://doi.org/10.1038/s41586-019-1610-8.

Karlin, S. and B. Levikson (1974). "Temporal fluctuations in selection intensities: case of small population size". In: *Theor. Popul. Biol.* 6.3, pp. 383–412. URL: https://doi.org/10.1016/0040-5809(74)90017-3.

Keilhack, L. (1909). "Hydrobiologische Referat review of Woltereck (1909)". Trans. by Maximilian Oliver Press 2020. In: *Internationale Revue der gesamten Hydrobiologie und Hydrographie* 11.4 (5), pp. 770–778. URL: https://files.osf.io/v1/resources/5szyd/providers/osfstorage/5eac58da0ffc0500236f7a5d?action=download&direct&version=1.

Keller, E. F. (1996). *Reflections on gender and science*. 2nd. Yale University Press.

Kern, A. D. and M. W. Hahn (2018). "The neutral theory in light of natural selection". In: *Mol. Biol. Evol.* 35.6, pp. 1366–1371. URL: https://doi.org/10.1093/molbev/msy092.

Kevles, D. J. (1985). *In the name of eugenics: genetics and the uses of human heredity*. 1st. California: University of California Press.

Khan, A. I. et al. (2011). "Negative epistasis between beneficial mutations in an evolving bacterial population". In: *Science* 332, pp. 1193–1196. URL: https://doi.org/10.1126/science.1203801.

Kimura, M. (1958). "On the change of population fitness by natural selection". In: *Heredity* 12.2, pp. 145–167. URL: https://doi.org/10.1038/hdy.1958.21.

Knopp, M. and D. I. Andersson (2018). "Predictable phenotypes of antibiotic resistance mutations". In: *mBio* 9.3, e00770–18. URL: https://doi.org/10.1128/mBio.00770-18.

Knöppel, A., J. Näsvall, and D. I. Andersson (2017). "Evolution of antibiotic resistance without antibiotic exposure". In: *Antimicrob. Agents Chemother.* 61.11, e01495–17. URL: https://doi.org/10.1128/AAC.01495-17.

Kuhn, T. S. (1996). *The Structure of Scientific Revolutions*. 3rd. University of Chicago Press.

Kulmuni, J. and P. Pamilo (2014). "Introgression in hybrid ants is favored in females but selected against in males". In: *PNAS* 111.35, pp. 12805–12810. URL: https://doi.org/10.1073/pnas.1323045111.

Kulmuni, J. et al. (2020). "Instability of natural selection at candidate barrier loci underlying speciation in wood ants". In: *Mol. Ecol.* 29.20, pp. 3988–3999. URL: https://doi.org/10.1111/mec.15606.

Lamsdell, J. C. et al. (2017). "Phylogenetic paleoecology: tree-thinking and ecology in deep time". In: *Trends Ecol. Evol.* 32.6, pp. 452–463. URL: https://doi.org/10.1016/j.tree.2017.03.002.

Langerhans, R. B. (2018). "Predictability and parallelism of multitrait adaptation". In: *J. Hered.* 109.1, pp. 59–70. URL: https://doi.org/10.1093/jhered/esx043.

Lässig, M. and V. Mustonen (2020). "Eco-evolutionary control of pathogens". In: *PNAS* 117.33, pp. 19694–19704. URL: https://doi.org/10.1073/pnas.1920263117.

Lässig, M., V. Mustonen, and A. M. Walczak (2017). "Predicting evolution". In: *Nat. Ecol. Evol.* 1.3, p. 0077. URL: https://doi.org/10.1038/s41559-017-0077.

Latour, B. and S. Woolgar (1987). *Laboratory Life: The Construction of Scientific Facts*. Princeton, New Jersey: Princeton University Press.

Leiss, W. (1994). *Domination over nature*. 2nd. University of Toronto Press.

Levin, B. R. et al. (2017). "A numbers game: ribosome densities, bacterial growth, and antibiotic-mediated stasis and death". In: *mBio* 8.1, e02253–16. URL: https://doi.org/10.1128/mBio.02253-16.

Levine, P. and A Bashford (2010). "The Oxford Handbook of the History of Eugenics". In: ed. by A Bashford and P. Levine. Oxford University Press. Chap. Introduction: Eugenics and the Modern World.

Levins, R. (1966). "The strategy of model building in population biology". In: *Am. Sci.* 54.4, pp. 421–431. URL: https://www.jstor.org/stable/27836590.

Levins, R. and R. Lewontin (1985). *The dialectical biologist*. Delhi, India: Aakar Books.

Lewontin, R. C. (1970). "The units of selection". In: *Annu. Rev. Ecol. Evol. Syst.* 1.1, pp. 1–18. URL: https://doi.org/10.1146/annurev.es.01.110170.000245.

— (1974). *The Genetic Basis of Evolutionary Change*. New York, New York: Columbia University Press.

— (1985). "The dialectical biologist". In: Delhi, India: Aakar Books. Chap. The organism as the subject and object of evolution.

Li, J., M. Schumer, and C. Bank (2022). "Imbalanced segregation of recombinant haplotypes in hybrid populations reveals inter-and intrachromosomal Dobzhansky-Muller incompatibilities". In: *PLOS Genet.* 18.3, e1010120. URL: https://doi.org/10.1371/journal.pgen.1010120.

Liu, C. et al. (2017). "A generalised individual-based algorithm for modelling the evolution of quantitative herbicide resistance in arable weed populations". In: *Pest Manag. Sci.* 73.2, pp. 462–474. URL: https://doi.org/10.1002/ps.4317.

Lloyd, G. (1993). *The Man of Reason: 'Male' and 'Female' in Western Philosophy*. 2nd. Routledge.

Loga, K. von and M. Gerlinger (2017). "Cancer (r)evolution". In: *Nat. Ecol. Evol.* 1.8, pp. 1051–1052. URL: https://doi.org/10.1038/s41559-017-0252-1.

Losos, J. B. (2011). "Convergence, adaptation, and constraint". In: *Evolution* 65.7, pp. 1827–1840. URL: https://doi.org/10.1111/j.1558-5646.2011.01289.x.

Lunzer, M., G. B. Golding, and A. M. Dean (2010). "Pervasive cryptic epistasis in molecular evolution". In: *PLOS Genet.* 6, e1001162. URL: https://doi.org/10.1371/journal.pgen.1001162.

MacKenzie, D. (1976). "Eugenics in Britain". In: *Soc. Stud. Sci.* 6.3-4, pp. 499–532. URL: https://doi.org/10.1177/030631277600600310.

— (1979). "Karl Pearson and the professional middle class". In: *Ann. Sci.* 36.2, pp. 125–143. URL: https://doi.org/10.1080/00033797900200441.

Magnello, M. E. (1999). "The non-correlation of biometrics and eugenics: Rival forms of laboratory work in Karl Pearson's career at University College London, part 2". In: *Hist. Sci.* 37.2, pp. 123–150. URL: https://doi.org/10.1177/007327539903700201.

Marjoram, P., A. Zubair, and S. V. Nuzhdin (2014). "Post-GWAS: where next? More samples, more SNPs or more biology?" In: *Heredity* 112.1, pp. 79–88. URL: https://doi.org/10.1038/hdy.2013.52.

Martin, G. and T. Lenormand (2015). "The fitness effect of mutations across environments: Fisher's geometrical model with multiple optima". In: *Evolution* 69, pp. 1433–1447. URL: https://doi.org/10.1111/evo.12671.

Martin-Roy, R. et al. (2021). "Differences in thermal tolerance between parental species could fuel thermal adaptation in hybrid wood ants". In: *Am. Nat.* 198.2, pp. 278–294. URL: https://doi.org/10.1086/715012.

Matuszewski, S. et al. (2016). "A statistical guide to the design of deep mutational scanning experiments". In: *Genetics* 204.1, pp. 77–87. URL: https://doi.org/10.1534/genetics.116.190462.

Mayr, E. (1982). *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. 1st. Belknap Press of Harvard University Press.

Mazumdar, P. M. H. (2002). "'Reform' eugenics and the decline of Mendelism". In: *Trends Genet.* 18.1, pp. 48–52. URL: https://doi.org/10.1016/S0168-9525(01)02551-3.

Melnyk, A. H., A. Wong, and R. Kassen (2015). "The fitness costs of antibiotic resistance mutations". In: *Evol. Appl.* 8.3, pp. 273–283.

Miko, I. (2008). "Epistasis: Gene Interaction and Phenotype Effects". In: *Nature Education* 1.1, p. 197. URL: https://doi.org/10.1098/rstb.1889.0007.

Monroe, J. G. et al. (2016). "Adaptation to warmer climates by parallel functional evolution of CBF genes in *Arabidopsis thaliana*". In: *Mol. Ecol.* 25.15, pp. 3632–3644. URL: https://doi.org/10.1111/mec.13711.

Moore, J. (2007). "R.A. Fisher: a faith fit for eugenics". In: *Stud. Hist. Philos. Sci. C* 38.1, pp. 110–135. URL: https://doi.org/10.1016/j.shpsc.2006.12.007.

Muller, H. J. (1942). "Isolating mechanisms, evolution, and temperature". In: *Biol Symp.* Vol. 6, pp. 71–125.

Murray, C. J. L. et al. (2022). "Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis". In: *Lancet* 399, pp. 629–655. URL: https://doi.org/10.1016/S0140-6736(21)02724-0.

Mustonen, V. and M. Lässig (2009). "From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation". In: *Trends Genet.* 25.3, pp. 111–119. URL: https://doi.org/10.1016/j.tig.2009.01.002.

Nasr, S. B. et al. (2019). "Role of TILS in predicting evolution of metastatic colorectal cancer". In: *Ann. Oncol.* 30, p. iv85. URL: https://doi.org/10.1093/annonc/mdz155.307.

Nghe, P. et al. (2020). "Predicting evolution using regulatory architecture". In: *Annu. Rev. Biophys.* 49, pp. 181–197. URL: https://doi.org/10.1146/annurev-biophys-070317-032939.

Nichols, R. J. et al. (2011). "Phenotypic landscape of a bacterial cell". In: *Cell* 144, pp. 143–156. URL: https://doi.org/10.1016/j.cell.2010.11.052.

Norton, B. (1983). "Dimensions of Darwinism : themes and counterthemes in twentieth century evolutionary theory". In: ed. by M. Grene. Cambridge, UK: Cambridge University Press. Chap. Fisher's entrance into evolutionary science: the role of eugenics. URL: https://archive.org/details/dimensionsofdarw0000gren/.

Ochman, H. and N. A. Moran (2001). "Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis". In: *Science* 292.5519, pp. 1096–1099. URL: https://doi.org/10.1126/science.1058543.

Ogbunugafor, C. B. (2022). "The mutation effect reaction norm (mu-rn) highlights environmentally dependent mutation effects and epistatic interactions". In: *Evolution* 76.s1, pp. 37–48. URL: https://doi.org/10.1111/evo.14428.

Okasha, S. (2008). "Fisher's fundamental theorem of natural selection – a philosophical analysis". In: *Br. J. Philos. Sci.* URL: https://doi.org/10.1093/bjps/axn010.

Ottenburghs, J. (2018). "Exploring the hybrid speciation continuum in birds". In: *Ecol. Evol.* 8.24, pp. 13027–13034. URL: https://doi.org/10.1002/ece3.4558.

Otto, S. P. (2021). "Selective interference and the evolution of sex". In: *J. Hered.* 112.1, pp. 9–18. URL: https://doi.org/10.1093/jhered/esaa026.

Otto, S. P. and T. Day (2007). *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton, New Jersey: Princeton University Press.

Otto, S. P. and M. C. Whitlock (1997). "The probability of fixation in populations of changing size". In: *Genetics* 146.2, pp. 723–733. URL: https://doi.org/10.1534/genetics.114.172890.

Palmer, A. C. and R. Kishony (2013). "Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance". In: *Nat. Rev. Genet.* 14.4, pp. 243–248. URL: https://doi.org/10.1038/nrg3351.

Park, Y., B. P. H. Metzger, and J. W. Thornton (2022). "Epistatic drift causes gradual decay of predictability in protein evolution". In: *Science* 376, pp. 823–830. URL: https://doi.org/10.1126/science.abn6895.

Partridge, D. (2018). "Darwin's two theories, 1844 and 1859". In: *J. Hist. Biol.* 51, pp. 563–592. URL: https://doi.org/10.1007/s10739-018-9509-z.

Paul, D. B. and J. Moore (2010). "The Oxford Handbook of the History of Eugenics". In: ed. by A Bashford and P. Levine. Oxford University Press. Chap. The Darwinian Context: Evolution and Inheritance.

Paun, V. I. et al. (2021). "First report on antibiotic resistance and antimicrobial activity of bacterial isolates from 13,000-year old cave ice core". In: *Sci. Rep.* 11.1, pp. 1–15. URL: https://doi.org/10.1038/s41598-020-79754-5.

Pavlidis, P. and N. Alachiotis (2017). "A survey of methods and tools to detect recent and strong positive selection". In: *J. Biol. Res. (Thessalon)* 24, pp. 1–17. URL: https://doi.org/10.1186/s40709-017-0064-0.

Pearl, R. (1917). "The selection problem". In: *Am. Nat.* 51.602, pp. 65–91. URL: https://www.journals.uchicago.edu/doi/pdf/10.1086/279588.

Pearson, K. (1897). "The chances of death and other studies in evolution". In: 37 Bedford St., London: Edward Arnold. Chap. Reproductive selection. URL: https://archive.org/details/b21698223_0001/.

Pinheiro, F. et al. (2021). "Metabolic fitness landscapes predict the evolution of antibiotic resistance". In: *Nat. Ecol. Evol.* 5, pp. 677–687. URL: https://doi.org/10.1038/s41559-021-01397-0.

Plutynski, A. (2006). "What was Fisher's fundamental theorem of natural selection and what was it for?" In: *Stud. Hist. Philos. Sci. C* 37.1, pp. 59–82. URL: https://doi.org/10.1016/j.shpsc.2005.12.004.

Popov, I. (2018). *Orthogenesis versus Darwinism*. 1st. Cham, Switzerland: Springer International Publishing AG.

Popper, K. R. (1962). *Conjectures and refutations: the growth of scientific knowledge*. 1st. New York, New York: Basic Books. URL: https://archive.org/details/conjecturesrefut0000popp/.

Portinha, B. et al. (2022). "Whole-genome analysis of multiple wood ant population pairs supports similar speciation histories, but different degrees of gene flow, across their European ranges". In: *Mol. Ecol.* 31.12, pp. 3416–3431. URL: https://doi.org/10.1111/mec.16481.

Price, G. R. (1972). "Fisher's 'fundamental theorem' made clear". In: *Ann. Hum. Genet.* 36.2, pp. 129–140.

Provine, W. B. (1986). *Sewall Wright and Evolutionary Biology*. 1st. Chicago: University of Chicago Press.

— (2001). *The Origins of Theoretical Population Genetics: With a New Afterword*. 2nd. Chicago: University of Chicago Press.

Ramiro, R. S., H. Costa, and I. Gordo (2016). "Macrophage adaptation leads to parallel evolution of genetically diverse *Escherichia coli* small-colony variants with increased fitness *in vivo* and antibiotic collateral sensitivity". In: *Evol. Appl.* 9.8, pp. 994–1004. URL: https://doi.org/10.1111/eva.12397.

Renaut, S., G. L. Owens, and L. H. Rieseberg (2014). "Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers". In: *Mol. Ecol.* 23.2, pp. 311–324. URL: https://doi.org/10.1111/mec.12600.

Richmond, M. L. (2001). "Women in the early history of genetics: William Bateson and the Newnham College Mendelians, 1900-1910". In: *Isis* 92.1, pp. 55–90. URL: https://doi.org/10.1086/385040.

Rieder, J. et al. (in press). "Metagenomics and metabarcoding experimental choices and their impact on microbial community characterization in freshwater recirculating aquaculture systems". In: *Environ. Microbiol.* URL: https://doi.org/10.1101/2022.07.05.498813.

Rocha, E. P. C. (2018). "Neutral theory, microbial practice: challenges in bacterial population genetics". In: *Mol. Biol. Evol.* 35.6, pp. 1338–1347. URL: https://doi.org/10.1093/molbev/msy078.

Rodríguez-Verdugo, A., B. S. Gaut, and O. Tenaillon (2013). "Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress". In: *BMC Evol. Biol.* 13.1, pp. 1–11. URL: https://doi.org/10.1186/1471-2148-13-50.

Roll-Hansen, N. (2010). "The Oxford Handbook of the History of Eugenics". In: ed. by A Bashford and P. Levine. Oxford University Press. Chap. Eugenics and the Science of Genetics.

Rouse, J. (2004). "The Feminist Standpoint Theory Reader: intellectual and political controversies". In: ed. by S. Harding. New York, New York: Routledge. Chap. Feminism and the social construction of scientific knowledge.

Rull, V. (2022). "On predictions and laws in biological evolution: Is biology able to formulate general laws and develop inductive predictions as in physics or chemistry?" In: *EMBO Rep.* 23.4, e54392. URL: https://doi.org/10.15252/embr.202154392.

Russell, C. A. and M. D. de Jong (2017). "Infectious disease management must be evolutionary". In: *Nat. Ecol. Evol.* 1.8, pp. 1053–1055. URL: https://doi.org/10.1038/s41559-017-0265-9.

Sailer, Z. R. and M. J. Harms (2017). "Molecular ensembles make evolution unpredictable". In: *PNAS* 114.45, pp. 11938–11943. URL: https://doi.org/10.1073/pnas.1711927114.

Salmon, W. C. (2006). *Four Decades of Scientific Explanation*. 3rd. University of Pittsburgh Press.

Santos-Lopez, A. et al. (2021). "The roles of history, chance, and natural selection in the evolution of antibiotic resistance". In: *eLife* 10, e70676. URL: https://doi.org/10.7554/eLife.70676.

Sarkar, S. (2017). "Haldane's *The Causes of Evolution* and the modern synthesis in evolutionary biology". In: *J. Genet.* 96, pp. 753–763. URL: https://doi.org/10.1007/s12041-017-0840-5.

Scott, T. A. et al. (2017). "Host-microbe co-metabolism dictates cancer drug efficacy in *C. elegans*". In: *Cell* 169.3, pp. 442–456. URL: https://doi.org/10.1016/j.cell.2017.03.040.

Scriven, M. (1962). "Explanations, predictions, and laws". In: *University of Minnesota Press* Minneapolis. URL: https://hdl.handle.net/11299/184631.

Semmel, B. (1958). "Karl Pearson: Socialist and Darwinist". In: *Br. J. Sociol.* 9.2, pp. 111–125. URL: https://doi.org/10.2307/587909.

Servedio, M. R. and M. A. F. Noor (2003). "The role of reinforcement in speciation: theory and data". In: *Annu. Rev. Ecol. Evol. Syst.* 34.1, pp. 339–364. URL: https://doi.org/10.1146/annurev.ecolsys.34.011802.132412.

Shapin, S. and S. Schaffer (2011). *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. 2nd. Princeton University Press.

Shaw, R. G. (2019). "From the past to the future: considering the value and limits of evolutionary prediction". In: *Am. Nat.* 193.1, pp. 1–10. URL: https://doi.org/10.1086/700565.

Shefferson, R. P., R. Mizuta, and M. J. Hutchings (2017). "Predicting evolution in response to climate change: the example of sprouting probability in three dormancy-prone orchid species". In: *R. Soc. Open Sci.* 4.1, p. 160647. URL: https://doi.org/10.1098/rsos.160647.

Shefferson, R. P. et al. (2018). "The evolutionary impacts of conservation actions". In: *Popul. Ecol.* 60, pp. 49–59. URL: https://doi.org/10.1007/s10144-018-0614-9.

Shen, C. et al. (2020). "Dynamics of mcr-1 prevalence and mcr-1-positive *Escherichia coli* after the cessation of colistin use as a feed additive for animals in China: a prospective cross-sectional and whole genome sequencing-based molecular epidemiological study". In: *Lancet Microbe* 1.1, e34–e43. URL: https://doi.org/10.1016/S2666-5247(20)30005-7.

Shmueli, G. (2010). "To explain or to predict?" In: *Stat. Sci.* 25.3, pp. 289–310. URL: https://www.jstor.org/stable/41058949.

Simões, P. et al. (2008). "How repeatable is adaptive evolution? The role of geographical origin and founder effects in laboratory adaptation". In: *Evolution* 62.8, pp. 1817–1829. URL: https://doi.org/10.1111/j.1558-5646.2008.00423.x.

Sniegowski, P. D. and P. J. Gerrish (2010). "Beneficial mutations and the dynamics of adaptation in asexual populations". In: *Phil. Trans. R. Soc. B* 365.1544, pp. 1255–1263. URL: https://doi.org/10.1098/rstb.2009.0290.

Sommer, M. O. A. et al. (2017). "Prediction of antibiotic resistance: time for a new preclinical paradigm?" In: *Nat. Rev. Microbiol.* 15.11, pp. 689–696. URL: https://doi.org/10.1038/nrmicro.2017.75.

Spencer, H. G. and D. B. Paul (1998). "The failure of a scientific critique: David Heron, Karl Pearson and Mendelian eugenics". In: *Br. J. Hist. Sci.* 31.4, pp. 441–452. URL: https://doi.org/10.1017/S0007087498003392.

Starr, T. N. et al. (2018). "Pervasive contingency and entrenchment in a billion years of Hsp90 evolution". In: *PNAS* 115, pp. 4453–4458. URL: https://doi.org/10.1073/pnas.1718133115.

Stigler, S. M. (2010). "Darwin, Galton and the statistical enlightenment". In: *J. R. Stat. Soc. Ser. A Stat. Soc.* 173.3, pp. 469–482. URL: https://doi.org/10.1111/j.1467-985X.2010.00643.x.

Sturtevant, A. H. (2001). "Reminiscences of TH Morgan". In: *Genetics* 159.1, pp. 1–5. URL: https://doi.org/10.1093/genetics/159.1.1.

Sundqvist, M. et al. (2010). "Little evidence for reversibility of trimethoprim resistance after a drastic reduction in trimethoprim use". In: *J. Antimicrob. Chemother.* 65.2, pp. 350–360. URL: https://doi.org/10.1093/jac/dkp387.

Szendro, I. G. et al. (2013). "Quantitative analyses of empirical fitness landscapes". In: *J. Stat. Mech. Theory Exp.* 2013, P01005. URL: https://doi.org/10.1088/1742-5468/2013/01/P01005.

Tachida, H. and M. Iizuka (1991). "Fixation probability in spatially changing environments". In: *Genet. Res.* 58.3, pp. 243–251. URL: https://doi.org/10.1017/S0016672300029992.

Tang, K. L. et al. (2017). "WHO Guidelines on Use of Medically Important Antimicrobials in Food-Producing Animals". In: chap. Restriction in the use of antibiotics in food animals and antibiotic resistance in food animals and humans – a systematic review and meta-analysis (University of Calgary, Canada). URL: https://www.ncbi.nlm.nih.gov/books/NBK487966/.

Taylor, S. A. and E. L. Larson (2019). "Insights from genomes into the evolutionary importance and prevalence of hybridization in nature". In: *Nat. Ecol. Evol.* 3.2, pp. 170–177. URL: https://doi.org/10.1038/s41559-018-0777-y.

Tegze, B. et al. (2012). "Parallel evolution under chemotherapy pressure in 29 breast cancer cell lines results in dissimilar mechanisms of resistance". In: *PLOS One* 7.2, e30804. URL: https://doi.org/10.1371/journal.pone.0030804.

Tenaillon, O. et al. (2012). "The molecular diversity of adaptive convergence". In: *Science* 335, pp. 457–461. URL: https://doi.org/10.1126/science.1212986.

The American Eugenics Society (2015). *Members, Officers, and Directors Activities Database*. URL: https://archive.org/details/AMERICANEUGENICSSOCIETYMEMBERS.

The Royal Society (2023). *Forthcoming issues in Philosophical Transactions of the Royal Society B: Biological Sciences*. Accessed on February 05, 2023. URL: https://royalsocietypublishing.org/rstb/forthcoming-issues.

Thompson, C. L. et al. (2021). "Back to the future: Reintegrating biology to understand how past eco-evolutionary change can predict future outcomes". In: *Integr. Comp. Biol.* 61.6, pp. 2218–2232. URL: https://doi.org/10.1093/icb/icab068.

Travisano, M. et al. (1995). "Experimental tests of the roles of adaptation, chance, and history in evolution". In: *Science* 267.5194, pp. 87–90. URL: https://doi.org/10.1126/science.7809610.

Trindade, S., A. Sousa, and I. Gordo (2012). "Antibiotic resistance and stress in the light of Fisher's model". In: *Evolution* 66, pp. 3815–3824. URL: https://doi.org/10.1111/j.1558-5646.2012.01722.x.

Turda, M. (2010). "The Oxford Handbook of the History of Eugenics". In: ed. by A Bashford and P. Levine. Oxford University Press. Chap. Race, Science, and Eugenics in the Twentieth Century.

Ulett, M. A. (2014). "Making the case for orthogenesis: the popularization of definitely directed evolution (1890–1926)". In: *Stud. Hist. Philos. Sci. C* 45, pp. 124–132. URL: https://doi.org/10.1016/j.shpsc.2013.11.009.

Valentine, D. (2021). "Technologies of Reproduction: Race, Disability, and Neoliberal Eugenics". In: *J. Phil. Disability*. URL: https://doi.org/10.5840/jpd20217122.

van Valen, L. (1973). "A new evolutionary law". In: *Evol. Theory* 1, pp. 1–30.

van Wyhe, J. (2007). "Mind the gap: Did Darwin avoid publishing his theory for many years?" In: *Notes Rec. R. Soc. Lond.* 61, pp. 177–205. URL: https://doi.org/10.1098/rsnr.2006.0171.

Vatsiou, A. I., E. Bazin, and O. E. Gaggiotti (2016). "Detection of selective sweeps in structured populations: a comparison of recent methods". In: *Mol. Ecol.* 25.1, pp. 89–103. URL: https://doi.org/10.1111/mec.13360.

Wahl, L. M. and A. D. Zhu (2015). "Survival probability of beneficial mutations in bacterial batch culture". In: *Genetics* 200.1, pp. 309–320. URL: https://doi.org/10.1534/genetics.114.172890.

Wallace, A. R. (1889). *Darwinism*. URL: https://web.archive.org/web/20070313182223/http://www.wku.edu/~smithch/wallace/S724CH15.htm.

Weldon, W. F. (1893). "On certain correlated variations in *Carcinus moenas*". In: *Proc. R. Soc. Lond.* 45, pp. 209–218. URL: https://doi.org/10.1098/rspl.1893.0024.

Wheldale, M. (1907). "The inheritance of flower colour in *Antirrhinum majus*". In: *Phil. Trans. R. Soc. B* 79.532, pp. 288–305. URL: https://doi.org/10.1098/rspb.1907.0020.

Whitlock, M. C. et al. (1995). "Multiple fitness peaks and epistasis". In: *Annu. Rev. Ecol. Syst.* 26, pp. 601–629. URL: https://doi.org/10.1146/annurev.es.26.110195.003125.

Wikipedia (2022). *The Time Machine*. Accessed on Semptember 24, 2022. URL: https://en.wikipedia.org/wiki/The_Time_Machine.

Wortel, M. T. et al. (2022). "Towards evolutionary predictions: Current promises and challenges". In: *Evol. Appl.* URL: https://doi.org/10.1111/eva.13513.

Wright, S. (1931). "Evolution in Mendelian populations". In: *Genetics* 16.2, pp. 97–159. URL: https://doi.org/10.1093/genetics/16.2.97.

— (1980). "Genic and organismic selection". In: *Evolution* 34.5, pp. 825–843. URL: https://doi.org/10.2307/2407990.

Chapter 2

Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (& sex chromosomes)

It is often said that conflict is invitation to transform the system. I also believe that conflict is an opportunity to transform our selves and others.
– Kai Cheng Thom

This chapter was published in *Molecular Ecology* in 2018 as part of a special issue on sex chromosomes and speciation. The student's contributions to the work are as follows: Claudia Bank and Jonna Kulmuni designed the research; Claudia Bank developed and coded the 2-locus model; Alexandre Blanckaert extended the model and analyzed it for >2 loci; The student performed the simulations and fitted the model to the data; All authors interpreted the results and wrote the manuscript; The student helped revise the manuscript according to reviewers' comments.

Received: 30 September 2017 | Revised: 19 December 2017 | Accepted: 22 December 2017

DOI: 10.1111/mec.14482

SPECIAL ISSUE: SEX CHROMOSOMES AND SPECIATION

WILEY MOLECULAR ECOLOGY

Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (and sex chromosomes)

Ana-Hermina Ghenu[1]* | Alexandre Blanckaert[1]* | Roger K. Butlin[2,3] | Jonna Kulmuni[2,4] | Claudia Bank[1,5]

[1]Instituto Gulbenkian de Ciência, Oeiras, Portugal

[2]Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

[3]Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden

[4]Centre of Excellence in Biological Interactions, Department of Biosciences, University of Helsinki, Helsinki, Finland

[5]Kavli Institute for Theoretical Physics, University of California at Santa Barbara, Santa Barbara, CA, USA

Correspondence
Claudia Bank, Instituto Gulbenkian de Ciência, Oeiras, Portugal.
Email: evoldynamics@gmail.com

Funding information
Fundação Calouste Gulbenkian; National Science Foundation, Grant/Award Number: PHY-1125915; Human Frontier Science Program; Finnish Cultural Foundation; Academy of Finland, Grant/Award Number: 252411

Abstract

In many diploid species, the sex chromosomes play a special role in mediating reproductive isolation. In haplodiploids, where females are diploid and males haploid, the whole genome behaves similarly to the X/Z chromosomes of diploids. Therefore, haplodiploid systems can serve as a model for the role of sex chromosomes in speciation and hybridization. A previously described population of Finnish *Formica* wood ants displays genome-wide signs of ploidally and sexually antagonistic selection resulting from hybridization. Here, hybrid females have increased survivorship but hybrid males are inviable. To understand how the unusual hybrid population may be maintained, we developed a mathematical model with hybrid incompatibility, female heterozygote advantage, recombination and assortative mating. The rugged fitness landscape resulting from the co-occurrence of heterozygote advantage and hybrid incompatibility results in a sexual conflict in haplodiploids, which is caused by the ploidy difference. Thus, whereas heterozygote advantage always promotes long-term polymorphism in diploids, we find various outcomes in haplodiploids in which the population stabilizes either in favour of males, females or via maximizing the number of introgressed individuals. We discuss these outcomes with respect to the potential long-term fate of the Finnish wood ant population and provide approximations for the extension of the model to multiple incompatibilities. Moreover, we highlight the general implications of our results for speciation and hybridization in haplodiploids versus diploids and how the described fitness relationships could contribute to the outstanding role of sex chromosomes as hotspots of sexual antagonism and genes involved in speciation.

KEYWORDS
epistasis, fitness landscape, haplodiploidy, heterosis, hybridization, speciation

1 | INTRODUCTION

Haplodiploids are an emerging system for speciation genetics (Knegt et al., 2017; Koevoets & Beukeboom, 2009; Kulmuni & Pamilo, 2014; Lohse & Ross, 2015). Although ≈ 20% of animal species are haplodiploid (comprising most *Hymenopterans*, some arthropods, thrips and *Hemipterans*, and several clades of beetles and mites; Crozier & Pamilo, 1996; Evans, Shearman, & Oldroyd, 2004; de la Filia, Bain, & Ross, 2015), little evolutionary theory has been developed specifically for speciation in haplodiploids (Koevoets & Beukeboom, 2009). Under haplodiploidy with arrhenotoky (hereafter simply haplodiploidy; Suomalainen, Saura, & Lokki, 1987), males develop from

*These authors contributed equally.

the mother's unfertilized eggs and are haploid, whereas eggs fertilized by fathers result in diploid females. As this mode of inheritance is, from a theoretical viewpoint, similar to that of the X/Z chromosome, most work on speciation of haplodiploids draws on the rich literature of sex chromosome evolution (Jablonka & Lamb, 1991; Johnson & Lachance, 2012; Lohse & Ross, 2015; Presgraves, 2008). An important similarity between haplodiploids and X/Z chromosomes is that recessive mutations in the haploid sex are exposed to selection, but they are masked in diploids. This is expected to lead to faster evolution in the sex chromosomes (Charlesworth, Coyne, & Barton, 1987) that may partly underlie the large-X effect (Presgraves, 2008). The large-X effect refers to the observation that the sex chromosomes seem to play a special role in speciation by acting as the strongest barrier for gene flow between hybridizing lineages across different species (Höllinger & Hermisson, 2017). Similarly, haplodiploid species have been suggested to acquire reproductive isolation earlier and speciate faster than diploid species (Lima, 2014; Lohse & Ross, 2015). Although the factors influencing haplodiploid and X/Z chromosome evolution are not expected to be exactly the same (e.g., movement of sexually antagonistic genes to the sex chromosomes, dosage compensation between the sex chromosomes and autosomes, and turnover of sex chromosomes cannot occur in haplodiploids; Abbott, Nordén, & Hansson, 2017), by studying haplodiploid models we can both improve our understanding of how speciation happens in the large subgroup of the animal kingdom that is haplodiploid, and gain new insights into the role of X/Z chromosomes in speciation for diploid species.

Recent studies have shown that hybridization and resulting gene flow between diverging populations may be important players in the speciation process since signs of hybridization and introgression are being observed ubiquitously in natural populations (Dieckmann & Doebeli, 1999; Mallet, 2005; Schluter, 2009; Schluter & Conte, 2009; Seehausen et al., 2014). When a hybrid population is formed, various selective forces may act simultaneously to either increase or decrease hybrid fitness, which dictate the fate of the population. One commonly documented finding is hybrid incompatibility (Chen, Zhiguo, & Lin, 2016; Fraïsse, Elderfield, & Welch, 2014; Presgraves, 2008), where combinations of alleles at different loci interact to confer poor fitness when combined in a hybrid individual (Bateson, 1909; Dobzhansky, 1936; Muller, 1942; Orr, 1995). In a hybrid population, the existence of hybrid incompatibility reduces the mean population fitness. This deficit can be resolved either through reinforcement (evolution of increased premating isolation to avoid production of unfit hybrids; Servedio & Noor, 2003) or by purging (demographic swamping leading to extinction of one of the local populations/species or reinstatement of the ancestral allele combinations; Wolf, Takebayashi, & Rieseberg, 2001). On the other hand, hybridization can transfer adaptive genetic variation from one lineage to another (Heliconius Genome Consortium, 2012; Song et al., 2011; Whitney, Randell, & Rieseberg, 2010) and may result in overall heterosis (also known as hybrid vigour): a higher fitness of hybrids as compared to their parents (Bernardes, Stelkens, & Greig, 2017; Chen, 2013; Schwarz, Matta, Shakir-Botteri, & McPheron, 2005). Heterosis can stabilize polymorphisms by conferring a fitness advantage to hybrids and thereby favour the maintenance of hybridization either through the improved exploitation of novel ecological niches or the masking of recessive deleterious mutations. Therefore, hybrid incompatibility acts to avert ongoing hybridization while heterosis favours the maintenance of hybrids.

One example of the simultaneous action of hybridization-averse and hybridization-favouring forces is found in a hybrid population of *Formica polyctena* and *F. aquilonia* wood ants in Finland (Beresford et al., 2017; Kulmuni & Pamilo, 2014; Kulmuni, Seifert, & Pamilo, 2010). Here, it has been reported that hybrid (haploid) males do not survive to adulthood, whereas (diploid) females have higher survivorship when they carry many introgressed alleles as heterozygotes (i.e., heterozygous for alleles originating from one of the parental species in a genomic background mostly indicative of the other parental species). Thus, a combination of hybrid incompatibility and heterosis seems to dictate the dynamics of the population in a ploidy-specific manner: hybrid haploid males suffer a fitness cost while diploid hybrid females can have a selective advantage over parental ones. Here, the differences in ploidy create an apparent sexual conflict (sensu Arnqvist & Rowe, 2005) between haploid males and diploid females, because their fitness landscapes (i.e., the complex relationship between genotypes and fitness created via hybrid incompatibility and heterozygote advantage) are different. This conflict is absent if the same rugged fitness landscape occurs in diploid autosomes.

When both hybridization-averse and hybridization-favouring forces are acting, the long-term resolution of a hybridizing population is difficult to foresee: will hybridization eventually result in either complete speciation or extinction of one of the populations involved? Alternatively, can it represent an equilibrium maintained stably on an evolutionary timescale? Furthermore, will the probability of these outcomes depend on ploidy? In other words, is one of these outcomes more probable when interacting genes are found on a "haplodiploid" X/Z chromosome than when they exist on a "diploid" autosome?

We here develop and analyse a population-genetic model of an isolated hybrid population in which both hybridization-averse and hybridization-favouring forces are acting, and we study the evolutionary outcomes in both haplodiploid and (fully) diploid genetic systems. The rich dynamics of the haplodiploid model can result in four possible evolutionary stable states depending on the strength of heterozygote advantage *versus* hybrid incompatibility, the strength of recombination, and the degree of assortative mating. This includes a case of symmetric coexistence (where all diversity is maintained) in which both alleles can be maintained despite the segregating hybrid incompatibility, and in which long-term hybridization is favoured. We find that the dynamics differ between haplodiploid and diploid systems and that, unlike in previous models of sexual conflict in haplodiploid populations (Albert & Otto, 2005; Kraaijeveld, 2009), the conflict is not necessarily resolved in favour of the females. Indeed, a compromise may be reached at which the average fitness of females is decreased to rescue part of the fitness of males. Moreover, evaluation of the model using the data from the natural hybrid

population suggests that, under the assumption of an equilibrium, the Finnish ant population may represent an example of compromise between male costs and female benefits through asymmetric coexistence. We discuss our findings with respect to the long-term effects of hybridization, the potential for speciation in haplodiploid versus diploid species, and with respect to their relevance for X- or Z-linked alleles in diploid individuals.

2 | MATERIALS AND METHODS

2.1 | The model

We model an isolated haplodiploid or diploid hybrid population with individuals from two founder populations P_+ and P_-. Note that throughout the manuscript, we preferentially refer to (sub)populations rather than species; in those instances in which we use the term "species" it is in order to emphasize that the two populations have diverged sufficiently for (potentially strong) hybrid incompatibility to exist. We assume discrete generations and consider two loci, **A** and **B**. Each locus has two alleles, the "+" allele (A_+ or B_+) inherited from population P_+ and the "−" allele (A_- or B_-) inherited from population P_-. We refer to "hybrids" as individuals that carry two alleles from each of the two parental populations and cannot be assigned to either parental background. We refer to "introgressed" individuals as those genotypes for which three of the four alleles are from the same parental population; these genotypes are identical to those produced by hybridization followed by backcrossing. We ignore new or recurrent mutation and genetic drift. Thus, male and female populations are of effectively infinite size; selection modifies the relative abundance of the different haplotypes/genotypes but not the number of individuals (soft selection). The life cycle is as follows (Figure 1; see also Table 1 for a list of model parameters); consistent with the recursions defined below, we begin the life cycle at the adult stage:

TABLE 1 List of model parameters

Symbol	Parameter	Limits
σ, ω	Strength of heterozygote advantage, resulting in fitness $\omega = (1+\sigma)$ or $\omega^2 = (1+\sigma)^2$ of introgressed or double-heterozygous diploid hybrids, respectively.	$\omega - 1 = \sigma > 0$
γ_1, γ_2	Strength of fully recessive negative epistasis, resulting in fitness $(1-\gamma_1)$ for A_+B_- homozygous diploid hybrids and A_+B_- hybrid haploid males, and $(1-\gamma_2)$ for A_-B_+ homozygous diploid hybrids and A_-B_+ hybrid haploid males.	$0 \le \gamma_1, \gamma_2 \le 1$
ρ	Recombination rate between locus **A** and **B**.	$0 \le \rho \le 0.5$
α	Strength of assortment via genotype matching, where $\alpha = 0$ represents random mating, $\alpha > 0$ represents assortative mating among conspecifics, and $\alpha < 0$ represents assortative mating between heterospecifics.	$-1 \le \alpha \le 1$

1. mating, either randomly or via genotype matching with assortment strength α as detailed below;
2. recombination (in diploid individuals) at rate ρ;
3. viability (or survival) selection, where heterosis is modelled as a heterozygote advantage, σ, and hybrid incompatibility is modelled as a fully recessive negative epistasis, γ_1 and γ_2 (further details are provided below and in Figure 2).

2.2 | Viability selection

The fitness landscape described here (Figure 2) is inspired by the situation observed in Finnish *Formica* ants (Beresford et al., 2017; Kulmuni & Pamilo, 2014; Kulmuni et al., 2010). There, the authors discovered heterosis in the diploid females but recessive incompatibilities expressed in the haploid males. This creates a situation in which the

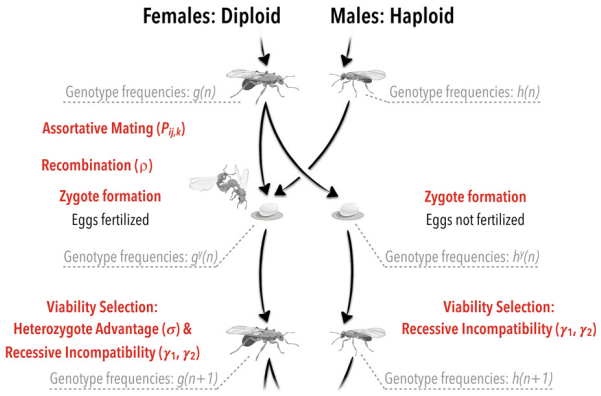
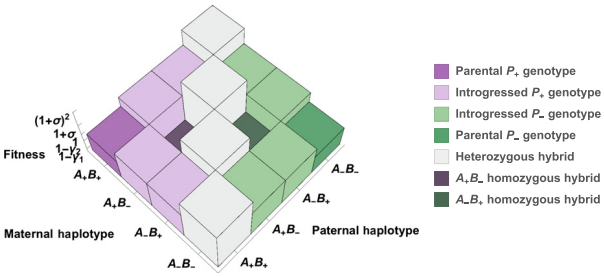
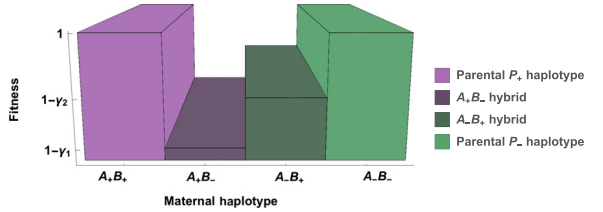


FIGURE 1 Illustration of the haplodiploid life cycle and its parameterization. Ant diagrams modified from "Ant Farm: Ask A Biologist activity for the classroom and home" (http://askabiologist.asu.edu/explore/ant_farm) ©Arizona Board of Regents/ASU Ask A Biologist, licensed under CC BY-SA 3.0 [Colour figure can be viewed at wileyonlinelibrary.com]



(a) Fitness Landscape for Diploid Individuals



(b) Fitness Landscape for Haploid Males

FIGURE 2 Three-dimensional fitness landscapes for the (a) diploid and (b) haploid genotypes. Panel a) corresponds to females in the haplodiploid model and all individuals in the diploid model. Individuals heterozygous at both loci (heterozygous hybrids) reside on a high-fitness ridge (in white), whereas individuals homozygous at both loci (homozygous hybrids) suffer from reduced fitness due to negative epistasis. Panel b) shows the fitness landscape for haploid individuals (i.e., males) in the haplodiploid model. This landscape is identical to a transect from Panel a) for genotypes homozygous at both loci

same alleles that are favoured in heterozygous females are selected against in hybrid haploid males and homozygous hybrid females. In the haplodiploid genetic system, males possess only one copy of each locus so they cannot be heterozygous and, therefore, cannot experience heterozygote advantage (Figure 2b). Therefore, the fitness landscape with heterozygote advantage and recessive hybrid incompatibility expresses itself as an apparent sexual conflict when sexes differ in ploidy, as in haplodiploids or for X/Z chromosomes.

In our model, selection for heterozygous individuals is multiplicative with respect to the number of heterozygous loci: introgressed individuals with one heterozygous locus have fitness $1 + \sigma$, whereas diploid hybrid individuals that are heterozygous at both loci have survivorship $(1 + \sigma)^2$ (Figure 2a). Finally, the recessive epistatic incompatibility parameter γ_1 acts on individuals homozygous or haploid for the A_+B_- haplotype, and γ_2 acts on individuals homozygous or haploid for the A_-B_+ haplotype (without loss of generality, we assume $\gamma_1 \geq \gamma_2$). Thus, epistasis in this model can be asymmetric, reflecting, for example, two Dobzhansky–Muller incompatibilities of different strength that have accumulated at a negligible recombination distance between the same chromosome pairs. Note that when $\gamma_1 = \gamma_2 = 1$, haploid hybrid males and homozygous hybrid zygotes are produced but do not survive to adulthood and that the classical case of a single Dobzhansky–Muller incompatibility is recovered when $\gamma_2 = 0$.

2.3 | Assortative mating

Prezygotic isolation via assortative mating is an important mechanism that could mediate the detrimental effects to the population caused by the co-occurrence of heterozygote advantage and epistasis modelled here. In the Finnish wood ant population that inspired our model (Kulmuni & Pamilo, 2014), almost all egg-laying queens collected had been inseminated by males of the same genetic group, indicating that prezygotic isolation barriers are likely operating to result in assortative mating. In this case, assortative mating could arise via choosiness of mating partners, via genotype-dependent development times, or via other postmating prezygotic mechanisms. We implemented assortment via genotype matching (reviewed in Kopp et al., 2017), where the proportion of matings depends on the genetic distance between two mating partners (and their respective frequencies in the population). We define the genetic distance between the genotypes of a mating pair as the average Hamming distance, that is the number of differences between two aligned sequences of characters, between all possible pairs of haplotypes with one partner from each sex. We use quadratic assortment (e.g., De Cara, Barton, & Kirkpatrick, 2008), which results in assortative mating without costs of choosiness but with sexual selection. The mating probability of a pair of male and female genotypes, $P_{ij,k}$, depends on the genetic distance between the two mates, the choosiness of the female and the abundance of the different haplotypes and genotypes as detailed below.

2.4 | Mathematical modelling and analysis

In a given generation n, the frequencies of the male and female adults are given by $h_k(n)$ and $g_{ij}(n)$, respectively, with i and k indicating the haplotype received maternally and j the one of paternal

origin. Without loss of generality, we assign index $i = 1$ to haplotype A_+B_+, index $i = 2$ to haplotype A_+B_-, $i = 3$ to haplotype A_-B_+ and $i = 4$ to A_-B_-. Below, we describe the modelled life cycle (illustrated in Figure S1), which determines how frequencies change from one generation to the next.

1. As detailed in Figure 1, the first step of the life cycle is the mating between two individuals. The mating probability between an *ij* female and a *k* male is given by:

$$P_{ij,k}(n) = \frac{\left(1 - \alpha\frac{d_{i,k} + d_{j,k}}{2}\right) g_{ij}(n) h_k(n)}{\sum_i \sum_j \sum_k \left(1 - \alpha\frac{d_{i,k} + d_{j,k}}{2}\right) g_{ij}(n) h_k(n)} \quad (1)$$

with $d_{i,k}$ the Hamming distance between two haplotypes. Note that for $\alpha = 0$, this simplifies to random mating and thus becomes equivalent to the dynamics described in Equation (S7).

2. The next step is the formation of the zygote. Recombination happens only in females. We denote the frequency of newly born females as $g^{\gamma}_{ik}(n+1)$.

$$\begin{cases} g^{\gamma}_{ik}(n+1) = \frac{1}{2}\sum_{j=1}^{4}\left(P_{ij,k}(n) + P_{\bar{i}\bar{j},k}(n)\right) - \frac{\rho}{2}\Delta_k(n) \text{ if } i \in \{1,4\} \\ g^{\gamma}_{ik}(n+1) = \frac{1}{2}\sum_{j=1}^{4}\left(P_{ij,k}(n) + P_{\bar{i}\bar{j},k}(n)\right) + \frac{\rho}{2}\Delta_k(n) \text{ if } i \in \{2,3\} \end{cases} \quad (2)$$

with $\Delta_k(n) = P_{14,k}(n) + P_{41,k}(n) - P_{23,k}(n) - P_{32,k}(n)$.

Males are composed from unfertilized female gametes, which have undergone recombination. The frequencies of newborn males are given by $h^{\gamma}_k(n)$:

$$\begin{cases} h^{\gamma}_k(n_\gamma) = \frac{1}{2}\sum_{j=1}^{4}\left(g_{kj}(n) + g_{\bar{k}\bar{j}}(n)\right) - \frac{\rho}{2}\tau(n) \text{ if } k \in \{1,4\} \\ h^{\gamma}_k(n_\gamma) = \frac{1}{2}\sum_{j=1}^{4}\left(g_{kj}(n) + g_{\bar{k}\bar{j}}(n)\right) + \frac{\rho}{2}\tau(n) \text{ if } k \in \{2,3\} \end{cases} \quad (3)$$

with $\tau(n) = g_{14}(n) + g_{41}(n) - g_{23}(n) - g_{32}(n)$.

3. Individuals of both sexes are under viability selection. The frequencies of male and female adults of the next generations are given by

$$h_k(n+1) = \frac{w^m_k h^{\gamma}_k(n)}{\sum_{k=1}^{4} w^m_k h^{\gamma}_k(n)} \quad (4)$$

with w^m_i the fitness of haplotype *i* in males and:

$$g_{ij}(n+1) = \frac{w^f_{ij} g^{\gamma}_{ij}(n)}{\sum_{i=1}^{4}\sum_{j=1}^{4} w^f_{ij} g^{\gamma}_{ij}(n)} \quad (5)$$

where w^f_{ij} denotes the fitness of the *ij* genotype. Note that there are no parental effects: $w^f_{ij} = w^f_{\bar{j}\bar{i}}$; we maintain the distinction only for modelling convenience.

The complete recursion for females is obtained by substituting $g^{\gamma}_{ij}(n)$ by its expression given in (2) in (5), and $P_{ij,k}(n)$ by (1). The complete recursion for males is given by substituting h^{γ}_i by its expression given in (3) in (4). For $\alpha = 0$, the detailed recursion is given in Equation (S7). Note that we use a different point of the life cycle (the gamete frequencies) as this is more easily tractable due to the reduced number of variables.

The diploid model can be obtained by applying Equations (2) and (5) to males as well, with the corresponding relevant substitutions.

For the analysis, we focus on the equilibrium of the system defined by:

$$\forall\{i,j,k\} \in \{1,2,3,4\}^3, g_{ij}(n+1) = g_{ij}(n) \text{ and } h_k(n+1) = h_k(n). \quad (6)$$

These equilibria can be obtained either by solving the system of equations presented above numerically, or by focusing on some of the known and potentially biologically relevant equilibria, like fixation of a given haplotype. The stability of the equilibria is then obtained by computing the eigenvalues of the Jacobian matrix at the focal equilibrium. If the absolute value of all eigenvalues is below 1, the equilibrium is locally stable. For a more detailed explanation, see Otto and Day (2007, chapter 7). We use this method to derive necessary and sufficient conditions for the existence and stability of the different evolutionary outcomes.

2.5 | Simulations

Derivations, simulations and data fitting were performed in MATHEMATICA (version 10.4.1.0; Wolfram Research, Inc., 2016). To enable complete reproducibility of the results, we provide an Online Supporting Information that documents all steps of the analysis as well as the code used for simulations and figures. Equilibrium genotype frequencies were obtained numerically when possible, or based on simulations until the differences between genotype frequencies of two consecutive generations were smaller than 10^{-8} (or stopped after 10^5 generations without convergence).

2.6 | Fitting the model to a natural ant population

To compare our model with data from the natural, hybridizing Finnish ant population, we estimated the different genotype frequencies of parental *F. polyctena*-like and *F. aquilonia*-like individuals from the data. Assuming that the natural population is at equilibrium, we fit the data (Table S2) to the model by calculating the sum of squared differences between the observed data and predicted equilibrium frequencies. Complete details of data estimation and model fitting are given in the Methods and Results in Supporting Information.

3 | RESULTS

In this section, we describe the dynamics of a hybrid population under our model, with a particular focus on quantifying the differences between the haplodiploid and the diploid model. Two parameter domains are of particular interest:

1. The case of free recombination and strong epistasis (i.e., large γ_1, γ_2) most likely resembles that of the natural ant hybrid population that inspired the model. Here, the hybrid incompatibility loci are located on different chromosomes, and epistasis is strong

enough to erase a large fraction of male zygotes during development.

2. The case of low recombination is most relevant for the effects of a fitness landscape with epistasis (i.e., a "rugged" landscape) in X or Z chromosomes. Here, epistasis could arise, for example, through interactions between regulatory regions and their respective genes.

3.1 | Evolutionary scenarios

Below, we describe four different types of evolutionary stable states (i.e., equilibrium scenarios) of the model, which represent long-term solutions to the opposing selective pressures of the hybridization-averse force of recessive negative epistasis and the hybridization-favouring heterozygote advantage. The population will attain these equilibria if no further pre- or postzygotic barriers or other functional mutations appear. Next, we provide various necessary and sufficient analytical conditions for these scenarios. Figure 3 illustrates the potential equilibria by means of phase diagrams.

3.2 | Exclusion

The *exclusion* scenario corresponds to the hybrid population becoming identical to one of the two parental populations, either P_{+} or P_{-}, and the other parental population being therefore excluded. It occurs when both alleles from one of the founder subpopulations are purged, leading to a monomorphic stable state of the population (Figure 3a). In this case, the initial frequency of $A_{+}B_{+}$ versus $A_{-}B_{-}$ individuals mainly determines the outcome (i.e., the population is swamped by the majority subpopulation). As a rule of thumb, this outcome is observed when recombination is frequent and when the hybridization-averse force of negative epistasis is strong as compared with the hybridization-favouring heterozygote advantage ($\gamma_1, \gamma_2 \gg \sigma$).

With regard to the apparent sexual/ploidy conflict in the haplodiploid model, exclusion can be interpreted as a victory of the haploid males because all polymorphism is lost and no low-fitness hybrid males are produced. Conversely, as all polymorphism is lost, diploid females "lose" in this case and neither high-fitness introgressed (i.e., those individuals carrying only one "foreign" allele) nor highest-fitness heterozygous hybrid females are produced. As discussed below, exclusion is never a possible outcome in the diploid model, in which there are no differences in ploidy between sexes.

3.3 | Single-locus polymorphism

A *single-locus polymorphism* occurs when one allele is purged from the population but the other locus remains polymorphic at equilibrium (Figure 3b). Because this is possible for either of the two loci, two such equilibria exist simultaneously, which are reached depending on the initial haplotype frequencies. This outcome is observed when recombination is frequent, epistasis is asymmetric ($\gamma_1 \neq \gamma_2$) and heterozygote advantage is small ($\gamma_1 \gg \sigma$). Like asymmetric

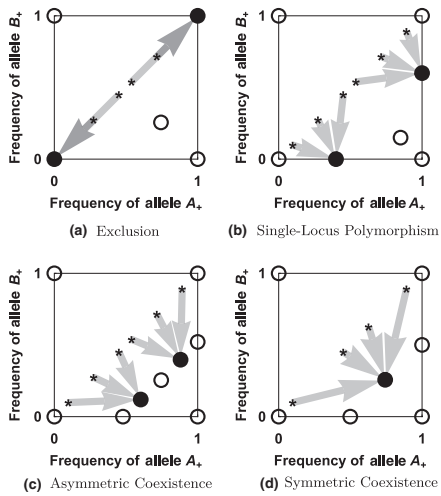


FIGURE 3 Phase-plane diagrams illustrating possible evolutionary scenarios in the haplodiploid model. The filled black dots show locally stable equilibria, and the empty dots show unstable ones. The grey arrows show the basin of attraction starting from secondary contact scenarios (black asterisks on the line at $p_{B_-} = p_{A_+}$). Panel (a) illustrates exclusion: there are two external locally stable equilibria, each corresponding to the fixation of a parental population haplotype. (Here, $\sigma = 0.02$, $\gamma_1 = 0.9$, $\gamma_2 = 0.11$, $\rho = 0.5$, and $\alpha = 0$.) Panel (b) represents a single-locus polymorphism. Only one locus is polymorphic, leading to the maintenance of the weaker of the two incompatibilities (the $A_{-}B_{+}$ interaction). (Here, $\sigma = 0.009$, $\gamma_1 = 0.11$, $\gamma_2 = 0.002$, $\rho = 0.5$, and $\alpha = 0$.) Panel (c) corresponds to asymmetric coexistence. Two internal equilibria are locally stable, with one allele close to fixation. This scenario minimizes the expression of the strongest interaction $A_{+}B_{-}$. (Here, $\sigma = 0.03$, $\gamma_1 = 0.11$, $\gamma_2 = 0.0013$, $\rho = 0.5$ and $\alpha = 0$.) Panel (d) shows symmetric coexistence. Frequencies of alleles A_{-} and B_{-} are symmetric around 0.5, with $p_{B_-} = 1 - p_{A_+}$. This scenario maximizes the formation of female heterozygous hybrids. (Here, $\sigma = 0.09$, $\gamma_1 = 0.3$, $\gamma_2 = 10^{-4}$, $\rho = 0.5$ and $\alpha = 0$.)

coexistence below, this case represents a compromise between the hybridization-averse and hybridization-favouring forces of negative epistasis and heterozygote advantage and is reached by maximizing the number of introgressed individuals of one founder subpopulation.

In the haplodiploid model, this scenario can be seen as a haploid-dominated compromise. As one locus is fixed, one epistatic interaction has disappeared and few low-fitness hybrid males are produced. In females, high-fitness introgressed female frequencies are maximized but, as one locus is fixed, the highest-fitness heterozygous hybrid female genotypes are no longer available.

The single-locus polymorphism is never stable in the diploid model, that is when the ploidy difference is removed from the

model. In a diploid population that resides transiently at single-locus polymorphism, a rare mutant at the second locus will always begin as heterozygote and therefore reap the advantage of being a heterozygote hybrid long before it suffers the epistatic cost of being a homozygote hybrid.

3.4 | Asymmetric coexistence

Asymmetric coexistence occurs when all four haplotypes remain in the population and the frequency of introgressed individuals of one founder subpopulation is maximized (Figure 3c). Because this can be achieved in two ways, two possible equilibria reside off the diagonal line $p_B = 1 - p_A$ (where p_A and p_B denote the allele frequencies of the "–" allele at the respective locus), and the initial contribution of different haplotypes determines which equilibrium will be attained. Like the single-locus polymorphism, this equilibrium represents a compromise between hybridization-averse and hybridization-favouring forces that is reached by maximizing the number of introgressed individuals. Our simulations demonstrate that this scenario is rarely present in haplodiploids, and it generally involves asymmetric epistasis and intermediate-strength heterozygote advantage.

In the haplodiploid model, asymmetric coexistence can be seen as a compromise that is dominated by the diploids. Unlike in the single-locus polymorphism scenario, both loci are polymorphic and some double-heterozygous hybrid females are produced. But, unlike the symmetric coexistence scenario described below, females are not victorious over males because such high-fitness hybrid females are produced only at low frequencies.

3.5 | Symmetric coexistence

Symmetric coexistence occurs when a locally stable equilibrium exists on the diagonal $p_B = 1 - p_A$, such that the number of heterozygous hybrids is maximized (Figure 3d). Our notion of "symmetric" refers to the total fraction of alleles from the P_+ and P_- founder populations segregating at equilibrium, which is equal in this case. Here, prolonged hybridization is a mutual best-case scenario for both populations. This equilibrium is most likely when recombination is weak or when the hybridization-favouring force of heterozygote advantage is strong as compared with the hybridization-averse negative epistasis ($\sigma \geq \gamma_1, \gamma_2$). In the haplodiploid model, symmetric coexistence represents a victory for the diploids, because they maximize their own fitness without regard to the production of unfit hybrid haploids.

The four evolutionary stable states described above usually result in either a single, globally stable equilibrium (in the case of symmetric coexistence) or a bistable system, in which two locally stable equilibria exist. In rare cases and close to bifurcation points, we observe cases of tristability, which are further described in Figure S2.

3.6 | Stability analysis of the model

Although the model dynamics are too complex to derive general analytical solutions, we were able to perform stability analyses for specific cases, which yield information about the general behaviour of the model. In the following, our use of ">" and "<" does not necessarily imply strict inequalities; we merely did not explicitly study the limiting cases. For ease of notation, we refer to heterozygote advantage in terms of ω below; recall that $\omega = 1 + \sigma$.

3.7 | Conditions for symmetric coexistence when epistasis is lethal

We begin by describing the equilibrium structure when epistasis is lethal, that is $\gamma_1 = \gamma_2 = 1$; this case may resemble that in the natural ant population, in which most hybrid males do not survive to reproduce. For the haplodiploid model, we obtain a full analytic solution of the identity, existence and stability of equilibria. Here, only two outcomes are possible: symmetric coexistence and exclusion (Figure 4a). As necessary and sufficient criterion for exclusion, we obtain

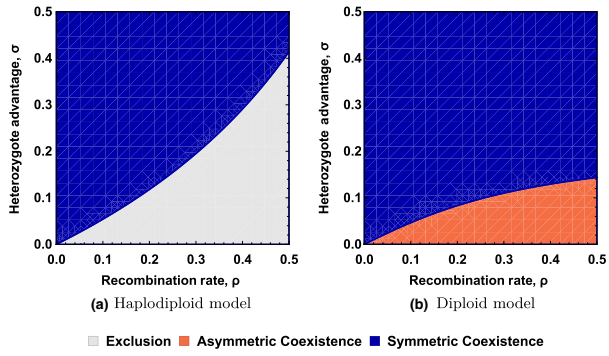


FIGURE 4 Symmetric coexistence can be locally stable if the heterozygote advantage, σ, is strong enough to compensate for recombination breaking up the parental haplotypes. Here we assume that epistasis is symmetric and lethal ($\gamma_1 = \gamma_2 = 1$). Panel (a) is an illustration of the condition for haplodiploids given in Equation (7) and panel (b) of Equation (8) for diploids [Colour figure can be viewed at wileyonlinelibrary.com]

$$\rho > \frac{\omega^2 - 1}{\omega^2}. \quad (7)$$

Thus, exclusion is only possible if heterozygote advantage is not too strong, and if recombination is breaking up gametes sufficiently often to significantly harm the haploid males.

For the diploid model, we can show that no boundary equilibrium is ever stable; asymmetric and symmetric coexistence are the only two possible outcomes. Although it was not possible to perform a stability analysis on the internal equilibria, we were able to propose a condition for asymmetric coexistence, which has been evaluated numerically:

$$\rho > \frac{(\omega^2-1)(2\omega^4 - 6\omega^3 + \omega^2 + 6\omega - 2)}{\omega^2(2\omega^2 - 4\omega + 1)(2\omega^2 - 3)} + 2\sqrt{\frac{(\omega-1)^5(\omega+1)^2(\omega^3 - \omega^2 - 3\omega + 1)}{\omega^4(2\omega^2 - 4\omega + 1)^2(2\omega^2 - 3)^2}}. \quad (8)$$

Although this expression is not very telling, its illustration in Figure 4b demonstrates how different this criterion is from that of the haplodiploid model. In the diploid model, males and females evolve on the same fitness landscape. Therefore, both males and females benefit from heterozygote advantage. This reduces the influence of the hybrid incompatibility on the optimal location of the population in genotype space, which thereby makes asymmetric coexistence less likely. Indeed, a heterozygote advantage of $\omega - 1 = \sigma > \approx 0.14$ is sufficient to ensure symmetric coexistence for all recombination rates, whereas in the haplodiploid model, $\sigma > \sqrt{2} - 1 \approx 0.41$ is necessary for symmetric coexistence independent of the recombination rate.

3.8 | General stability conditions in the haplodiploid model

Using the results derived for the case of lethal epistasis, and by means of critical examination of the existence and stability conditions that we were able to compute analytically, we arrived at several illustrative conjectures delimiting the evolutionary outcomes in the haplodiploid model when epistasis is not lethal ($\gamma_1, \gamma_2 \neq 1$). These were all confirmed by extensive numerical simulations (see Mathematica Online Supporting Information). Note that assortative mating was not considered here.

First, strong heterozygote advantage can always override the effect of epistasis. Specifically, if

$$\omega > \sqrt{2}, \quad (9)$$

the evolutionary outcome is always symmetric coexistence, regardless of the values of γ_1 and γ_2. This is true not only for a single pair of interacting loci, but also for an arbitrary number of independent incompatibility pairs, because the detrimental effects caused by each incompatibility pair are eventually resolved independently (see also the section on multiple loci below). This result can be deduced from Equation (7) for $\rho = 0.5$ and therefore corresponds to an upper bound: if heterozygote advantage is very strong, recombination no longer affects the outcome.

Second, recombination is a key player to determine whether compromise or exclusion can occur. In particular,

$$\rho < \frac{\omega^2 - 1}{\omega^2} \quad (10)$$

is a sufficient condition for the observation of symmetric coexistence, independent of the strength and symmetry of epistasis. This makes intuitive sense, because hybrid incompatibility is masked until gametes are broken up by recombination.

Thirdly, for symmetric epistasis ($\gamma_1 = \gamma_2$), there are three possible equilibrium patterns: symmetric coexistence, exclusion and tristability of the two former types of equilibria. A necessary and sufficient condition for observation of anything but symmetric coexistence is

$$\omega < \sqrt{2} \quad \text{and} \quad \rho > \frac{\omega^2 - 1}{\omega^2} \quad \text{and} \quad \gamma_1 = \gamma_2 > \frac{2(\omega - 1)}{\omega}. \quad (11)$$

If the recombination rate ρ and the epistatic effects γ_1, γ_2 are very close to this limit, there is tristability; if they are far away, there is exclusion (cf. Figure 5).

Finally, for asymmetric epistasis ($\gamma_1 \neq \gamma_2$), the dynamics display the whole range of possible evolutionary outcomes: symmetric coexistence, asymmetric coexistence, single-locus polymorphism, exclusion, as well as tristability of exclusion *and* symmetric coexistence, and single-locus polymorphism *and* symmetric coexistence. The local stability criterion for the stability of the monomorphic equilibria (i.e., the criterion for exclusion, or tristability of exclusion and symmetric coexistence) is

$$\omega < \sqrt{2} \quad \text{and} \quad \rho > \frac{\omega^2 - 1}{\omega^2} \quad \text{and} \quad \gamma_2 > \frac{2(\omega - 1)}{\omega}. \quad (12)$$

Thus, if epistasis is strong as compared with heterozygote advantage, no degree of asymmetry is sufficient to promote a compromise

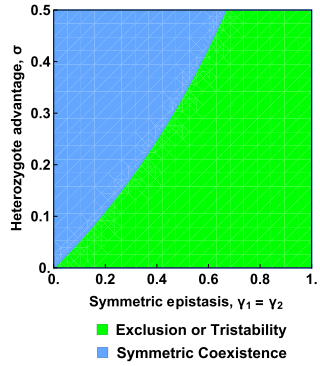


FIGURE 5 In haplodiploids, symmetric coexistence requires that heterozygote advantage, σ, is strong enough to both compensate for recombination such that the condition in Equation (10) is fulfilled (see also Figure 4a), and to overcome the deleterious effects of epistasis, as expressed by condition 11 for symmetric epistasis [Colour figure can be viewed at wileyonlinelibrary.com]

between males and females (i.e., single-locus polymorphism or asymmetric coexistence). In fact, we observe the following necessary (but not sufficient) condition for a single-locus polymorphism:

$$\omega < \sqrt{2} \text{ and } \rho > \frac{\omega^2 - 1}{\omega^2} \text{ and } \gamma_1 > \frac{2(\omega - 1)}{\omega} \text{ and } \gamma_2 < \frac{2(\omega - 1)}{\omega}. \quad (13)$$

Hence, only a tight balance between the selective pressures of epistasis and heterozygote advantage in combination with asymmetry of the hybrid incompatibility promotes a long-term equilibrium with compromise.

3.9 | An extension to multiple loci

3.9.1 | Incompatibilities involving four loci

Above, we have demonstrated that recombination is an essential player when determining whether exclusion or coexistence is the long-term outcome in the haplodiploid dynamics. In order to see how our results change in the (biologically relevant) case of multiple hybrid incompatibilities, we implemented the dynamics for four loci. Given the complexity of the system, we considered only lethal incompatibilities, that is $\gamma_i = 1$ for all interactions *i*. With this extension, we consider two scenarios. First, in the "*pairwise*" case, we consider pairs of independent hybrid incompatibilities, where we assume that the incompatible loci are located next to each other (locus **A** interacts with locus **B** at recombination distance ρ_{12}, and locus **C** with locus **D** at recombination distance ρ_{34}), which leaves four viable male haplotypes ($A_+B_+C_+D_+$, $A_+B_+C\_D\_$, $A\_B\_C_+D_+$ and $A\_B\_C\_D\_$). Second, in the "*network*" case we assume that all loci interact such that only two viable male haplotypes exist $A_+B_+C_+D_+$ and $A\_B\_C\_D\_$. In both cases, heterozygote advantage is defined as before, now acting on all four loci multiplicatively.

Under this model, we derived the conditions under which exclusion (the purging of all foreign alleles resulting in a monomorphic equilibrium) is locally stable (cf. Mathematica Online Supporting Information). For the pairwise case, exclusion is stable only if heterozygote advantage is relatively weak:

$$\omega < \min\left[\frac{1}{\sqrt{1-\rho_{12}}}, \frac{1}{\sqrt{1-\rho_{34}}}\right], \quad (14)$$

where ρ_{ij} is the recombination rate between neighbouring loci *i* and *j*. Note that this is independent of the recombination rate between noninteracting loci, here ρ_{23}. If $\rho_{12} = \rho_{34}$, this expression is equivalent to Equation (7) (Figure 4a). Overall, this condition indicates that exclusion, which we define as the fixation of one of the parental haplotypes, is less likely with four interacting loci than with two. This is because the fate of the two pairs of incompatibilities is decided independently, and exclusion requires that both pairs of incompatibilities fix for the same parental haplotype.

For the network case, the condition for stability of exclusion (see also Figure S3) is

$$\omega < \left((1-\rho_{12})(1-\rho_{23})(1-\rho_{34})\right)^{-\frac{1}{4}}. \quad (15)$$

In this scenario, exclusion is a more likely outcome with two pairs of incompatibilities than with one. This is because there are more unfit intermediate types in this scenario as compared with the pairwise model. Specifically in males, 14 of the 16 possible haplotypes do not survive to adulthood. To compensate for this fitness cost, any alternative evolutionary outcome requires strong heterozygote advantage.

3.9.2 | Incompatibilities involving an arbitrary number of loci

From the results for two and four loci, we derived a conjecture that generalizes to an arbitrary number of loci. For the pairwise case, Equation (14) can be generalized to

$$\omega < \min\left[\frac{1}{\sqrt{1-\rho_{ij}}}\right], \quad (16)$$

with *i* and *j* representing neighbouring interacting loci. Note that this result holds only if interacting loci are next to each other on the same chromosome, or if all loci are unlinked (in which case it simplifies to $\omega < \sqrt{2}$).

For the network case, Equation (15) generalizes to

$$\omega < \left(\prod_{\substack{i=1 \\ j=i+1}}^{n-1} (1-\rho_{ij})\right)^{-\frac{1}{n}}, \quad (17)$$

with *i* and *j* neighbouring loci and *n* the total number of loci in the network. Unlike in the pairwise case, the results for the network case do not depend on the genetic architecture (here, the ordering of loci along the genome).

We can therefore deduce that, for the pairwise case, exclusion becomes increasingly unlikely as the number of pairs of independent hybrid incompatibilities involved in the genetic barrier increases. Conversely, the opposite result is observed for the network case: more loci make exclusion a more likely outcome, but each additional interaction contributes less (cf. Figure S3).

3.10 | Increased assortative mating counteracts recombination and heterozygote advantage

Increasing the strength of assortative mating, $\alpha > 0$, counteracts the hybridization-favouring effect of heterozygote advantage, because matings between individuals with the same genotype are more common under stronger, positive assortment. Under sufficiently large positive α, exclusion is unavoidable. In general, increasing α leads to less maintenance of polymorphism in the population (Figure S4). Conversely, when $\alpha < 0$, which means that individuals prefer to mate with those whose genotype is most different from their own, polymorphism is more likely to be maintained in the population.

Also with assortative mating, recombination remains a key player in determining the evolutionary outcome. When $\alpha<0$ and recombination is small, symmetric coexistence is possible even in the absence of heterozygote advantage (i.e., $\sigma = 0$; Figure S4). Indeed, under these conditions and assuming epistasis is very strong, (almost) all hybrid males are dead and only parental males survive. This "disassortative" mating ($\alpha<0$) creates a bias for the rare male haplotype. For example, if one female genotype increases in frequency, it will seek mainly the males of the other parental haplotype to reproduce with (which are currently rare, as their frequency is directly tied to the frequency of the females in the previous generation). This will increase their reproductive success, which leads to an increase in this haplotype frequency. Therefore, under this mate choice regime, we would observe a stable population composed almost exclusively of the A_+B_+ and A_-B_- haplotypes.

3.11 | Differences between the haplodiploid and the diploid systems

As described above and illustrated in Figure 6, the resulting haplodiploid dynamics display a wider range of possible evolutionary outcomes than the diploid dynamics. Because both males and females profit from heterozygote advantage in the diploid model, polymorphism is always maintained; in other words, even the smallest amount of heterozygote advantage promotes the creation or maintenance of diversity in diploids (Table S3). Conversely, in the haplodiploid model, polymorphism can be lost either at one or both loci, resulting in a single-locus polymorphism or exclusion. Thus, alleles responsible for incompatibilities are more effectively purged in the haplodiploid model.

In the diploid model, a single-locus polymorphism is never stable: assume locus A is polymorphic and locus B is fixed for allele B_+. Then, a new mutant-carrying allele B_- will always have a selective advantage regardless of the genotype in which it first appears (Table S3). In contrast, in the haplodiploid model, this is no longer true as the mutant-carrying allele B_- will have a much lower fitness in males when associated with allele A_+. Therefore, if the cost of generating this unfit haplotype in males overrides the advantage in females, and allele A_+ is at high frequency, then invasion of the B_- mutant may be prevented, leading to the stability of the single-locus polymorphism.

When polymorphism is maintained at both loci at equilibrium (i.e., asymmetric and symmetric coexistence), epistasis creates associations between the compatible alleles, which results in elevated linkage disequilibrium (LD). Recombination breaks the association between alleles; thus, high recombination decreases normalized LD (D', where $D' = \frac{LD}{D_{max}}$ (Lewontin, 1964); Figure S5). D' increases with the strength of heterozygote advantage at low recombination rates because it maximizes the discrepancy between highly fit double-heterozygote females on the one hand that can, under low recombination rate, still produce many fit male offspring, and introgressed females on the other, who are less fit and produce many unfit hybrid males.

In Figure S6, we compare the normalized LD (i.e., D') between the haplodiploid and diploid models. When polymorphism is maintained at both loci in both the haplodiploid and diploid model, normalized LD is always larger in haplodiploids than diploids. The

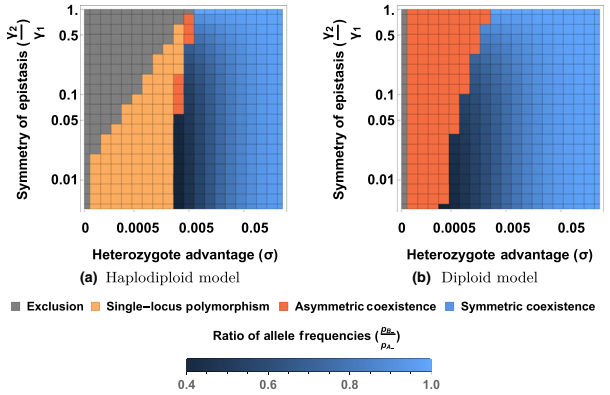


FIGURE 6 More evolutionary outcomes are possible in (a) the haplodiploid than (b) the diploid model. The *y*-axis shows the degree of asymmetry of epistasis, displayed as the ratio of the two epistasis parameters (γ_2/γ_1) for a constant value of $\gamma_1 = 0.01$. For symmetric coexistence, the locally stable equilibrium can be at any point on the diagonal $p_{B_-} = 1 - p_{A_-}$, where p_{A_-} and p_{B_-} denote the allele frequencies of the–allele at the respective locus. Blue shading illustrates the location of the equilibrium at symmetric coexistence: darker shades correspond to a bigger disparity in allele frequencies. This is the case when the asymmetry of the two epistasis parameters is large (i.e., smaller values on the *y*-axis) because smaller values of γ_2 favour the A_-B_+ haplotype over the A_+B_- haplotype. (Here, $\gamma_1 = 0.01, \rho = 0.5, \alpha = 0.$)

difference in normalized LD between haplodiploids and diploids is maximized for intermediate recombination rates, where recombination is strong enough to create unfit hybrid genotypes, but not efficient enough to break the associations that are generated. Due to the increased selection against hybrid incompatibility in haploid males in the haplodiploid model, the normalized LD is usually 2–3 times higher in the haplodiploid as compared with the diploid model.

Thus, the hybrid incompatibility leaves a statistical signature in a population, even if the population finds itself at an equilibrium. The increased association across the genome, exhibited if the interacting loci are on the same chromosome, may also result in an underestimate of the recombination rate. Although both the diploid and the haplodiploid models display the elevated LD signal, it is much more pronounced in the haplodiploid scenario. This is because only an eighth of the possible diploid male genotypes suffers the cost of the incompatibility as compared to half of the possible haploid male genotypes.

4 | DISCUSSION

Multiple recent studies have highlighted the pervasive nature of hybridization and its potential consequences for diversification and speciation (Abbott et al., 2013; Montecinos et al., 2017; Runemark et al., 2017). We here modelled the fate of a hybrid population in a scenario in which hybridization is simultaneously favoured and selected against, inspired by a natural population of hybrid ants that simultaneously displays heterosis and hybrid incompatibility. In addition, both adaptive introgression and hybrid incompatibilities have been identified in natural systems (Corbett-Detig, Zhou, Clark, Hartl, & Ayroles, 2013; Heliconius Genome Consortium, 2012; Whitney et al., 2015) and it is therefore likely that both processes may occur simultaneously during a single hybridization event. Furthermore, we were interested in comparing the long-term evolution of populations exposed to these opposing selective pressures under different ploidies (haplodiploid versus diploid), as it has been argued that haplodiploids might speciate more easily than diploids (Lohse & Ross, 2015). Finally, the comparison of ploidies can also be transferred to the case of diploid species with sex chromosomes, in which the described fitness landscape results in the diploid dynamics on the autosomes, and in the haplodiploid dynamics on the X/Z chromosome.

Our model considers a population in which heterozygote advantage and hybrid incompatibility act simultaneously on the same pair of loci, which creates a rugged fitness landscape with a ridge of high-fitness heterozygote genotypes, adjacent to which there are holes of incompatible double homozygotes (Figure 2a). In haplodiploids, haploid males cannot profit from heterozygote advantage but suffer strongly from hybrid incompatibility (Figure 2b). This results in a conflict of ploidies/sexes over the optimal location in the fitness landscape, because haploid males survive best if one parental haplotype is fixed, whereas diploid females profit from maximum heterozygosity. Although females suffer from the same incompatibility as males, their presence is mainly masked in the diploid individuals because of the recessivity of the hybrid incompatibility. This is similar to Haldane's rule (Charlesworth et al., 1987; Koevoets & Beukeboom, 2009).

4.1 | How ploidy matters

We found that, in the haplodiploid model, there exist four different stable outcomes of the conflict over hybrid status (Figure 3): exclusion, where "males/haploids win"; symmetric coexistence, where "females/diploids win"; and two outcomes, single-locus polymorphism and asymmetric coexistence, where a compromise between male costs and female benefits is mediated by high frequencies of introgressed females. In fact, as low-frequency heterozygotes are favoured both in males and in females in the diploid model, while only suffering the hybrid cost if introgressed alleles rise to high frequencies, exclusion and single-locus polymorphism never occur in the diploid model, which reduces the number of possible outcomes to asymmetric and symmetric coexistence. Therefore, consistent with Pamilo (1979); Pamilo and Crozier (1981); Patten, Carioscia, and Linnen (2015), we found that introgression and maintenance of polymorphism, and thus long-term hybridization, are less likely in haplodiploids as compared to diploids.

Prior work has found that in haplodiploid species traditional sexual conflict tends to be resolved in favour of females because genes spend two-thirds of their time in females (Albert & Otto, 2005). In our model, the co-occurrence of heterozygote advantage and hybrid incompatibility also creates an apparent sexual conflict that is caused by the difference in ploidy between the sexes. For several scenarios, we here derived the conditions for whether this conflict is resolved in favour of diploid females or haploid males. We find, that in addition to the strength of selection, recombination is a major player (cf. Figure 4 and Equation 12); only if recombination breaks up gametes, the hybrid incompatibility is expressed. With free recombination, that is, if the interacting genes are found on separate chromosomes, heterozygote advantage has to be very strong to counteract the hybrid incompatibility. We find that it has to be on the same order of magnitude as the strength of the incompatibility, but can be slightly lower in its absolute value. For example, heterozygote advantage with strength 41% is sufficient to result in symmetric coexistence even if the incompatibility is lethal (Figure 4b). Thus, under consideration of absolute magnitude across the full parameter range, our results are consistent with prior work. However, reported cases and potential mechanisms of hybrid incompatibility indicate that large effects are feasible, whereas observed cases of heterozygote advantage or heterosis of large effect are relatively rare (Hedrick, 2012). Therefore, it may well be that under natural circumstances, the conflict modelled here may indeed be likely to be resolved via purging of at least one incompatible allele and thus in favour of males/haploids.

As expected in the presence of epistasis, we observed that linkage disequilibrium (LD) is elevated at all polymorphic stable states (i.e., for symmetric and asymmetric coexistence) both in the diploid

and haplodiploid models, especially at intermediate recombination rates. This is particularly true for haplodiploids, which display about 2–3 times the LD of the diploid model with the same parameters. Transferred to the context of X/Z chromosomes, this is consistent with observations of larger LD on the X chromosome as compared with autosomes (Li & Merilä, 2010; Sandor et al., 2006; Wall, Andolfatto, & Przeworski, 2002). It has been argued that this is because selection is more effective on X-linked loci: recessive deleterious mutations are more visible to selection in haploid individuals (Charlesworth et al., 1987). However, a hybrid incompatibility accompanied by heterosis/heterozygote advantage as in our model may not be purged but create a continuous high-LD signal in an equilibrium population. This can potentially result in less efficient recombination and in underestimates of recombination rates on X chromosomes (because recombined individuals are not observed).

4.2 | Generalization to multiple incompatibilities

Exclusion remains a stable solution when we extend the model to multiple loci and incompatibilities. We describe an interesting difference between multiple independent pairs of incompatibilities, and multiple loci that all interact with each other: in the latter case, exclusion becomes increasingly probable because the number of viable males decreases. This scenario of higher-order epistasis has recently received attention with regards to speciation (Fraïsse et al., 2014; Kulmuni & Westram, 2017; Paixão, Bassler, & Azevedo, 2014), and it will be interesting in the future to identify molecular scenarios (e.g., involving biological pathways) that could result in such incompatibilities. In contrast, exclusion becomes less likely in the case of independent incompatibility pairs, where each incompatibility has to be purged independently and in the same direction, for exclusion to occur. Here, mechanisms that reduce the recombination rate, such as inversions, could potentially invade and tilt the balance towards coexistence and thus maintenance of polymorphism in the hybrid population. It is important to note that the independent purging of incompatibilities, which leads to a decreasing probability of exclusion with increasingly many incompatibility pairs, is only true in effectively infinite-sized populations. In small populations, we expect that exclusion becomes a more likely scenario, especially if lethal incompatibility pairs are present.

4.3 | Model assumptions

We chose a classical population-genetic modelling approach (Bürger, 2000; Nagylaki, 1992) to study how the co-occurrence of heterozygote advantage and hybrid incompatibility affect the long-term dynamics of a hybrid population. By treating the problem in a deterministic framework and considering only two loci throughout most of the manuscript, we greatly oversimplify the situation in the natural population that inspired our model. However, at the same time this allowed us to gain a general insight (often by means of analytical expressions) into how opposing selective pressures in genomes may be resolved, and to contrast these outcomes between haplodiploid and diploid systems. In addition to some obvious mechanisms at play in natural populations, which we ignore in our model (e.g., random genetic drift), some extensions of the model could be interesting to elaborate on in the future. For example, the ant populations that inspired our model represent networks of interacting nests and with many queens per nest and potentially different mating flight timing that could depend on genetic or environmental determinants (e.g., sun exposure in the spring). Thus, for the purpose of population-genetic inference of the evolutionary history (and potential evolutionary fate) of the hybrid ant population in Finland, it would be desirable to incorporate population structure, uneven sex ratios at birth, and sex-biased dispersal into the model, and obtain population-genomic data to infer evolutionary parameters.

4.4 | Is the natural population at an equilibrium of asymmetric coexistence?

Model fitting results (see Methods, Results, and Discussion in Supporting Information) are inconclusive about the fate of the natural ant population that inspired our model. Our results suggest that the natural population might be approaching an evolutionary outcome that allows a compromise between male and female interests; either as single-locus polymorphism or via asymmetric coexistence. In particular, our model is able to explain the unusual skew in the population, where *F. aquilonia*-like parental genotypes far outnumber *F. polyctena*-like genotypes (see Supporting Information). Furthermore, the high recombination rates and strong prezygotic mechanisms operating in the natural population (Kulmuni & Pamilo, 2014; Kulmuni et al., 2010), are consistent with a parameter domain in our models at which asymmetric coexistence can be stably maintained over a wide range of values of female hybrid advantage. More complex models, for example including more than two incompatibility loci, may be better able to explain the high frequencies of introgressed as compared to parental females observed in the natural hybrid population. As argued in the Results, interactions at or between multiple loci should result in steeper differences of introgressed-allele frequencies across life stages than our model is able to produce.

4.5 | Implications for hybrid speciation

Our model illustrates how the co-occurrence of heterozygote advantage and hybrid incompatibility affects haplodiploid and diploid populations. We can hypothesize how these different outcomes may provide an engine to hybrid speciation, or which other long-term evolutionary scenarios we expect to arise. The case of exclusion, which is possible only in the haplodiploid model, will lead to loss of diversity in the hybrid population, and, in the two-locus case, should result in the reversion of the hybrid population into one of its parental species. However, if multiple pairs of interacting loci are resolved independently, they may be purged randomly towards either parent, which could result in a true hybrid species that is isolated from both its parental species (Buerkle, Morris, Asmussen, & Rieseberg, 2000; Butlin & Ritchie, 2013; Schumer, Cui, Rosenthal, & Andolfatto,

2015). In fact, our finding that exclusion is less likely to occur in populations with multiple pairs of interacting loci may result from exactly this mechanism, but it is beyond the scope of this manuscript to explore this further.

The long-term fate of the population is less straightforward to anticipate in the case of polymorphic stable equilibria. For any of these, heterozygote advantage is strong enough to stabilize the polymorphism either at one or both loci. Without further occurrence of functional mutations, males (in the haplodiploid model) and double homozygotes for the incompatible alleles will continue to suffer a potentially large fitness cost. Mechanisms that could reduce this cost would be increased assortative mating or decreased recombination. However, neither of these would necessarily cause isolation from the parental species, unless they involved additional hybrid incompatibilities that isolate the hybrid population from its parental species. Alternatively, mutations that lower the hybrid fitness cost could invade, which would result in a weakening of species barriers and promote further introgression from the parental species. This indicates that any scenario in which polymorphic equilibria are stable may indeed be an unlikely candidate for hybrid speciation. Considering that such stable polymorphism (either as symmetric or asymmetric coexistence) is the only possible outcome in the diploid model, this results in the prediction that hybrid speciation would be more likely in a haplodiploid scenario. This is an interesting observation that is in line with other predictions that haplodiploids speciate more easily, that X/Z chromosomes are engines of speciation (Lima, 2014) and that hybrid speciation is rare (Schumer, Rosenthal, & Andolfatto, 2014).

4.6 | Relevance of the model for sex chromosomes

Haplodiploids and X/Z chromosomes have a similar mode of inheritance, where one sex carries a single copy of the chromosome and the other carries two copies. Therefore, our results apply equally to cases of X-to-X or Z-to-Z hybrid incompatibilities (Lohse & Ross, 2015). Although haplodiploid systems do not include all of the unique evolutionary phenomena exhibited by sex chromosomes (Abbott et al., 2017), our results for haplodiploids are relevant for sex chromosomes. Our model predicts the long-term evolution of a population under the simultaneous influence of heterozygote advantage and hybrid incompatibility and indicates the signatures that this type of fitness landscape could leave depending on whether it finds itself on an X chromosome or an autosome.

First, the complex selection pressure imposed by the co-occurrence of heterozygote advantage and hybrid incompatibility manifests itself as an apparent sexual conflict on the X chromosome/in haplodiploids. This conflict is caused by the ploidy difference between the sexes. Here, the same fitness landscape that would be masked on an autosome and result in a stable polymorphism creates a signal of sexually antagonistic selection on an X chromosome. Most importantly, this signal is created without the need for direct sexually antagonistic selection on single functional genes that have a sex-specific antagonistic effect. Thus, our model proposes an additional mechanism by which sex chromosomes can appear as hotspot of sexual conflict (e.g., Gibson, Chippindale, & Rice, 2002; Pischedda & Chippindale, 2006).

Second, we find that purging of incompatibilities is more likely in the haplodiploid model, and thus on X/Z chromosomes. This is consistent with the faster-X theory (Charlesworth et al., 1987). However, only if recombination is strong enough, incompatibilities will become visible to selection and purged in the presence of heterozygote advantage. If they are not purged, they may persist as a long-term polymorphism, invisible to most empirical approaches, and confound population-genetic inference by creating signals of elevated linkage disequilibrium.

5 | CONCLUSION

Hybridization is observed frequently in natural populations and can have both deleterious and advantageous effects. We here show how diverse outcomes can be produced even under a rather simple model of a single hybrid population, in which heterozygote advantage and hybrid incompatibility are occurring at the same time. Consistent with previous theory on haplodiploids and X/Z chromosomes, we found that incompatible alleles are more likely to be purged in a haplodiploid than in a diploid model. Nevertheless, our results suggest that long-term hybridization can occur even in the presence of hybrid incompatibility, and if there are many incompatible pairs or many loci involved in the incompatibility. The evolutionary fate of the Finnish hybrid ant population that inspired our model is difficult to predict; further population-genetic analysis will be necessary to gain a more complete picture of its structure and evolutionary history.

ACKNOWLEDGEMENTS

We thank Laura Cêtre for her work on a previous version of the model. We thank Pekka Pamilo, Bret Payseur, two anonymous reviewers and the members of the Bank and Kulmuni laboratories for discussion of the manuscript. This research was supported by the Fundação Calouste Gulbenkian and in part by the National Science Foundation under Grant No. NSF PHY-1125915. JK was supported by the Human Frontier Science Program, Finnish Cultural Foundation and Academy of Finland (252411 to CoE in Biological Interactions).

DATA ACCESSIBILITY

The complete documentation of all steps of the analysis is available as a Mathematica Online Supporting Information. Ant colony data are provided as Table S1; genotype frequency data were obtained from Kulmuni and Pamilo (2014).

AUTHOR CONTRIBUTIONS

C.B., J.K. and R.B. designed research; A.B. and C.B. developed the models; A.H.G. performed simulations and data analysis; and all authors interpreted the results and wrote the manuscript.

ORCID

Ana-Hermina Ghenu http://orcid.org/0000-0002-9088-2232
Alexandre Blanckaert http://orcid.org/0000-0003-2735-7491
Jonna Kulmuni http://orcid.org/0000-0002-8852-0793
Claudia Bank http://orcid.org/0000-0003-4730-758X

REFERENCES

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Abbott, J., Nordén, A., & Hansson, B. (2017). Sex chromosome evolution: Historical insights and future perspectives. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *284*(1854), 20162806. https://doi.org/10.1098/rspb.2016.2806

Albert, A., & Otto, S. (2005). Sexual selection can resolve sex-linked sexual antagonism. *Science*, *310*(5745), 119–121. https://doi.org/10.1126/science.1115328

Arnqvist, G., & Rowe, L. (2005). *Sexual conflict*. Princeton, NJ: Princeton University Press. https://doi.org/10.1515/9781400850600

Bateson, W. (1909). Heredity and variation in modern lights. *Darwin and Modern Science*, *85*, 101.

Beresford, J., Elias, M., Pluckrose, L., Sundström, L., Butlin, R., Pamilo, P., & Kulmuni, J. (2017). Widespread hybridization within mound-building wood ants in Southern Finland results in cytonuclear mismatches and potential for sex-specific hybrid breakdown. *Molecular Ecology*, *26*(15), 4013–4026. https://doi.org/10.1111/mec.14183

Bernardes, J., Stelkens, R., & Greig, D. (2017). Heterosis in hybrids within and between yeast species. *Journal of Evolutionary Biology*, *30*(3), 538–548. https://doi.org/10.1111/jeb.13023

Buerkle, C. A., Morris, R. J., Asmussen, M. A., & Rieseberg, L. H. (2000). The likelihood of homoploid hybrid speciation. *Heredity*, *84*(4), 441–451. https://doi.org/10.1046/j.1365-2540.2000.00680.x

Bürger, R. (2000). *The mathematical theory of selection, recombination, and mutation*, Vol. *228*. Chichester, UK: Wiley.

Butlin, R. K., & Ritchie, M. G. (2013). Pulling together or pulling apart: Hybridization in theory and practice. *Journal of Evolutionary Biology*, *26*(2), 294–298. https://doi.org/10.1111/jeb.12080

Charlesworth, B., Coyne, J., & Barton, N. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, *130*(1), 113–146. https://doi.org/10.1086/284701

Chen, Z. (2013). Genomic and epigenetic insights into the molecular bases of heterosis. *Nature Review Genetics*, *14*(7), 471–482. https://doi.org/10.1038/nrg3503

Chen, C., Zhiguo, E., & Lin, H.-X. (2016). Evolution and molecular control of hybrid incompatibility in plants. *Frontiers in Plant Science*, *7*, 1208.

Corbett-Detig, R. B., Zhou, J., Clark, A. G., Hartl, D. L., & Ayroles, J. F. (2013). Genetic incompatibilities are widespread within species. *Nature*, *504*(7478), 135–137. https://doi.org/10.1038/nature12678

Crozier, R., & Pamilo, P. (1996). *Evolution of social insect colonies: Sex allocation and kin selection*. Oxford, UK: Oxford University Press.

De Cara, M., Barton, N., & Kirkpatrick, M. (2008). A model for the evolution of assortative mating. *The American Naturalist*, *171*(5), 580–596. https://doi.org/10.1086/587062

Dieckmann, U., & Doebeli, M. (1999). On the origin of species by sympatric speciation. *Nature*, *400*(6742), 354–357. https://doi.org/10.1038/22521

Dobzhansky, T. (1936). Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics*, *21*(2), 113.

Evans, J., Shearman, D., & Oldroyd, B. (2004). Molecular basis of sex determination in haplodiploids. *Trends in Ecology and Evolution*, *19*(1), 1–3. https://doi.org/10.1016/j.tree.2003.11.001

de la Filia, A., Bain, S., & Ross, L. (2015). Haplodiploidy and the reproductive ecology of arthropods. *Current Opinion in Insect Science*, *9*, 36–43. https://doi.org/10.1016/j.cois.2015.04.018

Fraïsse, C., Elderfield, J., & Welch, J. (2014). The genetics of speciation: Are complex incompatibilities easier to evolve? *Journal of Evolutionary Biology*, *27*(4), 688–699. https://doi.org/10.1111/jeb.12339

Gibson, J., Chippindale, A., & Rice, W. (2002). The X chromosome is a hot spot for sexually antagonistic fitness variation. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *269*(1490), 499–505. https://doi.org/10.1098/rspb.2001.1863

Hedrick, P. W. (2012). What is the evidence for heterozygote advantage selection?}. *Trends in Ecology & Evolution*, *27*(12), 698–704. https://doi.org/10.1016/j.tree.2012.08.012

Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487* (7405), 94.

Höllinger, I., & Hermisson, J. (2017). Bounds to parapatric speciation: A Dobzhansky-Muller incompatibility model involving autosomes, X chromosomes, and mitochondria. *Evolution*, *71*(5), 1366–1380. https://doi.org/10.1111/evo.13223

Jablonka, E., & Lamb, M. J. (1991). Sex chromosomes and speciation. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *243*(1308), 203–208. https://doi.org/10.1098/rspb.1991.0032

Johnson, N. A., & Lachance, J. (2012). The genetics of sex chromosomes: Evolution and implications for hybrid incompatibility. *Annals of the New York Academy of Sciences*, *1256*(1), E1–E22. https://doi.org/10.1111/j.1749-6632.2012.06748.x

Knegt, B., Potter, T., Pearson, N., Sato, Y., Staudacher, H., Schimmel, B., … Egas, M. (2017). Detection of genetic incompatibilities in non-model systems using simple genetic markers: Hybrid breakdown in the haplodiploid spider mite tetranychus evansi. *Heredity*, *118*(4), 311. https://doi.org/10.1038/hdy.2016.103

Koevoets, T., & Beukeboom, L. (2009). Genetics of postzygotic isolation and Haldane's rule in haplodiploids. *Heredity*, *102*(1), 16–23. https://doi.org/10.1038/hdy.2008.44

Kopp, M., Servedio, M. R., Mendelson, T. C., Safran, R. J., Rodríguez, R. L., Hauber, M. E., … van Doorn, G. S. (2017). Mechanisms of assortative mating in speciation with gene flow: Connecting theory and empirical research. *The American Naturalist*, *191*(1), 1–20.

Kraaijeveld, K. (2009). Male genes with nowhere to hide; sexual conflict in haplodiploids. *Animal Biology*, *59*(4), 403–415. https://doi.org/10.1163/157075509X12499949744225

Kulmuni, J., & Pamilo, P. (2014). Introgression in hybrid ants is favored in females but selected against in males. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(35), 12805–12810. https://doi.org/10.1073/pnas.1323045111

Kulmuni, J., Seifert, B., & Pamilo, P. (2010). Segregation distortion causes large-scale differences between male and female genomes in hybrid ants. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(16), 7371–7376. https://doi.org/10.1073/pnas.0912409107

Kulmuni, J., & Westram, A. M. (2017). Intrinsic incompatibilities evolving as a by-product of divergent ecological selection: Considering them in empirical studies on divergence with gene flow. *Molecular Ecology*, *26*(12), 3093–3103. https://doi.org/10.1111/mec.14147

Lewontin, R. (1964). The interaction of selection and linkage. I. general considerations; heterotic models. *Genetics*, *49–67*(1), 49.

Li, M.-H., & Merilä, J. (2010). Sex-specific population structure, natural selection, and linkage disequilibrium in a wild bird population as revealed by genome-wide microsatellite analyses. *BMC Evolutionary Biology*, *10*(1), 66. https://doi.org/10.1186/1471-2148-10-66

Lima, T. G. (2014). Higher levels of sex chromosome heteromorphism are associated with markedly stronger reproductive isolation. *Nature Communications*, *5*, 4743. https://doi.org/10.1038/ncomms5743

Lohse, K., & Ross, L. (2015). What haplodiploids can teach us about hybridization and speciation. *Molecular Ecology*, *24*(20), 5075–5077. https://doi.org/10.1111/mec.13393

Mallet, J. (2005). Hybridization as an invasion of the genome. *Trends in Ecology and Evolution*, *20*(5), 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Montecinos, A. E., Guillemin, M.-L., Couceiro, L., Peters, A. F., Stoeckel, S., & Valero, M. (2017). Hybridization between two cryptic filamentous brown seaweeds along the shore: Analysing pre- and postzygotic barriers in populations of individuals with varying ploidy levels. *Molecular Ecology*, *26*(13), 3497–3512. https://doi.org/10.1111/mec.14098

Muller, H. (1942). Isolating mechanisms, evolution and temperature. *Biology Symposium*, *6*, 71–125.

Nagylaki, T. (1992). *Introduction to theoretical population genetics*, Vol. *142*. Berlin, Germany: Springer-Verlag. https://doi.org/10.1007/978-3-642-76214-7

Orr, H. (1995). The population genetics of speciation: The evolution of hybrid incompatibilities. *Genetics*, *139*(4), 1805–1813.

Otto, S. P., & Day, T. (2007). *A biologist's guide to mathematical modeling in ecology and evolution*, Vol. *13*. Princeton, NJ: Princeton University Press.

Paixão, T., Bassler, K. E., & Azevedo, R. B. R. (2014). Emergent speciation by multiple Dobzhansky-Muller incompatibilities. *bioRxiv*, 8268.

Pamilo, P. (1979). Genic variation at sex-linked loci: Quantification of regular selection models. *Hereditas*, *91*(1), 129–133.

Pamilo, P., & Crozier, R. H. (1981). Genic variation in male haploids under deterministic selection. *Genetics*, *98*(1), 199–214.

Patten, M., Carioscia, S., & Linnen, C. (2015). Biased introgression of mitochondrial and nuclear genes: A comparison of diploid and haplodiploid systems. *Molecular Ecology*, *24*(20), 5200–5210. https://doi.org/10.1111/mec.13318

Pischedda, A., & Chippindale, A. K. (2006). Intralocus sexual conflict diminishes the benefits of sexual selection. *PLoS Biology*, *4*(11), e356. https://doi.org/10.1371/journal.pbio.0040356

Presgraves, D. (2008). Sex chromosomes and speciation in *Drosophila*. *Trends in Genetics*, *24*(7), 336–343. https://doi.org/10.1016/j.tig.2008.04.007

Runemark, A., Trier, C. N., Eroukhmanoff, F., Hermansen, J. S., Matschiner, M., Ravinet, M., … Saetre, G.-P. (2017). Variation and constraints in hybrid genome formation. *bioRxiv*, 107508.

Sandor, C., Farnir, F., Hansoul, S., Coppieters, W., Meuwissen, T., & Georges, M. (2006). Linkage disequilibrium on the bovine X chromosome: Characterization and use in quantitative trait locus mapping. *Genetics*, *173*(3), 1777–1786. https://doi.org/10.1534/genetics.106.059329

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*(5915), 737–741. https://doi.org/10.1126/science.1160006

Schluter, D., & Conte, G. (2009). Genetics and ecological speciation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(Suppl. 1), 9955–9962. https://doi.org/10.1073/pnas.0901264106

Schumer, M., Cui, R., Rosenthal, G. G., & Andolfatto, P. (2015). Reproductive isolation of hybrid populations driven by genetic incompatibilities. *PLOS Genetics*, *11*(3), 1–21.

Schumer, M., Rosenthal, G., & Andolfatto, P. (2014). How common is homoploid hybrid speciation? *Evolution*, *68*(6), 1553–1560. https://doi.org/10.1111/evo.12399

Schwarz, D., Matta, B., Shakir-Botteri, N., & McPheron, B. (2005). Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature*, *436*(7050), 546–549. https://doi.org/10.1038/nature03800

Seehausen, O., Butlin, R., Keller, I., Wagner, C., Boughman, J., Hohenlohe, P., … Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, *15*(3), 176–192. https://doi.org/10.1038/nrg3644

Servedio, M., & Noor, M. (2003). The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 339–364. https://doi.org/10.1146/annurev.ecolsys.34.011802.132412

Song, Y., Endepols, S., Klemann, N., Richter, D., Matuschka, F.-R., Shih, C.-H., … Kohn, M. (2011). Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. *Current Biology*, *21*(15), 1296–1301. https://doi.org/10.1016/j.cub.2011.06.043

Suomalainen, E., Saura, A., & Lokki, J. (1987). *Cytology and evolution in parthenogenesis*. Boca Raton, FL: CRC Press.

Wall, J. D., Andolfatto, P., & Przeworski, M. (2002). Testing models of selection and demography in *Drosophila simulans*. *Genetics*, *162*(1), 203–216.

Whitney, K. D., Broman, K. W., Kane, N. C., Hovick, S. M., Randell, R. A., & Rieseberg, L. H. (2015). Quantitative trait locus mapping identifies candidate alleles involved in adaptive introgression and range expansion in a wild sunflower. *Molecular Ecology*, *24*(9), 2194–2211. https://doi.org/10.1111/mec.13044

Whitney, K., Randell, R., & Rieseberg, L. (2010). Adaptive introgression of abiotic tolerance traits in the sunflower *Helianthus annuus*. *New Phytologist*, *187*(1), 230–239. https://doi.org/10.1111/j.1469-8137.2010.03234.x

Wolf, D., Takebayashi, N., & Rieseberg, L. (2001). Predicting the risk of extinction through hybridization. *Conservation Biology*, *15*(4), 1039–1053. https://doi.org/10.1046/j.1523-1739.2001.0150041039.x

Wolfram Research, Inc. (2016). *Mathematica v. 10.4.1.0*. Champaign, IL, USA. Retrieved from https://www.wolfram.com

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ghenu A-H, Blanckaert A, Butlin RK, Kulmuni J, Bank C. Conflict between heterozygote advantage and hybrid incompatibility in haplodiploids (and sex chromosomes). *Mol Ecol*. 2018;27:3935–3949. https://doi.org/10.1111/mec.14482

Chapter 3

Epistasis decreases with increasing antibiotic pressure but not temperature

When the going gets tough, the tough take a coffee break.
– Stephen Hawking

This chapter was published in *Philosophical Transactions of the Royal Society B: Biological Sciences* on 03 April 2023 as part of a special issue on predicting evolution. The student's contributions to the work are as follows: Claudia Bank, Isabel Gordo, and the student designed the research; The student performed molecular biology and gathered the data; The student analyzed the data with assistance from André Amado and consultation from Claudia Bank and Isabel Gordo; All authors wrote and edited the manuscript; The student fulfilled the role of the corresponding author.

PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Ghenu A-H, Amado A, Gordo I, Bank C. 2023 Epistasis decreases with increasing antibiotic pressure but not temperature. *Phil. Trans. R. Soc. B* **378**: 20220058.
https://doi.org/10.1098/rstb.2022.0058

Received: 31 August 2022
Accepted: 28 December 2022

One contribution of 13 to a theme issue 'Predicting evolution'.

Subject Areas:
evolution, genetics, microbiology, theoretical biology

Keywords:
environmental interactions, cryptic epistasis, RNA polymerase B (*rpoB*), genetic interactions, rifampicin

Author for correspondence:
Ana-Hermina Ghenu
e-mail: ana-hermina.ghenu@unibe.ch

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6440214.

THE ROYAL SOCIETY PUBLISHING

Epistasis decreases with increasing antibiotic pressure but not temperature

Ana-Hermina Ghenu[1,2,3], André Amado[2,3], Isabel Gordo[1] and Claudia Bank[1,2,3]

[1]Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras 2780-156, Portugal
[2]Division of Theoretical Ecology and Evolution, Institut für Ökologie und Evolution, Universität Bern, Baltzerstrasse 6, 3012 Bern, Switzerland
[3]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

A-HG, 0000-0002-9088-2232; AA, 0000-0002-9027-5037; IG, 0000-0003-1622-5143; CB, 0000-0003-4730-758X

Predicting mutational effects is essential for the control of antibiotic resistance (ABR). Predictions are difficult when there are strong genotype-by-environment (G × E), gene-by-gene (G × G or epistatic) or gene-by-gene-by-environment (G × G × E) interactions. We quantified G × G × E effects in *Escherichia coli* across environmental gradients. We created intergenic fitness landscapes using gene knock-outs and single-nucleotide ABR mutations previously identified to vary in the extent of G × E effects in our environments of interest. Then, we measured competitive fitness across a complete combinatorial set of temperature and antibiotic dosage gradients. In this way, we assessed the predictability of 15 fitness landscapes across 12 different but related environments. We found G × G interactions and rugged fitness landscapes in the absence of antibiotic, but as antibiotic concentration increased, the fitness effects of ABR genotypes quickly overshadowed those of gene knock-outs, and the landscapes became smoother. Our work reiterates that some single mutants, like those conferring resistance or susceptibility to antibiotics, have consistent effects across genetic backgrounds in stressful environments. Thus, although epistasis may reduce the predictability of evolution in benign environments, evolution may be more predictable in adverse environments.

This article is part of the theme issue 'Predicting evolution'.

1. Introduction

It has been debated for decades whether gene-by-gene interactions (G × G or epistasis) influence evolutionary processes (reviewed in [1,2]). Compelling evidence for the successive fixation of epistatic substitutions comes from studies of amino acid replacements over phylogenetic time scales (e.g. [3–5]). Moreover, experimental evolution has demonstrated that adaptive paths can be constrained by epistasis (e.g. [6]). Also, the environment can modulate mutational effects through gene-by-environment (G × E) interactions. Whereas multiple studies have demonstrated the existence of G × E interactions (e.g. [7–9]), our knowledge of the extent and the consequences of G × G × E interactions, i.e. those in which epistasis interacts with the environment, are limited ([10–13], reviewed in [2]).

G × G, G × E and G × G × E interactions complicate evolutionary predictions because they alter expected phenotypes or fitness of individuals [2,14]. Studying their extent and incorporating their effect into evolutionary predictions is daunted by the complexity of the genotype space and the myriad of environments that could be tested [15]. However, our lack of quantitative knowledge of G × G and G × E effects poses dangers to fields in which genetic and evolutionary models are applied: they could lead to the failed genetic rescue of endangered species, or the unexpected spread or maintenance of antibiotic resistance (ABR).

At the same time, the existence of G × E interactions is a crucial assumption in ABR evolution. By definition, the original genotype grows well in the antibiotic-free environment and poorly in the antibiotic environment. Conversely, a resistant genotype grows well in the antibiotic environment and is usually assumed to have low fitness in the antibiotic-free environment, resulting in so-called costs of ABR (reviewed in [16], but see [17]). In addition, G × E interactions with other environmental variables, such as high temperature or minimal media, can facilitate the evolution of resistant genotypes in the absence of antibiotics [6,18]. Although G × E interactions are central in ABR evolution, many questions about the role of G × E remain unexplored (but see [19–21]); for example, how does the G × E relationship between susceptible and resistant genotypes change across non-antimicrobial environmental axes, and how frequently is ABR modulated by the genetic background (i.e. through G × G effects).

ABR was first identified as a problem shortly after the introduction of antibiotics for clinical use [22], and it continues to be a leading cause of death worldwide [23]. For this reason, inferring potential resistances, cross-resistances, compensatory mutations and, ultimately, predicting the evolutionary trajectories of bacterial genomes in the presence of antibiotics is an application of evolutionary biology that is important for the health of humans and the planet [24,25]. Critically, G × G and G × E effects influencing ABR can complicate the identification of new resistances. For example, with G × G effects, resistance may depend on more than one easily identifiable allele. In addition, with G × E effects, laboratory evolution settings may not be informative of the fitness of genotypes in the wild.

Including G × G effects in the study of ABR requires the consideration of many potentially interacting genotypes. With epistasis, interactions can be specific between particular mutations due to mechanistic interactions of residues or proteins, for example (reviewed in [26]), or global, where any combination of mutations ultimately shows a pattern of G × G [27,28]. In contrast to increasing numbers of reports of ubiquitous epistasis [12,29,30], much existing work to date has considered single ABR genotypes at a time under the assumptions of additive effects (but see [31,32]). The theoretical concept of a fitness landscape, which maps every possible genotype to its fitness, captures G × G interactions and enables the study of how such interactions affect evolutionary trajectories (reviewed in [2,33]). Unless there is a known underlying pattern of interactions, epistasis makes evolution on the fitness landscape less predictable. Moreover, with added G × E interactions, the fitness landscape may change across environmental gradients, leading to changes in both single genotypes and epistasis [8,12] and to different possible evolutionary trajectories [13,34].

We here take a step towards quantifying G × G × E interactions by studying 15 small (i.e. two mutational-step) fitness landscapes across two environmental gradients and their combination, antibiotic concentration and temperature. In our fitness landscapes, we combine three known ABR mutations in the *rpoB* gene with five gene knock-outs, resulting in 24 genotypes that are screened for competitive fitness in 12 abiotic environments. Unlike previous studies, we study interactions of two different mutation types, single amino acid substitutions (more specifically, single-nucleotide polymorphisms, or SNPs) and whole gene deletions. By screening the fitness landscapes across a grid of environments, we quantify their change across environmental gradients, which is related to the robustness of fitness predictions under environmental change. Since the ABR mutations and gene knock-outs were selected based on their *not* exhibiting any direct, mechanistic interactions, one might *a priori* expect few G × G interactions and that the fitness of the double mutants would be additive as compared to the single mutants. On the other hand, non-specific epistasis (epistasis between mutations that do not directly interact mechanistically) has been observed ubiquitously [26]. Our fitness landscapes feature moderate epistasis in the benign environments, which decreases at higher antibiotic concentrations. Altogether, we conclude that our fitness landscapes are very predictable across environments.

2. Results

(a) Choice of genotypes

We selected rifampicin-resistant single-nucleotide mutations that were *a priori* known to differ in their performance between different temperature environments in M9 minimal medium. The ABR mutations modify single amino acids at the antibiotic target site in the gene RNA polymerase B (*rpoB*). Prior to our study, it was unknown how the mutants would perform in different combinations of antibiotic and high-temperature environments (i.e. we only knew about the effect of one environmental axis on its own). *RpoB* H526Y was established to show a trade-off between environment types: it confers high rifampicin resistance at 37°C but grows poorly as compared to the wild type at higher temperatures in the absence of antibiotic [7,35]. Conversely, *rpoB* S512F was established to have a synergy between environment types: it grows well both at high rifampicin concentrations and higher temperatures [7], separately. Finally, *rpoB* I572N is a weak rifampicin resistance mutation that was first identified during evolution to high temperature [6,36].

We initially tried to select knock-out mutations similarly as we had done for *rpoB* mutants: we used data from a large phenotypic screen of *Escherichia coli* (*E. coli*; accessed from https://ecoliwiki.org/tools/chemgen/) to select knock-out mutations that were known to differ in their performance between environment types. However, neither the bacterial colony size s-scores [37] nor the bacterial colony opacity and density s-scores [38] were able to predict the growth effects of Keio collection [39] knock-out mutants in batch culture environments (data not shown).

Ultimately, we selected knock-out mutations based on their *a priori* functional effects and the presence of segregating knock-out polymorphism in natural populations (accessed from the panX database at https://pangenome.org/; [40]). We selected four knock-outs of functional genes, all involved in how the cell interacts with its environment as follows: *marR* (multiple ABR regulator) detects and responds to chemical stressors, and its knock-out is sensitive to heat shock [41], conferring either resistance [42] or susceptibility [43] to some antibiotics but not rifampicin [37]; *nuoC* and *yidK* are both located on the plasma membrane and involved in transport [44,45]; and *waaP* impacts outer membrane stability through lipopolysaccharide biosynthesis [46]. The fifth gene, *ybfG*, is a pseudogene and was selected as a control since its knock-out should perform similarly to the wild-type genotype. None of the knock-outs interact directly with *rpoB* or other genes involved in transcription, protein synthesis or DNA

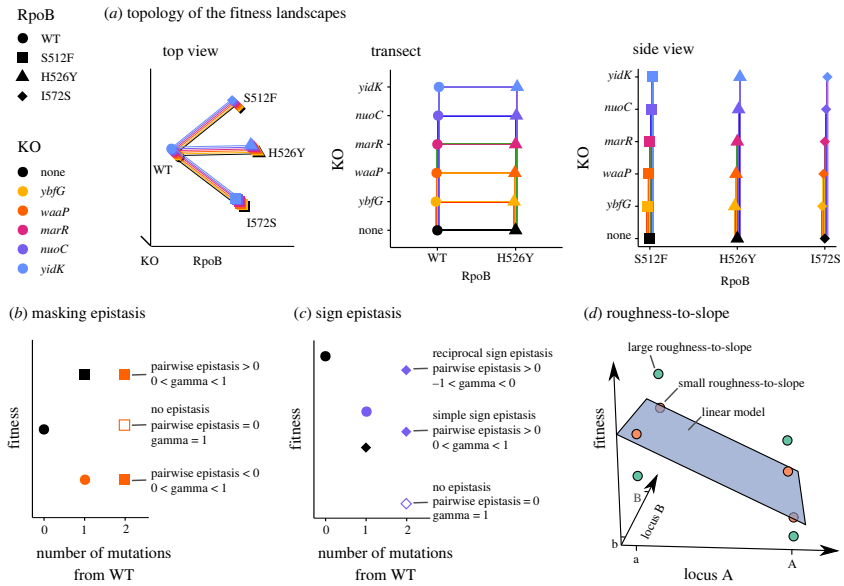


Figure 1. (*a*) Three-dimensional schematics depicting the topology of the studied fitness landscapes. The schematics show the same topology from different perspectives. All mutations are at unique, di-allelic loci. There are 15 fitness landscapes each composed of the wild type (WT), a *rpoB* single mutant, a knock-out (KO) single mutant, and the resultant double mutant. Lines connect genotypes that are one mutational-step apart. Depiction of 'classical' pairwise epistasis and gamma epistasis metrics for (*b*) masking and (*c*) sign epistasis. Masking epistasis is when the fitness effect of a beneficial or deleterious mutation on the wild-type background is zero on a different background. Sign epistasis occurs when the beneficial or deleterious effect of a mutation reverses in the presence of another mutation to become deleterious or beneficial, respectively. Reciprocal sign epistasis is a special case of sign epistasis where both mutations change sign on the other's background. The gamma statistic measures the correlation of fitness effects among the different genetic backgrounds where a mutation appears. The white-filled shape shows the expected fitness value *without* epistasis. For (*a–c*), the shape indicates the *rpoB* point mutation and the colour indicates which gene has been knocked-out; the black circle indicates the wild-type genotype without any mutations on *rpoB* and without any knock-outs. (*d*) Fitness landscapes exhibiting large (teal points) and small (orange points) roughness-to-slope ratios. The roughness-to-slope statistic measures how different the fitness landscape is from an additive, linear model (a purely additive landscape has a ratio of zero). The plane shows the fit of the multidimensional linear model to the fitness landscapes for two di-allelic loci. The vertical distance between the genotype (i.e. point) and the plane shows its residual. Epistasis is quantified as the ratio of the standard deviation of the residuals to the mean of the slopes of the plane. Note that (*d*) does not follow the colour and shape scheme of the other panels.

supercoiling. Therefore, any epistasis that will be observed should be an instance of 'non-specific epistasis' [26]. Non-specific epistasis is attributable to factors *other than* mechanistic interactions: for example, nonlinearities in the genotype-to-phenotype map. Figure 1*a* shows a schematic of the resulting topology of the fitness landscape when combining the ABR and knock-out mutations.

Illumina whole-genome re-sequencing was used to confirm the constructed mutants and identify any *de novo* mutations compared to the wild-type genotype (electronic supplementary material, table S3). See results in electronic supplementary material for details.

(b) Choice of environments

Since the wild type and knock-out single mutants are sensitive to rifampicin, the screened antibiotic concentrations were selected based on the wild-type dose response (electronic supplementary material, figure S14). The screened temperature environments were selected based on their physiological relevance for *E. coli* and its human host: 37°C is the temperature of a healthy human host, 40°C is the temperature of high fever and 42°C is near the maximum temperature that *E. coli* can withstand. Antibiotic-susceptible genotypes (i.e. genotypes that carry the wild-type allele at the *rpoB* locus) can be considered as 'specialists' of antibiotic-free environments because their growth deteriorates in the presence of the antibiotic, i.e. they show strong G × E interactions. Based on previous results, we expected that the *rpoB* S512F single mutant that grows well in the presence of antibiotic and at high temperature, separately, would exhibit 'generalist' growth in all environments with little G × E interaction.

(c) Measure of competitive fitness

The wild type, labelled by Green Fluorescent Protein (GFP), single and double mutants were competed against the same reference genotype, mCherry-labelled *rpoB* H526Y, in all environments by starting at approximately equal ratios and

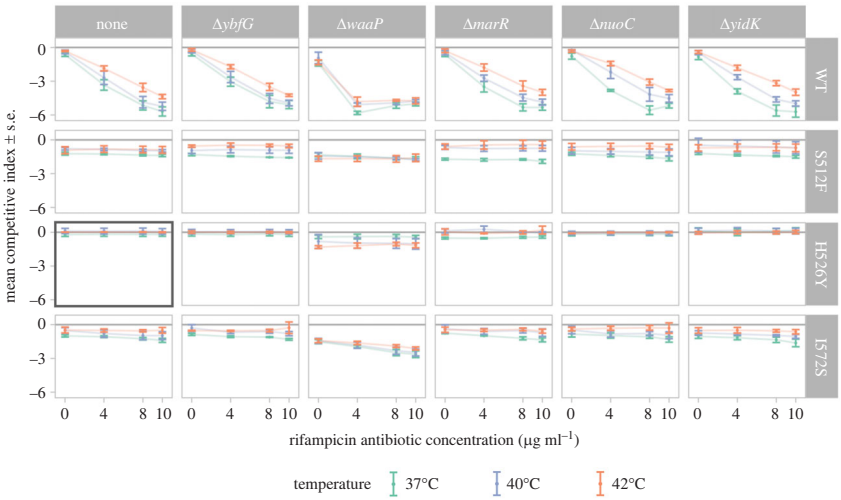


Figure 2. Estimated competitive index ($\hat{w}$) across all genotypes and environments. The *x*-axis shows the four rifampicin antibiotic concentration environments (μg ml^{-1}) and the *y*-axis shows the mean estimated competitive index with error-bars for standard errors ($n = 3$–4). Colours show the three temperature environments. Knock-out genotypes are shown in each column (with 'none' indicating the wild-type-like state of having no knock-outs) and *rpoB* antibiotic resistant mutants are shown in each row (with 'WT' indicating the susceptible wild-type *rpoB* sequence); therefore, the plot on the top left shows the wild type without any mutations and, similarly, the left column shows all *rpoB* single mutants while the top row shows all knock-out single mutants. Finally, the grey box indicates the *rpoB* H526Y genotype used as the reference for all competitions. All experiments were performed in M9 liquid minimal media with 0.4% glucose. The minimum inhibitory concentrations (MIC) for rifampicin (μg ml^{-1}) in LB agar medium are as follows: WT ≈ 6 [47], 20 < I572S < 50, S512F > 100 [48], H526Y > 768 [47].

a fixed inoculum size. The competitive index ($\hat{w}$), a proxy for fitness, was estimated after 20 h of batch culture growth as follows:

$$\hat{w} = \ln\left(\frac{D^f_{\text{comp}}/D^f_{\text{ref}}}{D^i_{\text{comp}}/D^i_{\text{ref}}}\right),$$

where D^i represents the initial density and D^f represents the final density as measured by flow cytometry. We found nearly identical $\hat{w}$ estimates when the fluorophores of the reference and competitor genotypes were swapped ($\hat{w}_{\text{GFP}} = \beta \cdot \hat{w}_{\text{mCherry}} + b$; adjusted $R^2 = 0.95$; $\beta = 0.96$, $p < 10^{-15}$; electronic supplementary material, figure S15 and table S4); including the effect of environment in the regression was not significant ($F_{11,127} = 1.27$, $p = 0.25$; electronic supplementary material, table S5 and figure S16), but adding the effect of genotype was significant ($F_{11,127} = 4.00$, $p < 10^{-4}$; electronic supplementary material, figure S17 and table S6), although genotype explains only about 1% more of the variation in the data (adjusted $R^2 = 0.96$ for overall regression) as compared to not including this effect. Therefore, we concluded that fluorescence is a neutral marker.

The mean $\hat{w}$ for all genotypes and environments is shown in figure 2. As expected, genotypes that did not have an ABR mutation (figure 2*a*) performed worse than the ABR reference competitor ($\hat{w} < 0$) when antibiotic was present. The Δ*waaP* single mutant (figure 2 top right facet) was the most susceptible to the antibiotic of all the investigated genotypes.

For all knock-out backgrounds, we compared the competitive fitness of the ABR mutations against the *rpoB* wild-type sequence in the absence of antibiotic to ascertain whether there was a cost of resistance. We observed no consistent cost of resistance across all backgrounds (electronic supplementary material, figure S18) and (in contrast to previous studies [7,36]) all three ABR mutations grew similarly to the wild type at higher temperatures.

(d) Multiple summary statistics found few gene-by-gene interactions

Several measures exist to determine the extent of G × G interactions (also termed epistasis). 'Classical' pairwise epistasis compares the observed mutational effect of the double mutant against the additive expectation from each of the two single mutants (figure 1*b*,*c*). We inferred pairwise epistasis by using the fitness of the wild-type genotype as a baseline and subtracting the fitness effects of both ABR and knock-out single mutants from the fitness effect of the double mutant (electronic supplementary material, figures S19 and S20 show examples without and with pairwise epistasis, respectively). We estimated pairwise epistasis for each of the 15 sets of double mutations in each of the 12 environments (electronic supplementary material, figure S21). We found significant pairwise epistasis only for the Δ*waaP* mutant. Across all ABR backgrounds, the Δ*waaP* mutant achieved a maximum value of positive pairwise epistasis at the low antibiotic concentration (4 μg ml^{-1}). In the absence of antibiotic and at higher antibiotic concentrations (8 – 10 μg ml^{-1}), the Δ*waaP* mutant exhibited no significant pairwise epistasis on most of the ABR backgrounds (small negative pairwise epistasis was observed with some ABR backgrounds). This dose-dependent peak of the pairwise epistasis arises because the Δ*waaP* single mutant has a large

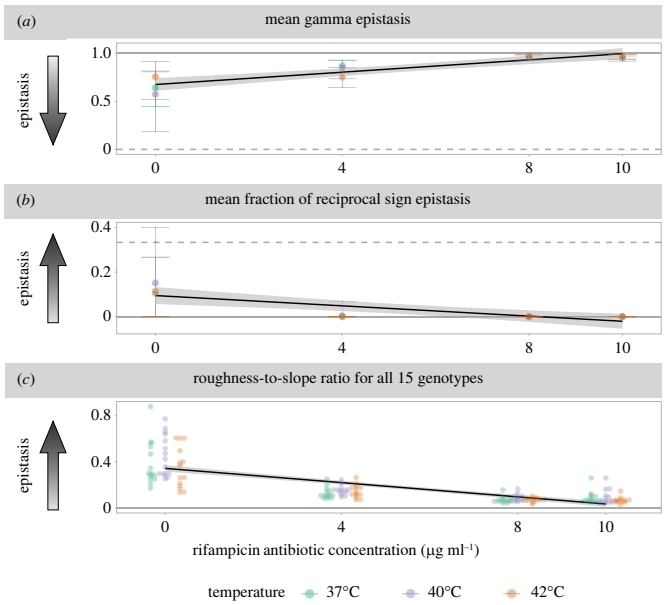


Figure 3. Antibiotic concentration has a significant negative trend with three measures of epistasis: (*a*) the mean gamma epistasis across all loci, (*b*) the mean fraction of loci exhibiting reciprocal sign epistasis and (*c*) the roughness-to-slope ratios by genotype. For each panel the *x*-axis shows the rifampicin antibiotic concentration in μg ml^{-1}, the colours show the temperature, and the *y*-axis shows the value of the epistasis measure. Arrows indicate the direction of stronger epistasis, solid horizontal lines show values of no epistasis, and dashed horizontal lines show the values of the maximum possible epistasis for each metric. The black trend lines show the significant negative correlations between antibiotic concentration and each epistasis measure, and the grey shaded regions show the 95% confidence interval of the corresponding least-squares estimated linear regression. For (*a*,*b*), the error bars show the 95% parametric bootstrap confidence intervals. For figure clarity, bootstrap confidence intervals are omitted in (*c*) but can be found in electronic supplementary material, figure S24, which shows the roughness-to-slope ratios for individual genotypes and environments. Significance is reported at the level $\alpha = 0.01$.

deleterious fitness effect at 4 μg ml^{-1} and, when this knock-out is found in combination with an ABR mutation, the ABR mutation rescues bacterial growth (e.g. electronic supplementary material, figure S20 second column). We detected almost no pairwise interactions of the Δ*waaP* mutant at higher antibiotic concentrations because the relative fitness effect of the Δ*waaP* mutation becomes much smaller as compared to the fitness effect of the ABR mutations (e.g. electronic supplementary material, figure S20 third and fourth columns).

Second, we inferred epistasis over the whole landscape using the gamma statistic [49], which measures the correlation of fitness effects among the different genetic backgrounds where a mutation appears (figure 1*b*,*c*). Mutations that exhibit little G × G interaction have similar effects on different backgrounds; consequently, their effects will be strongly correlated across backgrounds (gamma near one). On the other hand, a weak or negative correlation of the effects of a mutation in different backgrounds (gamma near zero or negative, respectively) indicates that a mutation exhibits G × G interactions. Overall, we observed gamma epistasis values near one, which are indicative of very little G × G interaction. Antibiotic concentration was positively correlated with gamma epistasis (figure 3*a*; $F_{1,10} = 52.7$, $p < 10^{-4}$, adjusted $R^2 = 0.825$) regardless of temperature. This means that the strength of G × G interactions decreased as antibiotic concentration increased and that it was independent of temperature. Focusing solely on the gamma epistasis of ABR mutants or knock-outs, respectively (electronic supplementary material, figure S22), we observed different qualitative but non-significant trends between the mutational classes and between the different temperatures. ABR mutations exhibited the strongest epistasis (i.e. gamma near zero) at the intermediate antibiotic concentration and increased with temperature. Conversely, knock-out mutations exhibited the strongest epistasis (again, gamma near zero) at the two highest antibiotic concentrations and decreased with temperature. Independent of the focal mutational class, there was less gamma epistasis at 42°C than at 37°C.

Third, we inferred the presence of simple sign and reciprocal sign epistasis (figure 1*c*). We identified simple sign epistasis as instances where the sign of the fitness effect of a mutation changed depending on its background in over five per cent of bootstrap samples. Then we identified reciprocal sign epistasis as instances where two mutations show sign epistasis on each other's background. Importantly, the presence of reciprocal sign epistasis is a prerequisite for multiple peaks in the fitness landscape [50]. In our study,

significant reciprocal sign epistasis occurred only in antibiotic-free environments but across all temperatures. This trend can be summarized as a negative correlation between antibiotic concentration and the mean fraction of reciprocal sign epistasis (figure 3*b*; $F_{1,10} = 20.3$, $p < 0.005$, adjusted $R^2 = 0.637$). Reciprocal sign epistasis was observed between the weak ABR mutation *rpoB* I572S and most of the knock-out mutations (electronic supplementary material, table S7). Apart from that, the fraction of simple sign epistasis was significant in all environments and showed no trend across environments ($F_{1,10} = 2.15$, $p = 0.17$; electronic supplementary material, figure S23).

Next, we computed the roughness-to-slope ratio. The roughness-to-slope ratio fits a linear model to the fitness landscape and then quantifies epistasis as the ratio of the standard deviation of the residuals to the linear component of the fit (figure 1*d*). Thus, completely additive landscapes exhibit a roughness-to-slope ratio of zero [50]. In our data, the mean estimated roughness-to-slope ratio was negatively correlated with antibiotic concentration (figure 3*c*; $F_{1,178} = 206.9$, $p < 10^{-15}$, adjusted $R^2 = 0.535$). Roughness-to-slope ratios were very small for all genotypes in the presence of the antibiotic; nevertheless the 95% bootstrap confidence intervals were significantly different from zero for all values.

Finally, we inferred whether the studied fitness landscapes exhibited diminishing-returns epistasis. Diminishing-returns epistasis measures whether the effect size of a beneficial mutation decreases as the fitness of the genetic background increases and is one of the most frequently observed patterns of epistasis in fitness landscapes [2]. We observed that the fitness effects of the knock-out mutations were not correlated with the overall fitness of the background, regardless of the environment (figure 4*a*,*c*). By contrast, for most environments, the ABR mutations exhibited a negative correlation with the fitness of the background that the mutation is on, indicating a pattern of diminishing-returns epistasis (figure 4*b*,*c*). The observed trend of diminishing returns for the knock-out and ABR mutants did not depend on the low fitness of the Δ*waaP* mutants. Indeed, when the Δ*waaP* backgrounds were removed from the analysis, the trend became more pronounced as significant diminishing-returns epistasis was observed in all but one of the 12 environments (electronic supplementary material, figure S25).

(e) Competitive fitness is predicted by antibiotic concentration

We used multiple linear regression to study the quantitative effects of antibiotic concentration and temperature environments on the competitive index, $\hat{w}$. We found that the 'null model' regression with $\hat{w}$ as the dependent variable and only additive effects of genotype and environment variables was able to explain a considerable amount of the variation observed in the data (adjusted $R^2 = 0.68$, $F_{13,982} = 162.5$, $p < 10^{-15}$; electronic supplementary material, table S8). Most of the residual variation was attributable to the wild-type *rpoB* genotypes (electronic supplementary material, figure S26). Contrary to resistant genotypes, the antibiotic-susceptible genotypes interacted with the rifampicin environments by having higher fitness in the absence of rifampicin but very low fitness at higher rifampicin concentrations.

Next, we quantified the interaction between the *rpoB* antibiotic susceptible and resistant genotypes with the antibiotic concentration by fitting a model with the additive effects and the $G_{rpoB} \times E_{AB}$ interaction. This model fitted the data better than the additive null model, explaining almost all of the variation (adjusted $R^2 = 0.881$, $F_{9,973} = 186.0$, $p < 10^{-15}$; electronic supplementary material, figure S27). Polynomial contrasts indicated that higher temperature was correlated with increased $\hat{w}$ ($t = 15.0$, $p < 10^{-15}$; electronic supplementary material, figures S28 and S29). The fixed effect of antibiotic concentration alone on $\hat{w}$ was not significant (electronic supplementary material, figure S28). Instead, the effect of antibiotic concentration was only significant as an interaction term with the susceptible *rpoB* wild-type genotype or the weak ABR genotype I572S.

Then, we compared the above model with a slightly more complex model that included an additional $G_{rpoB} \times G_{KO}$ interaction between genotypes (i.e. 'classical' pairwise epistasis between *rpoB* and knock-out mutations),

$$\hat{w} \sim \underbrace{G_{rpoB} + G_{KO} + E_{AB} + E_T}_{\text{additive effects}} + \underbrace{G_{rpoB} \times E_{AB} + G_{rpoB} \times G_{KO}}_{\text{1st order interactions}},$$

Although this model with G × E and G × G effects was significantly different from the model with only G × E effects ($F_{15,958} = 2.60$, $p < 0.001$; electronic supplementary material, figure S30), none of the additional coefficients were significant (electronic supplementary material, figure S31), and the model predictions were very similar to those of the model with only G × E. Indeed, the amount of variation explained by the more complex model (adjusted $R^2 = 0.883$) showed only a 0.2%-point improvement in the total variation explained as compared to the simpler model with only G × E effects. This finding confirms our previous results of small but statistically significant pairwise epistasis in our data.

Despite our relatively large dataset ($n = 996$), we did not have sufficient data to systematically study second-order interactions without overfitting. We fitted models with second-order interactions for descriptive purposes (see electronic supplementary material, results 8.6.3). In brief, we found that there may be a higher-order interaction between epistasis and temperature that our data has insufficient power to reveal statistically.

(f) Environmental quality is less important for predicting gene-by-gene interactions than antibiotic concentration alone

To study the interactions between epistasis and environment (G × G × E) without the problem of overfitting, we summarized the temperature and antibiotic environments as a single, continuous environmental quality metric that we used to regress the different epistasis measures [51]. We compared three metrics of environmental quality (mean growth of the reference strain, mean growth of all competitor strains, and mean combined growth of reference and competitor strains) and found that the mean growth of all competitor strains in each environment was the best metric of environmental quality (electronic supplementary material, results 8.7.1). This metric of environmental quality was able to distinguish environments from each other (electronic supplementary material, figure S32) and correctly distinguish specialist from generalist genotypes (electronic supplementary material, figure S33). As expected,

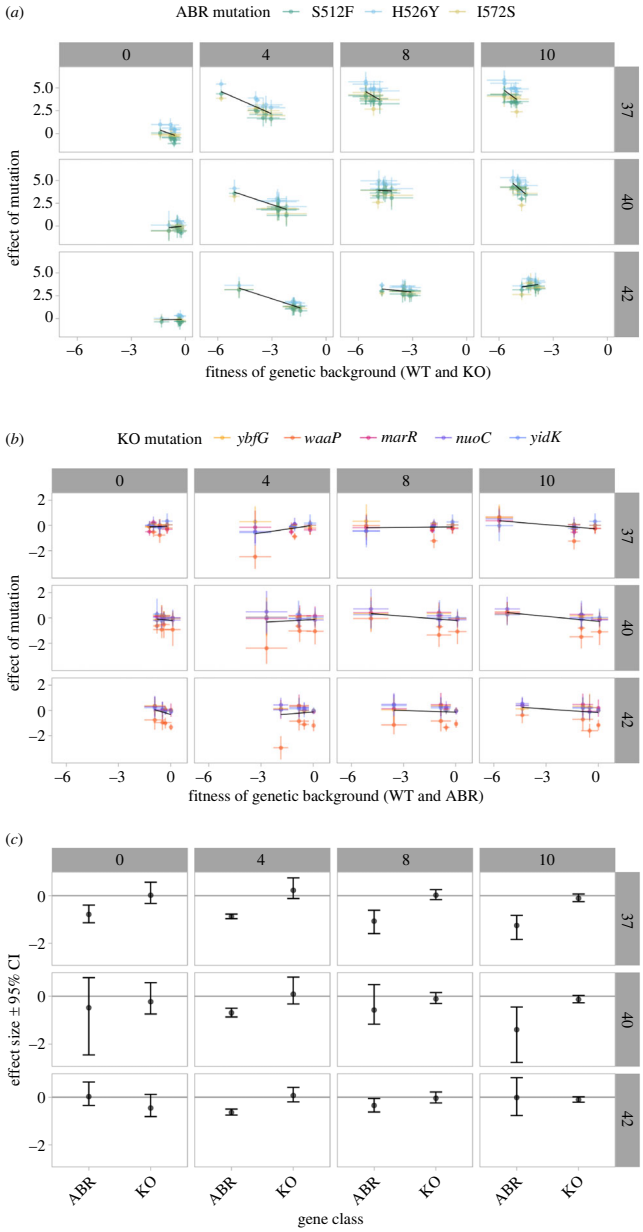


Figure 4. (*Caption overleaf.*)

the environmental quality metric was correlated with antibiotic concentration (adjusted $R^2 = 0.559$) and somewhat correlated with temperature (adjusted $R^2 = 0.325$).

We computed the correlation of environmental quality with all of the above measures of G × G by fitting the environmental quality metric as a linear predictor of pairwise epistasis

Figure 4. (*Overleaf.*) ABR *rpoB* mutations exhibit diminishing-returns epistasis in most environments whereas knock-out mutations never exhibit diminishing-returns epistasis. In each panel, the columns show different rifampicin antibiotic concentrations (in units of μg ml^{-1}) and the rows show different temperature environments (in °C). (*a*) ABR *rpoB* mutations exhibit diminishing-returns epistasis in eight of the 12 environments. The *x*-axis shows the fitness $\hat{w}$ of the background (either the wild type or one of the five knock-out mutations) and the *y*-axis shows the effect size of the ABR mutation on that background. (*b*) Knock-out mutations do not exhibit diminishing-returns epistasis in any of the environments. The *x*-axis shows the fitness $\hat{w}$ of the background (either the wild type or one of the three ABR *rpoB* mutations) and the *y*-axis shows the effect size of the knock-out mutation on that background. For (*a*,*b*), the points show mean values and are coloured corresponding to the mutation; the error bars show 95% parametric bootstrap confidence intervals for each value. The black lines show the means of the regressions across all bootstrap samples. (*c*) Quantification of the mean slope effect sizes (±95% confidence intervals) of the diminishing-returns epistasis for the *rpoB* mutations (ABR) and the knock-out mutations (KO). The mean slopes are shown as black trend lines in (*a*) and (*b*). Confidence intervals are determined from linear regressions of each of the parametric bootstrap samples. The confidence intervals that overlap with zero indicate slopes that are *not* significantly different from zero.

(electronic supplementary material, figures S37 and S38), gamma epistasis (electronic supplementary material, figure S39), the fraction of sign epistasis (electronic supplementary material, figure S40), and the roughness-to-slope ratio (electronic supplementary material, figure S41). Most of the regressions were either not significant to $\alpha = 0.01$ (gamma epistasis for ABR mutations, simple sign epistasis, reciprocal sign epistasis and pairwise epistasis excluding Δ*waaP*) or explained very little of the variation in the data (pairwise epistasis including Δ*waaP*: adjusted $R^2 = 0.036$). For metrics that exhibited a significant effect of environmental quality on G × G (total gamma epistasis, electronic supplementary material, figure S39; and roughness-to-slope ratio, figure S41), antibiotic concentration alone regressed as a dependent variable explained more of the variation in the G × G metric than was explained by environmental quality. Moreover, the results of the regression on environmental quality were broadly in agreement with our above findings that G × G decreases as antibiotic concentration increases and environmental quality deteriorates (i.e. positive correlations were found between the environmental quality metric and the G × G summary statistics). The only exception was the gamma epistasis for knock-out mutations (electronic supplementary material, figure S39, middle facet), which exhibited the opposite trend to other G × G metrics (i.e. G × G was found to increase as environmental quality deteriorated).

3. Discussion

Whether mutational effects and epistasis vary across environments, and how this matters for evolution, has been debated for decades (reviewed in [1,2]). Here, we created 15 small (i.e. two mutational-step) fitness landscapes composed of three single amino acid substitutions at a gene involved in ABR and whole-gene knock-out mutations of five different genes. We screened these fitness landscapes at increasing gradients of temperature and an antibiotic. We found relatively little epistasis for four of the five different epistasis metrics. The benign environments exhibited the most epistasis, and adding antibiotic decreased epistasis (figure 3). ABR mutations, but not gene knock-outs, exhibited diminishing-returns epistasis in most environments (figure 4). These results suggest that, while epistatic interactions may confound predictions in the absence of an antibiotic, the effects of ABR mutations are predictable in the presence of an antibiotic.

(a) Biological significance of our experiment

We combined ABR mutations with gene knock-out mutations because the presence or absence of accessory genes is the most common standing genetic variation in natural populations of *E. coli* [52]. To our knowledge, this is the first study to systematically investigate how G × G interactions between SNPs and gene deletions change across different environments. We selected the gene knock-outs especially to avoid direct, mechanistic interactions with *rpoB* or other genes involved in transcription, protein synthesis or DNA supercoiling. Therefore, our null expectation was that there would not be any G × G interactions and that fitness would be additive throughout. If there was specific epistasis as a result of direct, mechanistic interactions between the ABR mutations and gene knock-outs, we expected to observe masking epistasis where the effect of the complete gene deletion is dominant over all of the ABR mutations. We did observe masking epistasis, but in the opposite direction than expected under specific epistasis: the ABR mutations dominated over the knock-outs. This demonstrates that the effects of single amino acid substitutions at the *rpoB* locus are robust to non-core gene knock-out mutations. This result is an example of non-specific epistasis. Given its large population sizes and diverse pangenome composition, the core genes of *E. coli*, like *rpoB*, may have evolved to tolerate common polymorphisms like gene gain and loss events. We hypothesize that low pairwise epistasis between single amino acid substitutions in core genes and knock-out mutations of non-core genes may generally be expected for *E. coli* and other prokaryotes with large pangenomes [53].

We studied ABR mutations at the *rpoB* locus because this locus is clinically important and was shown to exhibit G × G interactions. Amino acid mutations at the *rpoB* locus (e.g. at site H526 included in our study) are the main, evolutionarily conserved mechanism of rifampicin resistance relevant for many bacterial pathogens [54,55]. This includes the clinically relevant *Mycobacterium tuberculosis*, against which rifampicin treatment remains an important first-line antibiotic [54]. Previous studies have found that the fitness of *rpoB* mutants depends on the presence of mutations in other mechanistically interacting proteins, like other subunits of the RNA polymerase complex [56] and other protein complexes involved in protein synthesis and DNA stability [31,57]. This type of 'specific epistasis' is easily explained by the close physical and functional proximity of these proteins to *rpoB*. 'Non-specific epistasis', however, has been observed between genes and proteins that do not interact directly [26]. *RpoB* is an excellent candidate to investigate non-specific epistasis as it was already shown to be highly pleiotropic: *rpoB* mutations impact the expression of hundreds of genes and many cell phenotypes [36,58–60]. The demonstrated functional effect of *rpoB* upon many other genes implies that mutations at other loci could impact the fitness of ABR *rpoB* mutations. Therefore, the quantification of G × G interactions at the *rpoB* locus is important to know how the danger of ABR depends on the genetic background.

We detected a signal of pairwise epistasis in only one of five gene knock-outs, Δ*waaP* (electronic supplementary material, figure S21). We had expected Δ*marR* to be the most likely knock-out mutation to exhibit pairwise epistasis, given its functional role in responding to chemical stressors and antibiotics [42,43,61] and its previously observed epistatic interactions with ABR mutations at genes involved in DNA stability [62]. Δ*marR* did not display significant pairwise epistasis in any environment; only Δ*waaP* did. The Δ*waaP* genotype on the susceptible *rpoB* wild-type background was more sensitive to low antibiotic concentration than the wild type and other knock-out single mutants, but this sensitivity was masked in combination with any of the three ABR mutations. Gene deletions of other lipopolysaccharide biosynthesis genes were shown to exhibit negative pairwise epistasis with ABR mutations [63]. Δ*waaP* could make cells more sensitive to rifampicin by increasing the permeability of the outer membrane [64,65]. Among the four functional genes studied, the *waaP* gene has the highest gene diversity ($\pi_{waaP} = 0.034$ versus $\pi_{nuoC} = 0.015$, $\pi_{marR} = 0.014$, $\pi_{yidK} = 0.013$) and is most frequently found knocked out (*waaP* is present in 78.5% of strains as compared to 99.5%, 99.3% and 90.8% for *nuoC*, *marR* and *yidK*, respectively) in natural populations, according to the panX database. Moreover, *waaP* impacts bacterial virulence and, so, is relevant in host–pathogen interactions [66]. *E. coli* strains from which *waaP* is absent may be more sensitive to low doses of rifampicin and other antibiotics than *waaP*-carrying genotypes. Our observed pattern of epistasis implies that this sensitivity is masked when the genotype carries a resistance mutation. Thus, the G × G interaction of *waaP* and *rpoB* erases selection against Δ*waaP* mutations that would occur at low doses of rifampicin.

(b) Measurement challenges of gene-by-gene interactions

One methodological challenge for detecting G × G interactions in our study was that, in the presence of antibiotics, the fitness effects of the ABR mutants quickly overshadowed those of the knock-outs. In this case, methods that involve a relative comparison of direct fitness effects to epistatic effects are bound to infer less epistasis when there are stronger direct effects. The negative correlation of the roughness-to-slope ratio as a function of antibiotic concentration could be explained by this phenomenon: we observed mostly neutral fitness effects and moderate epistasis in the absence of the antibiotic but large beneficial fitness effects and no epistasis in the presence of the antibiotic. However, the negative correlation between antibiotic concentration and epistasis was also exhibited by other epistasis summary statistics that are independent of the size of direct fitness effects (i.e. gamma epistasis and the mean fraction of reciprocal sign epistasis, figure 3). Therefore, the observed decrease of epistasis with antibiotic concentration is likely a real biological phenomenon.

Our estimates of gamma epistasis for different gene classes suggest that the fitness landscape as a whole was less epistatic than when only ABR mutations or knock-out mutations were considered. This could reflect a biological pattern (as supported by the low pairwise epistasis results) where the epistatic interactions between amino acid mutations and gene knock-outs are smaller than within those mutational classes, perhaps due to evolved mutational robustness for gene knock-outs in *E. coli*. Nevertheless, the different trends observed between the gamma epistasis of the different mutational classes are possibly impacted by the differences in sample size. The gamma statistic relies on averaging over multiple mutational effects, which leads to a dilution of the epistatic signal for larger fitness landscapes. The total dataset has a larger sample size than subsets of the data, and the knock-out mutational class has a larger sample size than the ABR class. Future work should explore how the gamma epistasis measure can be used more appropriately for comparing epistasis between gene classes and differently sized fitness landscapes, for example, by randomly subsampling the data to an equal size for all categories.

(c) Environmental interactions and their importance

It has long been established that there is a G × E interaction between rifampicin resistance mutations at the *rpoB* locus and temperature [67], among other environmental factors [18,68,69]. In a previous experiment, Rodríguez-Verdugo *et al*. [36] observed that *E. coli* adapted to high temperature evolved rifampicin resistance despite no rifampicin treatment. Therefore, we hypothesized that there would be an E × E interaction between high-temperature environments and rifampicin environments, which, if mediated by ABR mutations at the *rpoB* locus, could result in $G_{rpoB} \times E_{AB} \times E_T$ interactions as well. Overall we observed no cost of rifampicin resistance in minimal media, which is consistent with Lin *et al*. [70] and could be attributed to ABR *rpoB* mutations mimicking the stringent response [56]. There was a modest positive effect of high temperatures on the competitive fitness of all genotypes and a weak E × E interaction between temperature and antibiotic. However, contrary to previous studies that also used batch cultures with glucose and minimal medium [7,36], the *rpoB* mutations S512F and I572S did not grow better and H526Y did not grow worse than the wild type at higher temperatures. (In fact, H526Y was observed to grow *better* than the wild type at 42°C.) The discrepancy for I572S could be attributed to G × G × E with the genetic background: Rodríguez-Verdugo *et al*. [36] found a strong effect of *E. coli* genetic background on the fitness at site I572. On the other hand, the discrepancies for S512F and H526Y could be that, unlike in Trindade *et al*. [7], our batch cultures were not acclimatized to the environment in which the competitions occurred. This underscores the sensitivity of organisms to their environments and the importance of complete reporting of experimental methods.

We used the overall environmental quality as a regressor to quantify the interaction between epistasis and environment, G × G × E. We found that the antibiotic concentration alone was a better regressor of epistasis than the environmental quality, implying that temperature did not have an effect on epistasis. We had expected that the overall environmental quality would be a better regressor because it is highly correlated with antibiotic concentration while also taking into account the effects of temperature. On the other hand, fitting of linear models with higher-order interactions suggested that temperature exhibited a stronger interaction with $G_{rpoB} \times G_{KO}$ than antibiotic concentration. Unfortunately, we have limited confidence in the linear regression results due to the low sample size and the high order of the investigated interactions. However, this apparent contradiction raises new questions about G × G × E interactions: for example, is it possible that one environmental variable affects direct G × E interactions,

whereas the other acts on the G × G × E level? We are not aware of any previous work that has extracted such systematic patterns empirically, inferred their functional underpinnings, or developed models to account for such interaction. Certainly, the consequences of different interaction levels will be important to study in the future to predict potential evolutionary trajectories in varying environments.

(d) Why is there so little epistasis in the studied antibiotic resistance fitness landscapes?

Overall, we found little G × G and G × G × E interaction in our data. This result is in stark contrast to various recent experimental studies, which have presented strong evidence that both fitness effects of single mutations and their epistatic fitness interactions may vary greatly between environments ([9,10,19], reviewed in [2]). One reason for this discrepancy may be the choice of mutations for the fitness landscapes. Previous studies focused on SNPs between or within genes that had been found in an adaptive walk [4,5,11], indirectly inferred strong G × G(×E) by observing that adaptations were unique to the genetic background [21,30,71], or measured epistasis as the different effect of ABR mutations across vastly different genetic backgrounds [72,73]. We here quantified epistatic interactions between ABR SNPs and (non-ABR) gene knock-out mutations that occur in natural populations of *E. coli*. The studied combinations of mutations thus had no immediate relationship except the *rpoB* mutations' previously characterized G × E interaction for fitness under higher temperature or antibiotic. However, multiple studies in systems or molecular biology have shown that epistasis is common (reviewed in [26]). In particular, when two or more beneficial mutations are combined, negative epistasis in the shape of diminishing returns (i.e. where the fitness effect of a mutation is less beneficial when it occurs on a fitter background) has been observed ubiquitously [28,29,63,74], including for *rpoB* mutations conferring ABR to rifampicin [75]. In our study, we observed diminishing-returns epistasis only for the *rpoB* ABR mutations but no diminishing-returns epistasis for the knock-out mutations. Moreover, we did not observe any qualitative environmental trend in the strength of the diminishing-returns epistasis for the ABR mutations. However, the mutations shifted from neutral to beneficial as a function of antibiotic. Although there have been studies that derived null expectations of epistasis between random mutations of the same type [76], between mutations characterized by their fitness effect [77], or between SNPs in a pair of genes in a metabolic pathway [78], we are not aware of any general models or empirical studies that have proposed or quantified a null distribution of epistatic effects between SNPs and different structural mutation types (except for the assumption of no epistasis). Therefore, it is not clear whether the low G × G and G × G × E interactions observed in our data are to be expected.

Despite a lack of general epistatic null models, theoretical and empirical works have proposed a few hypotheses on how much epistasis to expect between mutations that confer ABR. For example, Mira *et al.* [34] found ubiquitous sign epistasis, in all 30 environments assayed, for a fitness landscape with all possible combinations of four ABR mutations. Knopp & Andersson [31] found that epistasis was rare between combinations of ABR mutations. Our finding that ∼20% of randomly selected combinations of mutations exhibit pairwise epistasis is in good quantitative agreement with the results of Knopp & Andersson [31]. A critical difference between our work and most previous work on ABR epistasis is that the knock-out mutations we investigated neither interacted directly with *rpoB* (as required for specific epistasis) nor, except Δ*waaP*, displayed any significant effects on fitness (electronic supplementary material, figure S28). Epistasis expressed by seemingly neutral mutations is termed 'cryptic epistasis', and has been previously observed between mutations that were fixed during an adaptive walk [3,5]. We have uncovered new instances of cryptic epistasis that depend on the environment. Using a theoretical model, Engelstädter [79] related the expected epistasis to the extent of G × E interactions of the involved mutations: they proposed that when there is no cost to the ABR mutations, there should be no epistasis between them. Moreover, a study by Das *et al.* [13], considering only ABR mutations with a cost of resistance predicted using mathematical models and empirically, observed the strongest epistasis at intermediate antibiotic concentrations. Our experiment screened a range of intermediate antibiotic concentrations, with 4 µg ml^{-1} below and 10 µg ml^{-1} near the minimum inhibitory concentration (MIC) of the wild-type *rpoB* (electronic supplementary material, figure S14), and ABR mutations without a cost of resistance (although a cost had been expected according to previous studies, see above) in combination with non-ABR mutations. Our finding of very little epistasis in the presence of antibiotic is thus more consistent with the prediction of Engelstädter [79].

(e) Do fitness landscapes become smoother as the concentration of an environmental stressor increases?

How generalizable are our findings that fitness landscapes become smoother in more stressful environments? To discuss this question, we compare our findings with those of Gorter *et al*. [8], who measured fitness under different cadmium concentrations for all combinations of knock-out mutations that had evolved in response to those toxic environments. Contrary to our results, their co-selected mutations conferring resistance to increasing cadmium exhibited increasingly strong selective effects and positive pairwise epistatic effects as the heavy-metal concentration increased [8]. Interestingly, both our results and those of Gorter *et al*. [8] contradict theoretical predictions. According to Fisher's geometric model [80], the average pairwise epistasis should be unchanged with environmental stress for random mutations, like the combinations of mutations used in our study, whereas it should decrease for co-selected mutations, like those studied in Gorter *et al*. [8]. Also, our results and those of Gorter *et al*. [8] differ from the predicted and empirical results of Das *et al*. [13] that epistasis should be strongest at an intermediate concentration of an antibiotic stressor. However, the model of Das *et al*. [13] was specific to increasing antibiotic (and not cadmium) stress and assumed a cost of resistance which was not observed in our study or in Gorter *et al*. [8]. These conflicting results and predictions call for the study of additional empirical fitness landscapes under increasing concentrations of an environmental stressor, and new theoretical fitness landscape models that incorporate mechanistic details of biological phenomena.

Our main conclusion is that epistasis decreases with increasing antibiotic concentration. Only in the absence of the antibiotic, we observed several instances of reciprocal sign epistasis for all three ABR mutations and at all temperature environments (figure 3*b*; electronic supplementary material, table S7). The presence of reciprocal sign epistasis implies that the fitness landscape has multiple peaks and that evolution may be less predictable in the absence of antibiotic. Overall, our results are consistent with the conclusion that the underlying fitness landscapes of ABR mutations and gene knock-outs are more rugged in the absence of antibiotic than in the presence of higher concentrations of antibiotic. Extrapolated to the larger sequence space, our results would imply that evolution is more predictable in the presence than in the absence of antibiotic, because ABR mutations have such strong beneficial effects in the presence of high antibiotic concentrations. Here, the strong effect of the ABR mutations potentially overrides any effects of the genetic background.

Data accessibility. All flow cytometry data, competitive fitness estimates and the annotated code used to generate the analyses are publicly available at the following Git repository: https://gitlab.com/evoldynamics/epistasis-decreases-with-increasing-antibiotic-pressure and is archived on Zenodo with the following doi:10.5281/zenodo.7661199. The whole-genome sequencing is archived on NCBI with the BioProject accession: PRJNA910115. The data are provided in electronic supplementary material [81].

Authors' contributions. A.-H.G.: conceptualization, funding acquisition, investigation, methodology, validation, visualization, writing—original draft, writing—review and editing; A.A.: methodology, visualization, writing—review and editing; I.G.: conceptualization, funding acquisition, investigation, methodology, resources, supervision, writing—review and editing; C.B.: conceptualization, funding acquisition, investigation, methodology, supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. A.-H.G. acknowledges funding from FCT PhD funding grant no. PD/BD/138215/2018. C.B. acknowledges funding from ERC Starting grant no. 804569 (FIT2GO), HFSP Young Investigator grant no. RGY0081/2020 and SNSF Project grant no. 315230_204838/1 (MiCo4Sys). I.G. and C.B. acknowledge funding from FCT PREPARE project (JPIAMR/0001/2016-ERA-NET).

Acknowledgements. We thank the Flow Cytometry Facility of Instituto Gulbenkian de Ciência and the Next Generation Sequencing (NGS) Platform of UBern for their services and assistance. We are grateful to the members of the Evolutionary Biology group for *E. coli* strains (R. Balbontín and P. Durão), training in microbiology (D. Guleresi) and consultation on methods, and to the members of the Evolutionary Dynamics/Theoretical Ecology and Evolution group for help with preliminary environmental screenings (C. Diwo), genetic engineering (M. Schmitz), consultation on WGS (A. Kapopoulou) and discussions. We thank T. Batallion, S. Yeaman, F. Blanquart and A. Wagner for feedback on results; and the PREPARE consortium for ideas and help with experimental design.

References

1. Whitlock MC, Phillips PC, Moore FBG, Tonsor SJ. 1995 Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* **26**, 601–629. (doi:10.1146/annurev.es.26.110195.003125)
2. Bank C. 2022 Epistasis and adaptation on fitness landscapes. *Annu. Rev. Ecol. Evol. Syst.* **53**, 457–479. (doi:10.1146/annurev-ecolsys-102320-112153)
3. Lunzer M, Golding GB, Dean AM. 2010 Pervasive cryptic epistasis in molecular evolution. *PLOS Genet.* **6**, e1001162. (doi:10.1371/journal.pgen.1001162)
4. Starr TN, Flynn JM, Mishra P, Bolon DNA, Thornton JW. 2018 Pervasive contingency and entrenchment in a billion years of Hsp90 evolution. *Proc. Natl Acad. Sci. USA* **115**, 4453–4458. (doi:10.1073/pnas.1718133115)
5. Park Y, Metzger BPH, Thornton JW. 2022 Epistatic drift causes gradual decay of predictability in protein evolution. *Science* **376**, 823–830. (doi:10.1126/science.abn6895)
6. Tenaillon O, Rodríguez-Verdugo A, Gaut RL, McDonald P, Bennett AF, Long AD, Gaut BS. 2012 The molecular diversity of adaptive convergence. *Science* **335**, 457–461. (doi:10.1126/science.1212986)
7. Trindade S, Sousa A, Gordo I. 2012 Antibiotic resistance and stress in the light of Fisher's model. *Evolution* **66**, 3815–3824. (doi:10.1111/j.1558-5646.2012.01722.x)
8. Gorter FA, Aarts MGM, Zwaan BJ, de Visser JAGM. 2018 Local fitness landscapes predict yeast evolutionary dynamics in directionally changing environments. *Genetics* **208**, 307–322. (doi:10.1534/genetics.117.300519)
9. Flynn JM, Rossouw A, Cote-Hammarlof P, Fragata I, Mavor D, Hollins III C, Bank C, Bolon DNA. 2020 Comprehensive fitness maps of Hsp90 show widespread environmental dependence. *eLife* **9**, e53810. (doi:10.7554/eLife.53810)
10. Flynn KM, Cooper TF, Moore FBG, Cooper VS. 2013 The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. *PLoS Genet.* **9**, e1003426. (doi:10.1371/journal.pgen.1003426)
11. Hall AE, Karkare K, Cooper VS, Bank C, Cooper TF, Moore FBG. 2019 Environment changes epistasis to alter trade-offs along alternative evolutionary paths. *Evolution* **73**, 2094–2105. (doi:10.1111/evo.13825)
12. Ono J, Gerstein AC, Otto SP. 2017 Widespread genetic incompatibilities between first-step mutations during parallel adaptation of *Saccharomyces cerevisiae* to a common environment. *PLoS Biol.* **15**, e1002591. (doi:10.1371/journal.pbio.1002591)
13. Das SG, Direito SOL, Waclaw B, Allen RJ, Krug J. 2020 Predictable properties of fitness landscapes induced by adaptational tradeoffs. *eLife* **9**, e55155. (doi:10.7554/eLife.55155)
14. Lässig M, Mustonen V, Walczak AM. 2017 Predicting evolution. *Nat. Ecol. Evol.* **1**, 0077. (doi:10.1038/s41559-017-0077)
15. de Visser JAGM, Krug J. 2014 Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490. (doi:10.1038/nrg3744)
16. Melnyk AH, Wong A, Kassen R. 2015 The fitness costs of antibiotic resistance mutations. *Evol. Appl.* **8**, 273–283. (doi:10.1111/eva.12196)
17. Lenormand T, Harmand N. 2018 Cost of resistance: an unreasonably expensive concept. *Rethinking Ecol.* **3**, 51–70. (doi:10.3897/rethinkingecology.3.31992)
18. Knöppel A, Näsvall J, Andersson DI. 2017 Evolution of antibiotic resistance without antibiotic exposure. *Antimicrob. Agents Chemother.* **61**, e01495-17. (doi:10.1128/AAC.01495-17)
19. Guerrero RF, Scarpino SV, Rodrigues JV, Hartl DL, Ogbunugafor CB. 2019 Proteostasis environment shapes higher-order epistasis operating on antibiotic resistance. *Genetics* **212**, 565–575. (doi:10.1534/genetics.119.302138)
20. Pinheiro F, Warsi O, Andersson DI, Lässig M. 2021 Metabolic fitness landscapes predict the evolution of antibiotic resistance. *Nat. Ecol. Evol.* **5**, 677–687. (doi:10.1038/s41559-021-01397-0)
21. Hernando-Amado S, Laborda P, Valverde JR, Martínez JL. 2022 Mutational background influences *P. aeruginosa* ciprofloxacin resistance evolution but preserves collateral sensitivity robustness. *Proc. Natl Acad. Sci. USA* **119**, e2109370119. (doi:10.1073/pnas.2109370119)
22. Aminov RI. 2010 A brief history of the antibiotic era: lessons learned and challenges for the future. *Front. Microbiol.* **1**, 134. (doi:10.3389/fmicb.2010.00134)
23. Murray CJL *et al.* 2022 Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655. (doi:10.1016/S0140-6736(21)02724-0)

24. Lukačišinová M, Bollenbach T. 2017 Toward a quantitative understanding of antibiotic resistance evolution. *Curr. Opin. Biotechnol.* **46**, 90–97. (doi:10.1016/j.copbio.2017.02.013)
25. McEwen SA, Collignon PJ. 2018 Antimicrobial resistance: a one health perspective. *Microbiol. Spectr.* **6**, 10. (doi:10.1128/microbiolspec.ARBA-0009-2017)
26. Domingo J, Baeza-Centurion P, Lehner B. 2019 The causes and consequences of genetic interactions (epistasis). *Annu. Rev. Genom. Hum. Genet.* **20**, 17.1–17.28. (doi:10.1146/annurev-genom-083118-014857)
27. Li C, Zhang J. 2018 Multi-environment fitness landscapes of a tRNA gene. *Nat. Ecol. Evol.* **2**, 1025–1032. (doi:10.1038/s41559-018-0549-8)
28. Reddy G, Desai MM. 2021 Global epistasis emerges from a generic model of a complex trait. *eLife* **10**, e64740. (doi:10.7554/eLife.64740)
29. Perfeito L, Sousa A, Bataillon T, Gordo I. 2014 Rates of fitness decline and rebound suggest pervasive epistasis. *Evolution* **68**, 150–162. (doi:10.1111/evo.12234)
30. Nyerges Á *et al.* 2018 Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. *Proc. Natl Acad. Sci. USA* **115**, E5726–E5735. (doi:10.1073/pnas.1801646115)
31. Knopp M, Andersson DI. 2018 Predictable phenotypes of antibiotic resistance mutations. *mBio* **9**, e00770-18. (doi:10.1128/mBio.00770-18)
32. Mira PM, Østman B, Guzman-Cole C, Sindi S, Barlow M. 2021 Adaptive processes change as multiple functions evolve. *Antimicrob. Agents Chemother.* **65**, e01990-20. (doi:10.1128/AAC.01990-20)
33. Fragata I, Blanckaert A, Louro MAD, Liberles DA, Bank C. 2019 Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82. (doi:10.1016/j.tree.2018.10.009)
34. Mira PM, Meza JC, Nandipati A, Barlow M. 2015 Adaptive landscapes of resistance genes change as antibiotic concentrations change. *Mol. Biol. Evol.* **32**, 2707–2715. (doi:10.1093/molbev/msv146)
35. Durão P, Trindade S, Sousa A, Gordo I. 2015 Multiple resistance at no cost: rifampicin and streptomycin a dangerous liaison in the spread of antibiotic resistance. *Mol. Biol. Evol.* **32**, 2675–2680. (doi:10.1093/molbev/msv143)
36. Rodríguez-Verdugo A, Gaut BS, Tenaillon O. 2013 Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress. *BMC Evol. Biol.* **13**, 50. (doi:10.1186/1471-2148-13-50)
37. Nichols RJ *et al.* 2011 Phenotypic landscape of a bacterial cell. *Cell* **144**, 143–156. (doi:10.1016/j.cell.2010.11.052)
38. Kritikos G, Banzhaf M, Herrera-Dominguez L, Koumoutsi A, Wartel M, Zietek M, Typas A. 2017 A tool named Iris for versatile high-throughput phenotyping in microorganisms. *Nat. Microbiol.* **2**, 17014. (doi:10.1038/nmicrobiol.2017.14)
39. Baba T *et al.* 2006 Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 20060.008. (doi:10.1038/msb4100050)
40. Ding W, Baumdicker F, Neher RA. 2018 panX: pan-genome analysis and exploration. *Nucleic Acids Res.* **46**, e5. (doi:10.1093/nar/gkx977)
41. Krisko A, Copic T, Gabaldón T, Lehner B, Supek F. 2014 Inferring gene function from evolutionary change in signatures of translation efficiency. *Genome Biol.* **15**, R44. (doi:10.1186/gb-2014-15-3-r44)
42. Praski Alzrigat L, Huseby DL, Brandis G, Hughes D. 2021 Resistance/fitness trade-off is a barrier to the evolution of MarR inactivation mutants in *Escherichia coli*. *J. Antimicrob. Chemother.* **76**, 77–83. (doi:10.1093/jac/dkaa417)
43. Reyes-Fernández EZ, Schuldiner S. 2020 Acidification of cytoplasm in *Escherichia coli* provides a strategy to cope with stress and facilitates development of antibiotic resistance. *Sci. Rep.* **10**, 9954. (doi:10.1038/s41598-020-66890-1)
44. Friedrich T. 1998 The NADH:ubiquinone oxidoreductase (complex I) from *Escherichia coli*. *Biochim. Biophys. Acta* **1364**, 134–146. (doi:10.1016/S0005-2728(98)00024-3)
45. Daley DO, Rapp M, Granseth E, Melén K, Drew D, von Heijne G. 2005 Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308**, 1321–1323. (doi:10.1126/science.1109730)
46. Yethon JA, Whitfield C. 2001 Purification and characterization of WaaP from *Escherichia coli*, a lipopolysaccharide kinase essential for outer membrane stability. *J. Biol. Chem.* **276**, P5498–5504. (doi:10.1074/jbc.M008255200)
47. Moura de Sousa J, Balbontín R, Durão P, Gordo I. 2017 Multidrug-resistant bacteria compensate for the epistasis between resistances. *PLoS Biol.* **15**, e2001741. (doi:10.1371/journal.pbio.2001741)
48. Trindade S, Sousa A, Xavier KB, Dionisio F, Godinho Ferreira M, Gordo I. 2009 Positive epistasis drives the acquisition of multidrug resistance. *PLoS Genet.* **5**, e1000578. (doi:10.1371/journal.pgen.1000578)
49. Ferretti L, Schmiegelt B, Weinreich D, Yamauchi A, Kobayashi Y, Tajima F, Achaz G. 2016 Measuring epistasis in fitness landscapes: the correlation of fitness effects of mutations. *J. Theor. Biol.* **396**, 132–143. (doi:10.1016/j.jtbi.2016.01.037)
50. Szendro IG, Schenk MF, Franke J, Krug J, de Visser JAGM. 2013 Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech. Theory Exp.* **2013**, P01005. (doi:10.1088/1742-5468/2013/01/P01005)
51. Malosetti M, Ribaut JM, van Eeuwijk FA. 2013 The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* **4**, 44. (doi:10.3389/fphys.2013.00044)
52. Touchon M *et al.* 2009 Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet.* **5**, e1000344. (doi:10.1371/journal.pgen.1000344)
53. Brockhurst MA, Harrison E, Hall JPJ, Richards T, McNally A, MacLean C. 2019 The ecology and evolution of pangenomes. *Curr. Biol.* **29**, R1094–R1103. (doi:10.1016/j.cub.2019.08.012)
54. Goldstein BP. 2014 Resistance to rifampicin: a review. *J. Antibiot.* **67**, 625–630. (doi:10.1038/ja.2014.107)
55. Cutugno L, Mc Cafferty J, Pané-Farré J, O'Byrne C, Boyd A. 2020 *rpoB* mutations conferring rifampicin-resistance affect growth, stress response and motility in *Vibrio vulnificus*. *Microbiology* **166**, 1160–1170. (doi:10.1099/mic.0.000991)
56. Koch A, Mizrahi V, Warner DF. 2014 The impact of drug resistance on *Mycobacterium tuberculosis* physiology: what can we learn from rifampicin?. *Emerg. Microbes Infect.* **3**, e17. (doi:10.1038/emi.2014.17)
57. Balbontín R, Frazão N, Gordo I. 2021 DNA breaks-mediated fitness cost reveals RNase HI as a new target for selectively eliminating antibiotic-resistant bacteria. *Mol. Biol. Evol.* **38**, 3220–3234. (doi:10.1093/molbev/msab093)
58. Hug SM, Gaut BS. 2015 The phenotypic signature of adaptation to thermal stress in *Escherichia coli*. *BMC Evol. Biol.* **15**, 177. (doi:10.1186/s12862-015-0457-3)
59. Pietsch F, Bergman JM, Brandis G, Marcusson LL, Zorzet A, Huseby DL, Hughes D. 2016 Ciprofloxacin selects for RNA polymerase mutations with pleiotropic antibiotic resistance effects. *J. Antimicrob. Chemother.* **72**, 75–84. (doi:10.1093/jac/dkw364)
60. Rodríguez-Verdugo A, Tenaillon O, Gaut BS. 2016 First-step mutations during adaptation restore the expression of hundreds of genes. *Mol. Biol. Evol.* **33**, 25–39. (doi:10.1093/molbev/msv228)
61. Deochand DK, Grove A. 2017 MarR family transcription factors: dynamic variations on a common scaffold. *Crit. Rev. Biochem. Mol. Biol.* **52**, 595–613. (doi:10.1080/10409238.2017.1344612)
62. Marcusson LL, Frimodt-Møller N, Hughes D. 2009 Interplay in the selection of fluoroquinolone resistance and bacterial fitness. *PLoS Pathog.* **5**, e1000541. (doi:10.1371/journal.ppat.1000541)
63. Lukačišinová M, Fernando B, Bollenbach T. 2020 Highly parallel lab evolution reveals that epistasis can curb the evolution of antibiotic resistance. *Nat. Commun.* **11**, 3105. (doi:10.1038/s41467-020-16932-z)
64. Wang Z, Wang J, Ren G, Li Y, Wang X. 2015 Influence of core oligosaccharide of lipopolysaccharide to outer membrane behavior of *Escherichia coli*. *Mar. Drugs* **13**, 3325–3339. (doi:10.3390/md13063325)
65. Xu G, Liu H, Jia X, Wang X, Xu P. 2021 Mechanisms and detection methods of *Mycobacterium tuberculosis* rifampicin resistance: the phenomenon of drug resistance is complex. *Tuberculosis* **128**, 102083. (doi:10.1016/j.tube.2021.102083)
66. Yethon JA, Gunn JS, Ernst RK, Miller SI, Laroche L, Malo D, Whitfield C. 2000 *Salmonella enterica* serovar typhimurium *waaP* mutants show increased susceptibility to polymyxin and loss of virulence in vivo. *Infect. Immun.* **68**, 4485–4491. (doi:10.1128/IAI.68.8.4485-4491.2000)

67. Jin DJ, Gross CA. 1989 Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*. *J. Bacteriol.* **171**, 5229–5231. (doi:10.1128/jb.171.9.5229-5231.1989)
68. Hall AR. 2013 Genotype-by-environment interactions due to antibiotic resistance and adaptation in *Escherichia coli*. *J. Evol. Biol.* **26**, 1655–1664. (doi:10.1111/jeb.12172)
69. Zorraquino V, Kim M, Rai N, Tagkopoulos I. 2017 The genetic and transcriptional basis of short and long term adaptation across multiple stresses in *Escherichia coli*. *Mol. Biol. Evol.* **34**, 707–717. (doi:10.1093/molbev/msw269)
70. Lin W, Zeng J, Wan K, Lv L, Guo L, Li X, Yu X. 2018 Reduction of the fitness cost of antibiotic resistance caused by chromosomal mutations under poor nutrient conditions. *Environ. Int.* **120**, 63–71. (doi:10.1016/j.envint.2018.07.035)
71. Hoeksema M, Jonker MJ, Brul S, Ter Kuile BH. 2019 Effects of a previously selected antibiotic resistance on mutations acquired during development of a second resistance in *Escherichia coli*. *BMC Genom.* **20**, 284. (doi:10.1186/s12864-019-5648-7)
72. Vogwill T, Kojadinovic M, MacLean RC. 2016 Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of *Pseudomonas*. *Proc. R. Soc. B* **283**, 20160151. (doi:10.1098/rspb.2016.0151)
73. Apjok G, Boross G, Lázár V, Papp B, Pál C, Csörgö B. 2019 Limited evolutionary conservation of the phenotypic effects of antibiotic resistance mutations. *Mol. Biol. Evol.* **36**, 1601–1611. (doi:10.1093/molbev/msz109)
74. Khan AI, Dinh DM, Schneider D, Lenski RE, Cooper TF. 2011 Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* **332**, 1193–1196. (doi:10.1126/science.1203801)
75. MacLean RC, Perron GG, Gardner A. 2010 Diminishing returns from beneficial mutations and pervasive epistasis shape the fitness landscape for rifampicin resistance in *Pseudomonas aeruginosa*. *Genetics* **186**, 1345–1354. (doi:10.1534/genetics.110.123083)
76. Martin G, Elena SF, Lenormand T. 2007 Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat. Genet.* **39**, 555–560. (doi:10.1038/ng1998)
77. Martin G, Lenormand T. 2015 The fitness effect of mutations across environments: Fisher's geometrical model with multiple optima. *Evolution* **69**, 1433–1447. (doi:10.1111/evo.12671)
78. Kemble H, Eisenhauer C, Couce A, Chapron A, Magnan M, Gautier G, Le Nagard H, Nghe P, Tenaillon O. 2020 Flux, toxicity, and expression costs generate complex genetic interactions in a metabolic pathway. *Sci. Adv.* **6**, eabb2236. (doi:10.1126/sciadv.abb2236)
79. Engelstädter J. 2014 Fitness landscapes emerging from pharmacodynamic functions in the evolution of multidrug resistance. *J. Evol. Biol.* **27**, 840–853. (doi:10.1111/jeb.12355)
80. Blanquart F, Achaz G, Bataillon T, Tenaillon O. 2014 Properties of selected mutations and genotypic landscapes under Fisher's geometric model. *Evolution* **68**, 3537–3554. (doi:10.1111/evo.12545)
81. Ghenu AH, Amado A, Gordo I, Bank C. 2023 Epistasis decreases with increasing antibiotic pressure but not temperature. Figshare. (doi:10.6084/m9.figshare.c.6440214)

Chapter 4

Challenges and pitfalls of inferring microbial growth rates from lab cultures

All models are approximations. Assumptions, whether implied or clearly stated, are never exactly true. All models are wrong, but some models are useful. *So the question you need to ask is not "Is the model true?" (it never is) but "Is the model good enough for this particular application?"*
– George E. P. Box, Alberto Luceño, and Maria del Carmen Paniagua-Quiñones

This chapter has been submitted for review to many journals and has been desk rejected by all except *PLOS One*, who rejected it after peer review. At time of thesis submission, the paper is under consideration at *Frontiers in Ecology and Evolution*. The student's contributions to the work are as follows: The student initiated the study and then performed some initial simulations and analyses. Loïc Marrec performed all the numerical and analytical work featured in the paper. The student performed the literature review. All authors analyzed and interpreted the data, then wrote and edited the manuscript. The student and Loïc Marrec together are performing the tasks of the corresponding author.

Challenges and pitfalls of inferring microbial growth rates from lab cultures

Ana-Hermina Ghenu*[1,2,3], Loïc Marrec†[2,3], and Claudia Bank[1,2,3]

[1]Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras, 2780-156, Portugal
[2]Institut für Ökologie und Evolution, Universität Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland
[3]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

October 9, 2023

Abstract

After more than 100 years of generating monoculture batch culture growth curves, microbial ecologists and evolutionary biologists still lack a reference method for inferring growth rates. Our work highlights the challenges of estimating the growth rate from growth curve data and shows that inaccurate estimates of growth rates significantly impact the estimated relative fitness, a principal quantity in evolution and ecology. First, we conducted a literature review and found which different types of methods are currently used to estimate growth rates. These methods differ in the meaning of the estimated growth rate parameter. Mechanistic models estimate the intrinsic growth rate μ whereas phenomenological methods – both model-based and model-free – estimate the maximum *per capita* growth rate $\mu_{\max}$. Using math and simulations, we show the conditions in which $\mu_{\max}$ is not a good estimator of μ. Then, we demonstrate that inaccurate absolute estimates of μ is not overcome by calculating relative values. Importantly, we find that poor approximations for μ sometimes lead to wrongly classifying a beneficial mutant as deleterious. Finally, we re-analyzed four published data-sets using most of the methods found by our literature review. We detected no single best-fitting model across all experiments within a data-set and found that the Gompertz models, which were among the most commonly used, were often among the worst fitting. Our study provides suggestions for how experimenters can improve their growth rate and associated relative fitness estimates and highlights a neglected but fundamental problem for nearly everyone who studies microbial populations in the lab.

1 Introduction

Measuring batch culture growth curves is a common method used by nearly all who work with single-celled organisms in the laboratory. Growth curves allow experimenters to readily

*Co-first and co-corresponding author: hermina.ghenu@gmail.com
†Co-first and co-corresponding author: loic.marrec@iee.unibe.ch

measure population phenotypes, like the dynamics and efficiency of growth in particular environments, for microscopic organisms whose individual cell phenotypes are often laborious or expensive to quantify. The growth rate is an especially important trait for evolutionary microbiologists and microbial ecologists. The growth rate is important because it is related to fitness in population biology, it is used to estimate the number of generations a microbial culture has been growing for (e.g., Wein and Dagan 2019), it is more responsive to selection that other traits in microbial evolution experiments (Wahl and Zhu, 2015), and it is central in describing competition for limited resources (Miller et al., 2005; Bernhardt et al., 2020). Overall, growth curves are commonly used because they are easy to obtain, have been used for a long time, and usually give consistent results within an experiment. The importance of growth curves is only increasing in the age of high-throughput experimental screens of microbial populations, from which conclusions are drawn about responses to ecological challenges, to antimicrobial drugs, and about optimal strains for agricultural purposes. Nevertheless, despite the popularity of gathering growth curve data and the proliferation of methods for extracting growth parameters from said data, it is not clear what is the best method for estimating values of interest for these data.

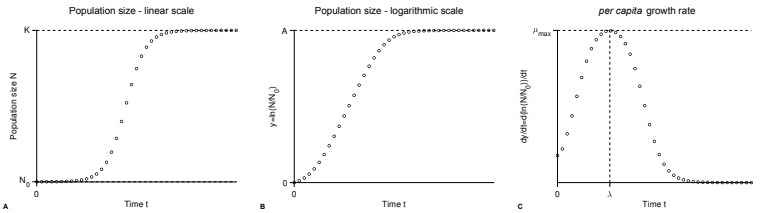


Figure 1: **A schematic illustration of the same simulated batch culture population growth curve, plotted in three different ways.**: **A)** Population size N versus time t. The quantity K corresponds to the carrying capacity and N_0 to the initial population size. The initial population density or fraction is given by N_0/K. **B)** Logarithm of the population size divided by the initial population size $y = \ln(N/N_0)$ versus time t. Note that in other publications the quantity $\ln(N/N_0)$ is sometimes denoted as y. The quantity $A = \ln(K/N_0)$ is the logarithm of the fold increase over the initial population size at carrying capacity. **C)** First derivative of the logarithm of the population size divided by the initial population size $\mathrm{d}y/\mathrm{d}t = \mathrm{d}\ln(N/N_0)/\mathrm{d}t$ versus time t. The function $\mathrm{d}y/\mathrm{d}t$ may be interpreted as the *per capita* growth rate. The quantity $\mu_{\max}$ is the *per capita* maximum growth rate and λ the lag phase duration.

The idea behind the batch culture growth curve is simple: inoculate a sterile culture medium with a small number of individuals N_0 and track the increase in population size over time using any available method to estimate population size (e.g., colony forming units, optical density, microscopy cells counts, flow cytometry). An idealized growth curve is shown in Figure 1, in which each panel shows the same simulated data but with a different y-axis. When an experimenter assesses a growth curve, what they may first observe is perhaps a "lag phase" with little growth. Then, they will *always* observe a phase of rapid growth, alternately called the "log phase" or "linear phase" by different researchers, in which the growth is linear

when shown with the y-axis on a log-scale (Figure 1B). Figure 1C shows the first derivative of Figure 1B: in other words, the instantaneous *per capita* growth rate. The fastest growth rate is usually reached during the "log/linear phase" (Figure 1C; Monod 1949). Then, this phase may be followed by more decelerations and subsequent accelerations under diauxic/diphasic growth conditions (not shown or further discussed herein). Eventually the population growth slows down to halt at the "stationary phase", reaching the final carrying capacity (denoted K in linear-scale, Figure 1A, or A in log-scale, Figure 1B) when all the usable resources are depleted from the batch culture.

Growth curves are often used to estimate the *per capita* growth rate and fitness. The *per capita* growth rate is important in population biology because it is used to calculate the growth of a mutant strain as compared to a wild-type strain: this is the relative growth rate, or the relative fitness (w). The relative fitness is particularly important in evolutionary biology since it classifies a mutant as deleterious ($w < 1$), neutral ($w = 0$) or beneficial ($w > 1$) with respect to natural selection. Indeed, when the relative fitness is greater than 1, the mutant reproduces faster than the wild-type, and conversely when w is less than 1. Although there is still discussion in the field about whether the relative growth rates measured from monoculture growth curves are predictive of competitive fitness (Concepción-Acevedo et al., 2015; Ram et al., 2019), many biologists use the growth rate as a measure of fitness (e.g., Knopp and Andersson 2018).

Microbial batch culture protocols have been used for over 100 years in microbiology (e.g., Slator 1916) and population ecology (e.g., Carlson 1913), and remain a mainstay of experimental evolution and ecology. During this time, many experimental protocols (Delaney et al., 2013; Hall et al., 2014; Stevenson et al., 2016; Kurokawa and Ying, 2017) and estimation methods (Zwietering et al., 1990; Baranyi and Roberts, 1994; Jung et al., 2015) have been developed for this type of data. Since the early 1990s, automated plate readers that incubate and periodically scan the opacity of the cultures growing in the microwells have simplified the process of gathering data for hundreds of bacterial populations simultaneously growing in (relatively) homogeneous batch culture environments. Sources of inconsistency, like the batch effect (Blomberg, 2011), can be mitigated, for example by growing all cultures of interest on the same day(s), in order to arrive at consistent data. Nevertheless, despite the long tradition and good recommendations for setting up experiments, programs and papers detailing methods for estimating growth rates (and other growth parameters) from this data continue to be published and highly sought after. Many of these estimation methods are implementations of classical models (e.g. Sprouffske and Wagner 2016; Petzoldt 2020). This shows that also after 100+ years of generating growth curve data, microbial ecologists and evolutionary biologists are still struggling to find the best way of estimating the growth rate from their data.

The main goal of our paper is to demonstrate that there are significant limitations to existing methods for using batch culture growth curve data to estimate the intrinsic growth rate μ, which is the fastest *per capita* number of divisions per time unit possible when the cell's resources are limitless or otherwise optimal. These limitations impact the calculation of quantities of interest such as the selection coefficient and the relative fitness. We first take stock of how the community currently analyzes growth curve data by semi-quantitatively reviewing the literature to survey which methods are used in evolution and ecology. After explaining different approaches for modelling growth curves, we then use math and simulations to show that many of the currently used approaches are inappropriate for accurately estimating the

intrinsic growth rate μ. We quantify the errors for the intrinsic growth rate μ when the maximum attained growth rate $\mu_{\max}$ is used as an estimator and the generating model is known. Next, we present the limited set of conditions in which an exponential approximation can be used for estimating μ. Importantly, we demonstrate that using inaccurate estimates of μ to estimate the relative fitness often leads to inaccurate fitness estimates and sometimes to wrongly classifying a beneficial mutation as deleterious in some cases. Finally, we apply our theoretical insights to previously published data and show that both absolute and relative growth rate estimates may vary greatly depending on the method. Overall, we present a systematic evaluation of different methods, with recommendations for best practices.

2 Results & Discussion

2.1 Literature review: How does the community analyze the data?

We reviewed 50 papers from evolution and ecology that estimated growth curves for all types of microbial data (see Table S1 and Methods section). Most of the data (90%) were acquired by an automated microplate reader tracking optical density (OD) over time. Other data types included cell counts or fluorescent yields over time. Several papers (6%) did not report the starting inoculum size. Of those papers that reported the inoculum size, 52% used a fixed absolute initial population size ($\widehat{N_0}$) for inoculation whereas 44% used a constant initial population fraction ($\widehat{N_0} = \widehat{K} \times$ *dilution factor*), which we hereafter refer to as the dilution fraction. A constant dilution fraction means that stationary-phase cultures, whose populations are at their carrying capacity, K, were diluted by a constant dilution factor (e.g., $1/1000\times$). This means that for experiments with a constant dilution fraction the absolute population size of the inoculum, N_0, differed between strains/treatments when the carrying capacities, K, were different. For experiments using a constant dilution fraction, the reported dilution factor varied between 10^{-4} to 10^{-1} with a geometric mean value of $10^{-2.37}$.

We found that the growth rate was by far the most commonly estimated growth parameter (94% of all papers reviewed). The other estimated growth parameters were: the carrying capacity (44%), the lag time (34%), and the area under the curve (AUC; 12%). The growth rate was usually reported as an absolute value for each strain/treatment. Moreover, 24% of papers estimated the relative growth rate, or relative fitness, of different strains as compared to the wild-type.

Methods that explicitly fit a model of population growth were used about as often as "model-free" or "nonparametric" approaches (i.e., methods that do not require a model; Figure 2A). One particular model-free approach, the "Easy Linear" method, was especially popular (right doughnut chart of Figure 2A): it was used in about a third of all papers. When models *were* used, only one model was usually reported to have been fit (left doughnut chart of Figure 2A).

We classified the growth curve analysis methods as either mechanistic or phenomenological (Figure 2B). A mechanistic model allows researchers to simulate the underlying process. In contrast, a phenomenological approach enables researchers to describe and quantify the pattern of interest but without simulating the underlying process. We found that phenomenological approaches were used more often than mechanistic models (Figure 2B). The most popular methods within the phenomenological approach were the various model-free methods (right doughnut of Figure 2B). The logistic model was by far the most popular mechanistic

model used (left doughnut of Figure 2B). Depending on its equation, the Gompertz model is either a phenomenological or a mechanistic model (table 1); however, we found that the mechanistic Gompertz model was never fitted whereas the phenomenological Gompertz model was popular (18% of all papers and 28% of all phenomenological methods used).

We found that many growth curve experimental methods, data, and analyses do not yet conform with recommendations for reproducible research (e.g., Wilkinson et al. 2016; Munafò et al. 2017). Over 10% of the papers reviewed reported insufficient information about how growth curves were analyzed and therefore could not be classified for Figure 2. Some highly-cited papers (e.g. Gullberg et al. 2011; Trindade et al. 2012) neither cited an established method nor included sufficient description of their *ad hoc* methods for estimating growth rates (see table S1). Beyond reporting of experimental methods, the data-set itself was often not shared: about half (46%) of all papers do not show any figures of nor provide any of the growth curve data (see table S1). 40% of papers provided plots of at least a subset of the growth curves and 14% of all papers published their raw growth curve data.

Our finding from Figure 2 that $\sim 13\%$ of articles provide insufficient information regarding their growth curve analysis methods likely *under*estimates the magnitude of the problem. This is because we found articles for inclusion in the review by searching among the citations to previously published growth curve analysis methods papers (see methods). Therefore, most of the papers we included cited an established method for analyzing growth curve data. Hopefully these issues of methods under-reporting will improve as scientists become more knowledgeable about recommendations for open science and data management (Wilkinson et al., 2016; Munafò et al., 2017).

Our finding of insufficiently reported information regarding the analysis of growth curves corroborates previous concerns about the lack of a standard method for growth curve analysis (Fernandez-Ricaud et al., 2016). The remainder of our article discusses different methods for analyzing growth curves and, thus, will hopefully contribute to an increased appreciation of why it is important to provide sufficiently detailed methodological information on data analyses.

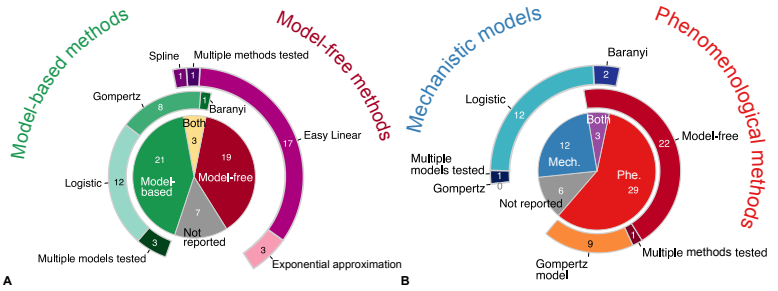


Figure 2: Pie and doughnut charts of literature review results. **A) Model-free methods were used about as often as model-based methods** across all 50 papers reviewed. "Both" refers to when both a model-free and a model-based method were used in the same paper. The doughnut charts surrounding the central pie-chart illustrate how frequently different models, including both phenomenological and mechanistic models, were used (left, shades of green) and how frequently different phenomenological model-free methods were used (right, shades of purple). **B) Phenomenological approaches (Phe) were used more often than mechanistic models (Mech)**. "Both" refers to when both a mechanistic model(s) and a phenomenological approach(es) were used in the same paper. The doughnut charts surrounding the central pie-chart illustrate how frequently different mechanistic models were used (left, shades of blue) and how frequently different phenomenological approaches were used (right, shades of red-orange). The Gompertz model is listed twice because it is either a mechanistic model or a phenomenological model, depending on the equation; however, no paper was found to use a mechanistic Gompertz model. Both of the phenomenological Gompertz models listed in Table 1, Gompertz and modified Gompertz, are grouped together. For more details on the methods and equations used by each paper, see table S1. All slices within each pie and doughnut chart show the counts of papers included.

2.2 Exposition of existing models and methods

2.2.1 Conceptual distinctions between parametric vs non-parametric methods and phenomenological vs mechanistic models

A variety of approaches have been developed over the years to describe and quantify growth curves, as shown in Figure 2. Below we explain the main differences between the most commonly used approaches and models, then we compare the advantages of each.

Model-free vs model-based methods:
One way to classify the different methods is to distinguish between model-free (or non-parametric) methods and model-based (or parametric) methods. Model-free methods use an algorithm to find an estimate of the growth rate that is relatively robust to any noise error in the data. For example, in the classical exponential approximation from Monod (1949) that is featured in many introductory microbiology textbooks, growth rate is estimated by measuring cell concentrations (N_1 and N_2) at two time points (t_1 and t_2) during the "log" phase of growth, then calculating $R = (\log_2 N_2 - \log_2 N_1)/(t_2 - t_1)$. This is an algorithm that can easily be used by hand or performed by a computer. The Easy Linear method (Hall et al., 2014; Mira et al., 2017) is a more complex algorithm that uses a sliding window of five successive data points to calculate the maximum slope among many linear regressions fitted to the log-scale growth curve data. Another example is the Spline method that calculates the maximum value of the first derivative of the log-scale growth curve data by either using the mean of three successive pairs of points (e.g., Ashino et al. 2019) or kernel smoothing of the growth curve (e.g., Kahm et al. 2010; Petzoldt 2020) in order to remove experimental noise. The parameters of these algorithms are usually tunable, for example the size of the sliding window used by the Easy Linear method. However, no explicit assumptions are made about the shape of the growth curves.

Model-based methods, on the other hand, use equations to explicitly describe the relationship between time, the independent variable, and population size (or a proxy of population size, like OD), the dependent variable. A model fitting algorithm is then used to find the model parameters that best fit the observed data, usually by numerical minimization of the residual sum of squares. Model-based methods tend to be preferred by theoretical and statistical biologists because models specifically define the assumptions that are being made, model-based methods have defined protocols for assessing goodness of fit, and model-fitting allows quantification of the (frequentist or Bayesian) error of the estimates (Otto and Day, 2011; Bolker, 2008).

Empiricists tend to prefer to apply model-free approaches to biological growth curve data both because it is technically more difficult to fit and compare models and because of the inflexibility of existing models to fit the data (source: personal communication). In other words, the main advantage of model-free approaches is that they do not require a model. Model-free approaches have drawbacks, however: since there is no model, it is not clear how to compare the likelihood or goodness-of-fit between different methods and bootstrappind is necessary to quantify the error around estimates (for example, in order to generate the 95% confidence intervals).

Mechanistic vs phenomenological models:
Within model-based methods, there is a distinction between mechanistic and phenomenological models. Mechanistic models may also be called process models and phenomenological models can be called statistical/pattern models (Bolker, 2008; Liberles et al., 2013). We de-

fine mechanistic models as representations that simulate or at least bear some resemblance to the underlying process(es) that produced the data. Therefore, the parameters estimated from mechanistic models are interpretable with respect to the underlying process. For population growth curve models in particular, the *per capita* growth rate of a mechanistic model is by definition the intrinsic growth rate μ when the cell's resources are limitless or otherwise optimal.

Phenomenological models can be thought of as 'black-box empirical models' (Chezeau and Vial, 2019). They are created by selecting functions that have a similar shape as the pattern of interest. Then, the parameters of those functions are given a biologically relevant meaning. For growth curves in particular, phenomenological population growth models are defined such that the point on the curve with the fastest rate of *per capita* growth (i.e., the inflection point of $y = \ln(N/N_0)$) corresponds to the maximum growth rate $\mu_{\max}$ (Zwietering et al., 1990, equations 4 & 5). Model-free (non-parameteric) methods all belong to the phenomenological category, as shown in Figure 2B, as they quantify specific parameters of interest without simulating the underlying process.

Although the distinction between mechanistic models and phenomenological methods may seem arcane, it is a crucial distinction because parameters estimated from phenomenological methods cannot be treated as unbiased estimators of parameters for mechanistic models (Rodrigue and Philippe, 2010). We explain below how serious pitfalls in estimating growth rates are a direct result of the differences between growth rates estimated by mechanistic models (i.e., the intrinsic growth rate μ) and phenomenological methods (i.e., the maximum growth rate $\mu_{\max}$).

2.2.2 Mathematical description of models and model parameters, including the initial fraction

Many models have been developed to describe growth curves (e.g., Tsoularis and Wallace 2002; Huang 2011; Baranyi and Roberts 1994). We have summarized the equations and parameters of the most prevalent models found by our literature review in Table 1, distinguishing mechanistic models (Mech) from phenomenological models (Phe). As explained in the section above, a main difference to note is that mechanistic models are defined in terms of the intrinsic growth rate μ, whereas phenomenological models are defined in terms of the maximum growth rate $\mu_{\max}$. Mechanistic models describe population size N as a function of time (i.e., linear scale), whereas phenomenological models describe $y = \ln(N/N_0)$ as a function of time (i.e., they operate on a logarithmic scale).

The common feature of all models, whether mechanistic or phenomenological, is that they have a sigmoid or 'S' shape (Zwietering et al., 1990). All models consider single-strain, well-mixed bacterial populations whose every individual divides at the same *per capita* rate, although the division rate varies over time. Each of these populations starts with N_0 microbes and their maximum population size is defined by a carrying capacity, written as K for mechanistic models or A for phenomenological models.

The models differ in important ways. Most of the models derive from the logistic model but seek to generalize it (Tsoularis and Wallace, 2002). For example, the Richards (both mechanistic and phenomenological) and Baranyi models include a parameter to set the growth inflection and a lag phase, respectively, whereas the Huang model incorporates both. The Gompertz models (both mechanistic and phenomenological) have no additional parameters

but display faster growth than the logistic model for the same set of parameters since they have a higher *per capita* growth rate. Futhermore, the Richards model requires an extra parameter β that determines how quickly deceleration occurs as the stationary phase is reached, whereas the Baranyi model includes a lag phase specified by the parameter h_0. The Huang model includes a lag phase defined by τ in addition to a parameter α determining the curvature.

The initial fraction, N_0/K, is an important value to keep in mind throughout our paper. This is the size of the population at inoculation (i.e., N at time $t = 0$) divided by the final carrying capacity. Only when the initial fraction is small can one distinguish models that have an initial exponential growth, like the logistic model, from models with a lag phase, like the Gompertz, Richards, Baranyi, and Huang models.

2.2.3 What is the difference between μ and $\mu_{\max}$?

It is important to distinguish between three growth rate estimators: the maximum population growth rate ($\max(\mathrm{d}N/\mathrm{d}t)$), the *per capita* maximum growth rate ($\mu_{\max}$), and the *per capita* intrinsic growth rate (μ). The maximum population growth rate $\max(\mathrm{d}N/\mathrm{d}t)$ is the fastest increase in size achieved by the *entire population*. We are not interested in the maximum population growth rate parameter since it is not estimated by any of the methods or models we discuss here; we only mention it so that the reader does not mistake it for the maximum growth rate ($\mu_{\max}$). The maximum growth rate $\mu_{\max}$ is the fastest *per capita* number of divisions per unit of time actually achieved in the observed growth curve. In more quantitative terms, $\mu_{\max}$ is the maximum value of the curve $\mathrm{d}\ln(N/N_0)/\mathrm{d}t$ (Figures 1B-C). As mentioned above, the maximum growth rate $\mu_{\max}$ is a value estimated using phenomenological approaches. Finally, the intrinsic growth rate μ (sometimes denoted as r (Sprouffske and Wagner, 2016), called the Malthusian parameter of population growth, or the intrinsic rate of increase) is the fastest *per capita* number of divisions per unit of time theoretically possible and, because it is a mechanistic model parameter, it is used for simulating population growth processes. We here focus on the intrinsic growth rate μ and the maximum growth rate $\mu_{\max}$ because these are the two quantities estimated by the most used methods.

There is an important conceptual difference between the intrinsic growth rate μ and the maximum growth rate $\mu_{\max}$. The intrinsic growth rate μ is the theoretical maximum number of cell divisions per time unit assuming population dynamics that follow an exponential law. However, No real population achieves an infinite size because the division process is limited by space and/or nutrients, for instance. Thus, the number of divisions per time unit is not constant over time, so the maximum division rate $\mu_{\max}$ is the largest *per capita* value observed during the population growth. Therefore, the intrinsic growth rate μ can quantify the strain-specific division rate theoretically independently of the environment or experimental conditions. On the other hand, the maximum growth rate $\mu_{\max}$ (like other values estimated by phenomenological methods) is always specific to the experiment itself and cannot be generalized as a strain-specific value that applies to different environments or conditions. As will be shown below, in the best case scenario $\mu_{\max}$ approximates μ, but in other scenarios $\mu_{\max}$ is a composite parameter that depends on other values, like the inoculum size and lag time.

Previous work has pointed out confusions between different growth rate estimators (Perni et al., 2005). The confusion between these terms is so prevalent that some papers mistook μ for $\mu_{\max}$ (Yang et al., 2006), vice versa (Wu et al., 2017), or distinguished between the two but swapped the names (Khan et al., 2017). Furthermore, some authors wrote the

Model	Equation	Parameters
Logistic (Mech)	$N(t) = \frac{N_0 K \exp(\mu t)}{K + N_0(\exp(\mu t) - 1)}$	K, N_0, μ
Gompertz (Mech)	$N(t) = (N_0/K)^{\exp(-\mu t)} K$	K, N_0, μ
Richards (Mech)	$N(t) = \frac{N_0 K}{(N_0^\beta + (K^\beta - N_0^\beta)\exp(-\mu\beta t))^{-1/\beta}}$	K, N_0, μ, β
Baranyi (Mech)	$N(t) = \frac{(-1 + \exp(h_0) + \exp(\mu t))N_0 K}{(\exp(\mu t) - 1)N_0 + \exp(h_0)K}$	K, N_0, μ, h_0
Huang (Mech)	$N(t) = \frac{N_0 K}{N_0 + (K - N_0)(1 + \exp(\alpha\tau))^{\mu/\alpha}(\exp(\alpha t) + \exp(\alpha\tau))^{-\mu/\alpha}}$	K, N_0, μ, α, τ
Logistic (Phe)	$y(t) = \frac{A}{1 + \exp(4\mu_{\max}(\lambda - t)/A + 2)}$	A, $\mu_{\max}$, λ
Gompertz (Phe)	$y(t) = A\exp(-\exp(\mu_{\max}\exp(1)(\lambda - t)/A + 1))$	A, $\mu_{\max}$, λ
modified Gompertz (Phe)	$y(t) = A\exp(-\exp(\mu_{\max}\exp(1)(\lambda - t)/A + 1)) + A\exp(\alpha(t - t_{shift}))$	A, $\mu_{\max}$, λ, α, t_{shift}
Richards (Phe)	$y(t) = A(1 + \nu\exp(1 + \nu + \mu_{\max}(1 + \nu)^{1+1/\nu}(\lambda - t)/A))^{-1/\nu}$	A, $\mu_{\max}$, λ, ν

Table 1: **Different population growth models:** The population growth models considered in this paper with their equation and parameters. Mechanistic models are indicated by (Mech) and describe the population size N as a function of time t. Every mechanistic model includes a carrying capacity K, an initial population size N_0 and an intrinsic growth rate μ. The mechanistic Richards model includes a parameter β to adjust its inflection point and the Baranyi model has a lag phase defined by h_0. The Huang model has both with parameters α and τ, respectively. Phenomenological models are indicated by (Phe) and describe $y = \ln(N/N_0)$ as a function of time t. Every phenomenological model includes a carrying capacity $A = \ln(K/N_0)$, a maximum growth rate $\mu_{\max}$ and a lag phase λ. The modified Gompertz model (Phe) displays a second increase after the growth reaches a first saturation plateau. The parameters t_{shift} and α control the time and the slope of the second increase. The inflection point in the phenomenological Richards model is adjustable by the parameter ν.

mechanistic logistic equation as $\mathrm{d}N/\mathrm{d}t = \mu_{\max}(1 - N_0/K)N$ (Petzoldt, 2020), whereas other authors preferred $\mathrm{d}N/\mathrm{d}t = \mu(1 - N/K)N$ and $\mu_{\max} = \max(\mathrm{d}\ln(N(t)/N_0)/\mathrm{d}t)$ (Sprouffske and Wagner, 2016). Different naming conventions become even more misleading for models in which the intrinsic growth rate is a function of the resource concentration, such as the Monod class of models that have their own specific, mechanistic definition for μ_{max} (Monod, 1949; Chezeau and Vial, 2019). We will not discuss substrate-use models herein. Having explained the conceptual differences between the $\mu_{\max}$ maximum growth and μ intrinsic growth rates above, we now expand on the mathematical differences.

Deriving the difference between $\mu_{\max}$ vs. μ:
In the following, we mathematically explain why $\mu_{\max}$ is not always a good proxy for μ, especially at large initial population fractions, N_0/K. Analytical math is combined with simulations to show for which initial population fractions an experimenter can estimate the intrinsic growth rate μ from the maximum growth rate $\mu_{\max}$.

We assumed that the population dynamics follow one of the mechanistic growth models from Table 1 and mathematically derived $\mu_{\max}$ for the five mechanistic models. As reported in Table 2, $\mu_{\max}$ depends on the system parameters, namely the initial population fraction N_0/K as well as the parameters β and h_0 for the Richards and the Baranyi models, respectively.

The results of Table 2 are illustrated by the points in Figure 3A. The estimated maximum growth rate ($\mu_{\max}$) values differ between models with the same parameter values (intrinsic growth rate $\mu = 1$ and carrying capacity $K = 10^5$). In general, the maximum growth rate is approximately equal to the intrinsic growth rate when the initial fraction of individuals satisfies $N_0/K \ll 1$ and $(N_0/K)^\beta \ll 1$ in the Logistic and Richards models, respectively. Indeed, these conditions lead to $\mu_{\max} = \mu(1 - N_0/K) \approx \mu$ and $\mu_{\max} = \mu(1 - (N_0/K)^\beta) \approx \mu$ for the Logistic and Richards models, respectively (see Table 2). The Gompertz model is a special case since the maximum division rate is a good proxy for the intrinsic division rate when the initial fraction is large, roughly equal to $\exp(-1) \approx 0.37$ (i.e., an inoculum size corresponding to a dilution factor for the stationary phase batch culture of between one-third and two-fifths).

In order to test the analytical predictions from Table 2, we evaluated the growth rates as estimated by model-free methods using data simulated under each of the five population growth models. Unlike experimental data, for which the true μ value that generated the data is never known, estimating the growth rate from simulated data allows us to check the accuracy of the estimates as compared to the known μ parameter that the data was simulated under.

We focus on two model-free methods, the popular Easy Linear (Hall et al., 2014) and the Spline (e.g., Adkar et al. 2017; Ashino et al. 2019) methods, to determine the maximum growth rate $\mu_{\max}$. Both methods assume that only the exponential stage of growth is useful to estimate the maximum growth rate. We generated data using individual based stochastic simulations for the Gompertz, Richards, Logistic, Huang, and Baranyi models. Then we used the two different model-free methods, Spline and Easy Linear, to compute the maximum growth rate $\mu_{\max}$ for different parameter values. In practice, both model-free methods provided us with the same results. Our simulated data averaged over several stochastic realizations did not include the myriad sources of noise present in experiments, therefore resulting in a low noise level.

As shown by the lines in Figure 3A, there is an excellent agreement between our analytical predictions (lines) and the estimates from simulated data (points). As predicted analytically,

Model	**Derived maximum growth rate $\mu_{\max}$**
Logistic (Mech)	$\mu(1 - N_0/K)$
Gompertz (Mech)	$-\mu \ln(N_0/K)$
Richards (Mech)	$\mu(1 - (N_0/K)^{\beta})$
Baranyi (Mech)	$\frac{\mu e^{-h_0}(-2(1+\sqrt{(-1+e^{h_0})(e^{h_0}(K/N_0)-1)})(N_0/K)+e^{h_0}(1+N_0/K))}{1-N_0/K}$
Huang (Mech)	Numerical solution
Logistic (Phe)	$\mu_{\max}$
Gompertz (Phe)	$\mu_{\max}$
modified Gompertz (Phe)	$\mu_{\max}$
Richards (Phe)	$\mu_{\max}$

Table 2: **Maximum growth rates:** The maximum growth rate $\mu_{\max}$ for the population growth models considered in this paper. Mechanistic models are indicated by (Mech), whereas the phenomenological models are indicated by (Phe). The maximum growth rate was derived for the mechanistic models by analytically determining $\max(\mathrm{d}y/\mathrm{d}t) = \max(\mathrm{d}\ln(N/N_0)/\mathrm{d}t)$. All phenomenological models have the same maximum growth rate, whereas the maximum growth rate differs between mechanistic models.

the estimated maximum growth rate $\mu_{\max}$ is not equal to the known intrinsic growth rate μ value used to create the simulations, unless $N_0/K \ll 1$. For the Baranyi, Huang, Logistic, and Richards models, the smaller the initial population fraction, the better the maximum growth rate performs as a proxy for estimating the intrinsic bacterial growth rate, μ. However, this is not the case for the Gompertz model. For the Baranyi and Richards models (supplementary Figures S4A and S4B), the smaller the parameter h_0 and the larger the parameter β, the closer is the maximum growth rate $\mu_{\max}$ to the intrinsic growth rate μ. Similarly, for the Huang model, the higher the curvature defined by α and the shorter the duration τ of the lag phase, the better $\mu_{\max}$ is as a proxy for μ (see Figures S4C and S4D).

We have shown that the maximum growth rate $\mu_{\max}$ is not always equivalent to the intrinsic growth rate μ. Therefore, methods that estimate the maximum growth rate $\mu_{\max}$ but then (often implicitly) assume that this value can be treated as the μ of a mechanistic model must be applied with caution. As we have demonstrated in Table 2 and Figure 3A, $\mu_{\max}$ tends to *under*estimate the true intrinsic growth rate μ – except when population growth follows the Gompertz mechanistic model, in which case the maximum growth rate $\mu_{\max}$ mostly *over*estimate the true intrinsic growth rate μ. This is because the *per capita* growth rate is generally smaller than the intrinsic one. Hence, we recommend that a clear distinction must be made between the intrinsic growth rate μ and the maximum growth rate $\mu_{\max}$.

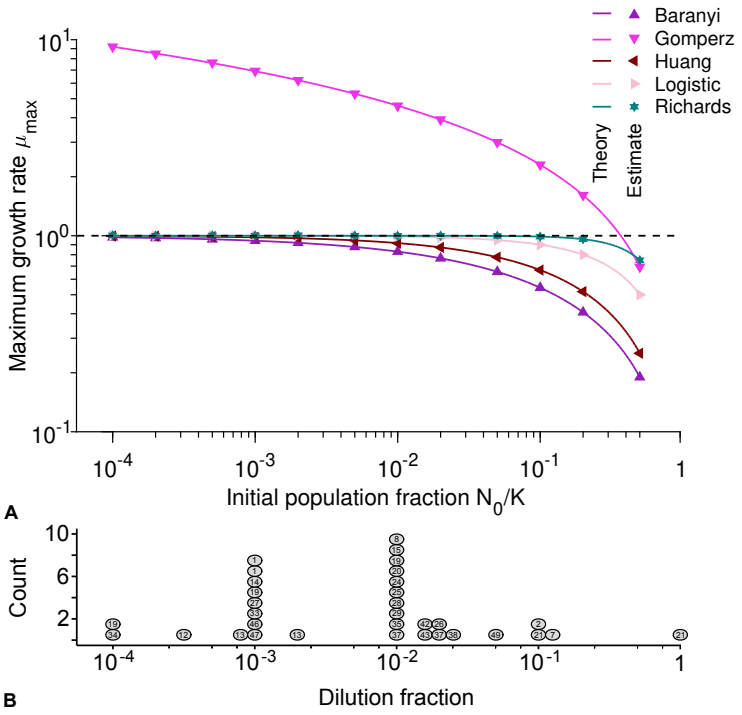


Figure 3: **A) The Gompertz model and large initial population fractions make the maximum growth rate a poor proxy for the intrinsic growth rate:** Maximum growth rate $\mu_{\max}$ versus initial population fraction N_0/K for different mechanistic population growth models, where $\mu = 1$. Each point represents estimated values by Spline (from Petzoldt 2020) from simulated data averaged over 10^4 stochastic realizations. The solid lines correspond to the analytical predictions of the maximum growth rate (see Table 2). The dashed line shows the intrinsic growth rate value μ. Parameter values: $K = 10^5$, $\mu = 1$, $\alpha = 2$, $\beta = 2$, $h_0 = 2$ and $\tau = 2$. The initial population fraction is defined as a combination of model parameters, N_0/K, while the initial dilution fraction is an empirical quantity extracted from experimental data as detailed in the literature review methods. **B) For the most commonly used dilution fractions in the literature, the maximum growth rate is a good proxy for the intrinsic growth rate:** Histogram of the estimated dilution fractions observed for the 27 (out of 50) papers that provided sufficient information to estimate this value. Each circle represents a publication and the number inside the circle indicates the number of the reference as given in supplementary table S1. Several publications appear more than once because they used more than one dilution fraction for different experiments.

The initial fraction is a key parameter determining the relationship of $\mu_{\max}$ and μ

Above we showed that (for most models) we can use the estimated maximum growth rate $\mu_{\max}$ as an approximation of μ when initial population fractions N_0/K are small. We estimated the initial dilution fractions used in different experiments in order to ascertain whether most studies are using appropriately small initial fractions. Figure 3B shows the estimated dilution fractions used by different papers included in our literature review. Papers that had experiments with multiple, different batch culture starting conditions are included as multiple points with the same number. For example, Ganucci et al. 2018 (labeled as 21) has a dilution fraction near 1. The authors used cell viability counts to track yeast growth in media with increasing ethanol concentrations, sometimes resulting in almost no growth. Two values from Ganucci et al. (2018) are summarized in Figure 3B, indicating the largest (0.1) and the smallest (0.94) dilution fractions observed. Since methods differ between publications, with some using a mid-exponential phase culture and others using a stationary phase culture for inoculation (see supplementary table S1), the estimated dilution fraction should be considered as an upper bound for the initial fraction used in each paper.

More than two-thirds of papers use at least one estimated dilution fraction smaller than 10^{-2}; the geometric-mean observed dilution fraction was $10^{-2.7}$. For such small dilution fractions, if there is no lag time, and if growth follows one of the population growth models except Gompertz, the maximum growth rate $\mu_{\max}$ tends to be a good estimator of the intrinsic growth rate μ. When the true growth curve dynamics in the experiments follows one of the mechanistic models other than Gompertz, the relative difference between the maximum growth rate and the intrinsic growth rate for the mean initial population fraction $N_0/K = 10^{-2.7}$ is between 0-12%. Here, the largest difference between $\mu_{\max}$ and μ is obtained for the Baranyi model. However, for the Gompertz model the relative difference between the maximum growth rate and the intrinsic growth ranges from -821% to 31% (see Figure 3).

Discussion of the difference between $\mu_{\max}$ and μ:

We showed that the $\mu_{\max}$ values calculated from mechanistic models depend on other parameters in addition to μ, such as the initial population fraction. Conversely, one cannot obtain μ from $\mu_{\max}$ alone. For mechanistic models, additional parameters such as the initial population size and carrying capacity are required (see Table 1) to be able to calculate μ from $\mu_{\max}$. For phenomenological approaches, which are specified directly in terms of $\mu_{\max}$, μ is not defined. Nevertheless, even when $\mu_{\max}$ is estimated by phenomenological models, its estimated value will be different when experimental quantities such as the initial population size and carrying capacity change. In reality, the estimated values for both $\mu_{\max}$ and μ may also vary with the experimental conditions (such as genotype, medium, temperature, etc).

We emphasize that the main difference between the maximum growth rate $\mu_{\max}$ and the intrinsic growth rate μ is that $\mu_{\max}$ is a phenomenological quantity, whereas μ is a model parameter. Therefore, obtaining different estimates of μ and $\mu_{\max}$ is expected and not a sign of bad performance of a model or method, especially for large initial population fractions. For example, the manual of one software (Delaney, 2014) provides options for fitting various phenomenological models and a single mechanistic model but discourages users from applying the mechanistic model because "it predicts fastest growth" as compared to the other implemented models. Our results can readily explain this observation and debunk the implied worse performance of the mechanistic model. Indeed, the publication associated with this software recommends a large inoculum size (Delaney et al. 2013; associated studies labeled

as 8, 22, 25, 42, and 43 in Figure 3B and supplementary table S1), for which we showed that $\mu_{\max}$ consistently overestimates μ.

Given the differences between $\mu_{\max}$ and μ, which of these should preferably be estimated? Unfortunately, there is not a one-size-fits-all answer to this question. From a theoretician's perspective, we recommend that researchers estimate the intrinsic growth rate μ. Most studies perform growth curve experiments to characterize growth for specific strains, treatments, or environments in the (either explicit or implicit) context of population ecology or mechanistic models – all of which are defined in terms of μ. Only the intrinsic growth rate μ can be used for simulating mechanistic models and therefore to identify the mechanisms which best explain the observed dynamics. The maximum growth rate $\mu_{\max}$, on the other hand, is phenomenological: it describes what is observed in the data contingent on the population starting conditions and the experimental environment. At best, $\mu_{\max}$ approximates μ. However, from an experimenter's perspective, estimating $\mu_{\max}$ has the advantage that model-free methods are technically easier to use than estimators of mechanistic models, especially for data that display diphasic or other non-sigmoid/non-'S' shaped growth. We strongly urge researchers who decide to estimate $\mu_{\max}$ to use (and report) a small initial population fraction and to assert that the data do not have a significant lag time.

When using phenomenological methods, a second decision is necessary about whether to use model-based or model-free methods. In our noise-free, simulated data, model-based and model-free phenomenological methods yielded the same estimates for $\mu_{\max}$ (Figure 3A); future work should elaborate on their performance in the presence of different sources of noise (e.g., to expand on the work that has already been done by Mira et al. (2017) for the Easy Linear method). Although models are preferable from a theoretician's view, certain types of experimental data (for example, displaying diauxic growth, curves without samples in the stationary phase, and other non-sigmoid shaped data as well as very noisy data) may not be fitted well by a model of sigmoidal growth. In this case the data may be better summarized by a model-free phenomenological method. Importantly, we recommend that the choice of method is justified clearly and in writing, no matter which method is used.

2.3 Guidelines for estimating growth rates

2.3.1 Ad hoc fitting an exponential curve to growth curves should be avoided

One common approximation (Monod, 1949; Kassen, 2014) that is used to obtain an estimate of the intrinsic growth rate μ is to fit an exponential equation, like $N(t) = N_0 e^{\mu t}$, to the early phases of growth (i.e., during the "log"/"linear" phase, prior to the deceleration and stationary phases). This method is explained in many introductory microbiology textbooks, as previously summarized in the explanation of model-free methods of section 2.2.1 above, and we refer to this approach as an exponential approximation. Under the exponential approximation, the intrinsic growth rate is given by $\mu = \ln(N(t)/N_0)/t$.

To test the accuracy of the exponential approximation when applied to batch culture population growth, we expressed the intrinsic growth rate as a function of population size and other possible parameters for different mechanistic population growth models (Table 3 and Figure S5).

We found that the exponential approximation is frequently a poor estimator of the intrinsic growth rate μ. The exponential approach is never valid for a population following Gompertz

Mechanistic model	**Intrinsic growth rate μ**
Exponential	$\ln\left(\frac{N(t)}{N_0}\right)/t$
Logistic	$\ln\left(\frac{N(t)}{N_0}\frac{1-N_0/K}{1-N(t)/K}\right)/t$
Gompertz	$\ln\left(\frac{\ln(N_0/K)}{\ln(N(t)/K)}\right)/t$
Richards	$\ln\left(\frac{(N(t)/N_0)^\beta(1-(N_0/K)^\beta)}{1-(N(t)/K)^\beta}\right)/(t\beta)$
Baranyi	$\ln\left(1-\frac{e^{h_0}(1-N(t)/N_0)}{1-N(t)/K}\right)/t$
Huang	$\alpha\frac{\ln((-1+K/N(t))/(1-K/N_0))}{\ln((1+e^{\alpha\,\tau})/(e^{\alpha\,t}-e^{\alpha\,\tau}))}$

Table 3: **Intrinsic growth rate as function of the model parameters at time t for different mechanistic population growth models** The difference between the equations indicates that the exponential approximation may often not be appropriate; see also Figure S5.

growth. There is no parameter range for which the equation $\ln\left(\ln(N_0/K)/\ln(N(t)/K)\right)/t$ reduces to $\ln\left(N(t)/N_0\right)/t$ (see Table 3). The exponential approach is valid for the logistic growth when the initial population size is very small in comparison to the carrying capacity (i.e., $N_0 \ll K$) and for time points at which the population size remains small in comparison to the carrying capacity (i.e., $N(t) \ll K$). These conditions lead to $\ln(N(t)(1-N_0/K)/(N_0(1-N(t)/K)))/t \approx \ln(N(t)/N_0)/t$ (see Table 3). This makes sense since the phase during which these conditions are satisfied corresponds to the regime in which logistic growth can be reduced to exponential growth. The same conditions apply to Baranyi growth, with the additional condition that the lag phase must be short (i.e., $h_0 \ll 1$), so that one obtains $\ln\left(1-e^{h_0}(1-N(t)/N_0)/(1-N(t)/K)\right)/t \approx \ln(N(t)/N_0)/t$ (see Table 3). If the lag phase is not short, the exponential phase starts later whereas the exponential approach assumes that it starts at the beginning of the growth. Richards growth is more complex. Here, the quantities $(N_0/K)^\beta$ and $(N(t)/K)^\beta$ must be much less than 1 to make the exponential approach valid. Thus, the larger the deceleration parameter β is when $N_0 \ll K$ and $N \ll K$, the more abruptly the "log"/"linear" phase transitions into the stationary phase, and the more valid the exponential approximation becomes, so that $\ln\left((N(t)/N_0)^\beta(1-(N_0/K)^\beta)/(1-(N(t)/K)^\beta)\right)/(t\beta) \approx \ln(N(t)/N_0)/t$ (see Table 3). Consequently, the exponential approximation is valid only in a very restricted set of conditions: when there is no lag phase, the initial population fraction is very small, and the measured population sizes remain small as compared to the carrying capacity. Most experimental data probably do not meet this necessary set of conditions.

It is of note that throughout the literature, including introductory textbooks, the term "exponential growth rate" tends to be used to describe the intrinsic growth rate μ and some-

times deemed the same as the maximum growth rate $\mu_{\max}$ (e.g., Basra et al. 2018; Novak et al. 2009). We here strongly caution against this conflation of potentially very different quantities. In particular, we recommend that approximating batch culture growth with an exponential curve in order to get an estimate for the intrinsic growth rate μ requires careful assurance that the assumption is valid for the data at hand.

2.3.2 Theory predicts that using $\mu_{\max}$ or the exponential approximation for estimating relative fitness can yield wrong results

In evolution and population biology, the relative fitness w of a mutant (M) compared to a wild-type (WT) is classically defined as the ratio of μ^M/μ^{WT}. The relative fitness or the selection coefficient, defined as $s \equiv w - 1 = \mu^M/\mu^{WT} - 1$, is used to classify a mutant as deleterious ($w < 1$ or $s < 0$), neutral ($w = 1$ or $s = 0$), or beneficial ($w > 1$ or $s > 0$). Thus, to infer how natural selection favors one strain over another from monoculture growth curves, microbial ecologists and evolutionary biologists need the intrinsic growth rate, which is obtained by fitting a mechanistic model.

One argument regarding the issues of miscalculation of the intrinsic growth rate μ discussed above is that these concerns are important for *absolute* growth rate estimates, but can be disregarded when considering *relative* estimates. Here, the argument is that whereas absolute estimates cannot be compared between data-sets, relative growth rates can be estimated within a data-set by using a common reference sample and these relative growth rates can then be compared between data-sets.Moreover, with regards to selection coefficients, the sign may be more important than the absolute value. Namely, incorrect absolute estimates should yield correct rankings of growth rate estimates, and thus correctly estimated signs of the selection coefficient. Below, we demonstrate that these assumptions are wrong and that incorrect estimates of the growth rates (for example, by assuming that $\mu_{\max} = \mu$) can severely affect the classification of strains into beneficial, neutral, or deleterious.

To demonstrate the effects on relative fitness estimates of assuming $\mu_{\max} = \mu$, we simulated growth curve data from separate batch monocultures for two strains and estimated their maximum growth rates $\mu_{\max}$ using the growth curves. Then we assumed (erronously) that the maximum growth rate $\mu_{\max}$ was a good approximation of the intrinsic growth rate μ. We calculated the relative fitness of the mutant with respect to the wild-type as $w = \mu^M_{\max}/\mu^{WT}_{\max}$ (or the selection coefficient as $s = \mu^M_{\max}/\mu^{WT}_{\max} - 1$).

Figure 4A shows that using $\mu_{\max}$ to estimate the relative fitness generally infers incorrect values for the relative fitness (as well as the selection coefficient), unless both strains have exactly the same initial population fraction (N_0/K). Even more concerning, this estimation sometimes categorizes the mutant as deleterious when it is beneficial, and vice versa. This is especially worrisome because we assumed noise-free data and an ideal case in which both bacterial strains follow the same population dynamics. In summary, equivocating $\mu_{\max}$ with μ is likely to lead to wrong estimates of fitness and selection coefficients.

Mis-estimation of the selection coefficient also occurs when calculating relative growth rate values using the exponential approximation. Lenski et al. (1991) extended the exponential approximation to calculate the fitness of a mutant strain (M) relative to the fitness of a wild-type strain (WT), $w = \ln(N^M(t)/N^M_0)/\ln(N^{WT}(t)/N^{WT}_0)$ (or the selection coefficient $s \equiv w - 1 = \ln(N^M(t)/N^M_0))/\ln(N^{WT}(t)/N^{WT}_0) - 1)$. Note that under the assumption of exponential growth, both the relative fitness and the selection coefficient measured by the

equation stated above are time-independent. Both Lenski et al. (1991) and Ram et al. (2019) empirically set the time interval of measurements t to 24 hours.

Similarly to using $\mu_{\max}$ as a proxy for μ, the exponential approximation yields wrong estimates for the relative fitness (as well as the selection coefficient) in many cases (Figure 4B). For growth curves (except for the Gompertz model) in which the initial population fraction of the mutant is smaller than that of the wild-type, the exponential approximation is a more conservative estimator than $\mu_{\max}$; it is less likely to overestimate the relative fitness. However, when the initial population fraction of the mutant is larger than that of the wild-type, the exponential approximation is likely to incorrectly infer that a beneficial mutant is deleterious (Figure 4B). Thus, the exponential approximation both misestimates and misclassifies throughout much of the experimentally reasonable parameter range.

Implications for the estimation of selection coefficients from growth rates: We showed that using $\mu_{\max}$ as a proxy for μ in order to calculate the relative growth rates or relative fitness can lead to biased estimates (Figure 4). The amount of bias in the estimate becomes larger as the difference in the true growth parameters between the two studied strains becomes larger. Accordingly, we strongly recommend that experimenters use the intrinsic growth rate μ to estimate relative fitness. According to population biology, relative fitness is defined as the ratio of the intrinsic growth rate of the mutant strain over the wild-type strain, μ^M/μ^{WT} (Lenski et al., 1991; Crow and Kimura, 2009; Chevin, 2011). From this point of view, it follows that relative fitness can only be estimated using the intrinsic growth rate μ. Nevertheless, fitness-related phenotypes, like $\mu^M_{\max}/\mu^{WT}_{\max}$, are sometimes used as a summary statistic of population growth that is contextual to the environmental, temporal, and population conditions (e.g., Adkar et al. 2017). Above we showed that a proxy of the intrinsic growth rate (like $\mu_{\max}$ or the exponential approximation) can accurately estimate the relative fitness only if specific criteria are met. When these criteria are *not* met, $\mu_{\max}$ becomes a composite parameter that depends on other experimental quantities that should be reported, like the initial fraction and lag time. Only if experiments are set up with small and equal initial fractions of the mutant and wild-type strains and if the strains do not exhibit a lag phase, using $\mu_{\max}$ as a proxy for μ to estimate the relative fitness may be justified.

2.4 Application of theory: Re-analyzing 4 published data-sets

In order to clarify the theoretical considerations discussed above, we re-analyzed four published data-sets using the diversity of methods discussed above and then compared how the different methods performed on the same data. Among the 50 studies reviewed, we identified four that were appropriate for re-analysis since these papers provided their complete data-set and reported their estimated values (Adkar et al., 2017; Hammer et al., 2021; Ram et al., 2019; Todd and Selmecki, 2020). Each of these bacterial growth curve data-sets reports optical density versus time for different bacterial strains, with a total of 142 curves across all studies. We used three publicly available R packages, Growthcurver, grofit and growthrates, to estimate growth parameters (Sprouffske and Wagner, 2016; Kahm et al., 2010; Petzoldt, 2020). We tested two model-free methods (Spline and Easy Linear) and the model-based methods listed in Table 1. These models are based on different equations that are phenomenological (Zwietering et al., 1990) or mechanistic (Tsoularis and Wallace, 2002; Baranyi and Roberts, 1994; Huang, 2011). We focused on inferring the maximum growth rate ($\mu_{\max}$), both because it is of greatest relevance to the work discussed above and because it is the only quantity

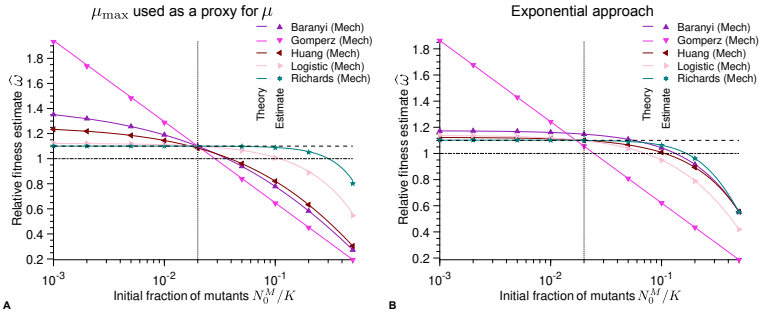


Figure 4: **Different initial population fractions between wild-type and mutant batch cultures result in poor estimates of relative fitness:** Relative fitness $\widehat{\omega}$ estimate versus initial fraction $N_{0,M}/K$ of mutants for different mechanistic population growth models. As a reminder, $\omega = \mu_M/\mu_{WT}$. **A) the maximum growth rate $\mu_{\max}$ is used as a proxy for the intrinsic growth rate μ.** Each point represents estimated values by Spline (growthrates package) from simulated data averaged over 10^4 stochastic realizations. The solid lines correspond to the analytical predictions of the relative fitness using estimates of the maximum growth rate (see Table 2). **B) the intrinsic growth rate is obtained applying the exponential approximation.** Solid lines correspond to the analytical predictions of the relative fitness using estimates of the maximum growth rate (see Table 3). In both panels the dashed line shows the real relative fitness value w. The dotted line represents the configuration in which the growth model parameters of the mutant are equal to the parameters of the wild-type (except their intrinsic growth rates). The dash-dotted line corresponds to the neutral case, i.e. when both the mutant and the wild-type have the same growth rate. Parameter values: $K_{WT} = K_M = K = 10^5$, $\mu_{WT} = 1$, $\mu_M = 1.1$, $\alpha_{WT} = \alpha_M = 2$, $\beta_{WT} = \beta_M = 2$, $h_{0,WT} = h_{0,M} = 2$, $\tau_{WT} = \tau_M = 2$ and $t = 1$.

common to all methods we tested. For mechanistic models (which are defined in terms of μ), we estimated $\mu_{\max}$ by using derivatives to find the maximum slope of $\ln(N/N_0)$ (as shown in Figure 1C; see Table 2).

The maximum growth rate ($\mu_{\max}$) values estimated from the same data vary widely depending on the method used (Figure 5A, and Figures S1A, S2A and S3A). It is not possible to assess the accuracy of the estimates for the model-free methods. However, we used goodness-of-fit tests to assess the model-based methods by calculating the residual sum of squares (RSS). The RSS measures the discrepancy between the data and the fitted model. Thus, the smaller the RSS, the better the model. Since the models we tested have different numbers of parameters, we also calculated Akaike's Information Criterion (AIC) for mechanistic models using the method of López et al. (2004). The results of the AIC are consistent with those of the RSS (see Supplementary Material). Despite the discrepancy in the inferred maximum growth rate, many of the models fit the data well in most cases because the RSS values are low and similar (Figures 5B-C, and Figures S1B-C, S2B-C and S3B-C).

A visual inspection of the fits corroborates our findings that all models fit the data gener-

ally well (visualizations of all fits available at available at https://github.com/LcMrc/GrowthRates). We emphasize that a visual inspection is important to ensure that the estimated values are appropriate. Indeed, in a few cases, the fits proved to be unsatisfactory although the summary statistics were good (see, e.g., Figure S6).

No model is consistently preferred for all samples of a data-set. This highlights the difficulty of choosing 'the one' right model, although this choice has a great impact on the growth parameter estimates. The Gompertz equation, whether phenomenological or mechanistic, frequently yielded the worst statistics, although our literature review indicates the phenomenological Gompertz equation as the most frequently used model. Moreover, the models used in the original publications of the data were not always the models that we found to obtain the best statistics.

As previously stated, model-free methods may be preferred for some research questions, especially when neither the underlying mechanisms nor relative fitness estimates are of interest. Model-free methods obtain the maximum growth rate by determining the maximum of the function $\mathrm{d}\ln(N(t)/N_0)/\mathrm{d}t$ (see Figure 1C). A quick comparison with the derivative of the experimental data ensures the validity of the estimate. We found that Easy Linear gave slightly different results from the Spline method because the former requires the user to specify how many data points to include for the analysis of the log-linear part of the growth curve. Note that model-free methods are likely to be more accurate than model-based methods for estimating $\mu_{\max}$ because the latter involve more parameters and a data fit.

Relative growth rate estimates from empirical data:
Adkar et al. (2017) used $\mu_{\max}$ estimates from a phenomenological Gompertz model in order to estimate relative fitnesses. Therefore, this study allowed us to evaluate our concerns about using $\mu_{\max}$ for relative fitness estimates. First, we used all approaches (model-free methods, phenomenological models, and mechanistic models) to estimate $\mu_{\max}$ for the wild-type and mutant strains from this data-set. Assuming (erroneously) that $\mu_{\max}$ was a proxy for μ, we then calculated the relative fitness. Figure 5D shows that mechanistic models (circles) estimate more beneficial fitness values than methods that estimate $\mu_{\max}$ directly (diamonds for model-free methods and squares for phenomenological models). Especially for strains that were estimated by Adkar et al. 2017 (crosses) to have especially low fitness, we found a large variation in the fitness values estimated by different methods.

Next, we used mechanistic models to estimate the intrinsic growth rate μ for the wild-type and mutant strains from this data-set and subsequently calculated the relative fitness. In Figure 5E we compared our estimated values with those published by Adkar et al. (2017). Again, many models estimate larger fitness values than those reported in the original study. This is especially pronounced for samples that were estimated by Adkar et al. 2017 (crosses) to have especially low fitness.

The overall conclusion of this section is that estimating relative fitness using inaccurate estimates of μ likely propagates to the level of relative fitnesses, and causes large discrepancies between relative fitness values estimated using different methods. Importantly, these discrepancies are most pronounced for samples that are of special interest in an experiment.

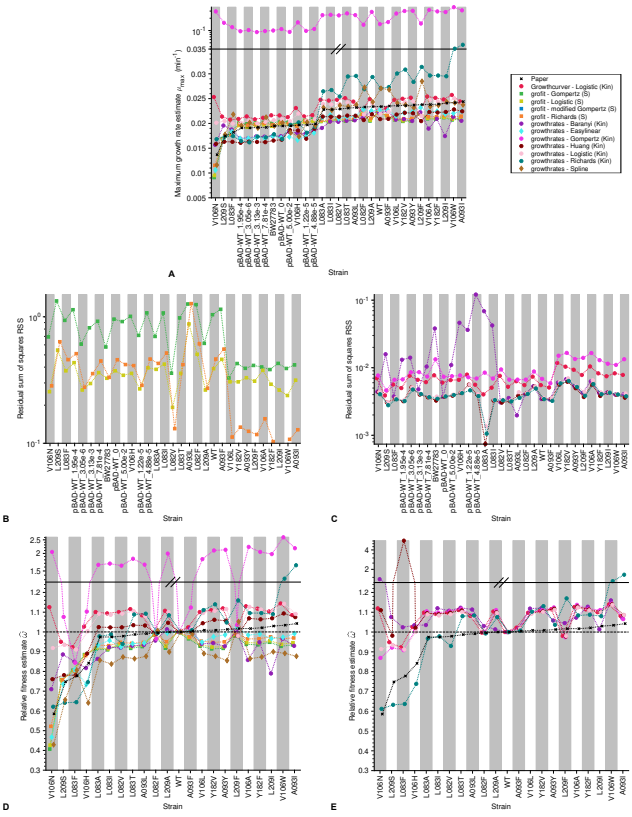


Figure 5: **Analysis of published data-sets (Adkar et al., 2017) shows that estimates differ vastly depending on the method:** A) Maximum growth rate $\widehat{\mu}_{\max}$ estimate for each strain. Each growth curve was analyzed using three different R packages including both model-free and model-based methods. The crosses show the values reported in the paper, the circles are obtained by methods based on mechanistic models, the squares by methods based on phenomenological models and the diamonds by model-free methods. B) Goodness-of-fit measured as residual sum of squares (RSS) by strain for the phenomenological models. C) RSS by strain for the mechanistic models. We include the RSS of phenomenological and mechanistic models in different plots as the scale of the y-axes differ between these models: phenomenological models operate on a logarithmic scale since $y = \ln(N/N_0)$ but mechanistic models operate on a linear scale $N(t)$. **Relative fitness:** D) and E) Relative fitness estimate $\widehat{\omega}$ versus strain. In D) the relative fitness is computed using $\mu_{\max}$ whereas in E) it is computed using μ.

Implications of data re-analysis

Our analysis indicates that the choice of the best method to analyze growth curve data is very difficult. Especially, identifying the 'right' model that best fits all strains/treatments within a data-set seems daunting. This difficulty might explain why there exists such a diversity of methods for analyzing growth curve data, as we found in the literature review. Interestingly, we saw that most articles only report using one method for analyzing their data. We suspect that different labs and researchers have their own preferences and habits on how to obtain growth rate estimates. In the interest of time (and sanity), researchers may be using the model and method they know best and for which they have obtained reasonable-looking results in the past, rather than trying out a multitude of unfamiliar computational tools.

One clear finding from our re-analysis of published data is that the Gompertz family of models – both phenomenological and mechanistic – are usually not the best choice. This point was previously made by López et al. (2004). We corroborate their empirical results with mathematical arguments. However, our literature review showed that the Gompertz models have remained popular long after López et al.'s study in 2004. We recommend that experimenters fit and compare more than one model when analyzing data. In the light of our results, it will be important to develop one easy-to-use framework that allows for model choice and comparison, which would easily single out inappropriate models.

Our results confirmed the finding of (Peleg and Corradini, 2011) that standard statistical techniques for model selection were often unhelpful (Figure 5B-C): when comparing the fit of different models to the same data, the goodness-of-fit statistics did not always select the model that looked best upon visual inspection and/or there was not much difference between models in terms of goodness-of-fit. This corroborates our personal communications with empiricists that they are reluctant to use models for fitting their growth curve data.

3 Recommendations & Conclusions

Despite the long-established study of batch culture growth curve data, estimating growth rates is still not straightforward. Using a literature review, math, simulations, and analysis of previously published data, our work highlights experimental and theoretical pitfalls encountered by many researchers who work with batch monoculture growth curves. To that end, we have summarized our recommendations for better growth rate estimates as a checklist in Figure 6.

General recommendations:

We urge readers to remember that the intrinsic growth rate μ, a model parameter estimated from mechanistic models, is not the same as the maximum growth rate $\mu_{\max}$, a summary-statistic estimated using phenomenological models or by model-free methods. Although $\mu_{\max}$ is often used as a proxy for μ, this assumption is not always justified.

We recommend that researchers make their raw data and methods available and reproducible. In particular, this involves reporting all experimental parameters like inoculum size, carrying capacity or initial fraction, and lag time. During our literature review, we were surprised by the lack of sufficient information on experimental methods, estimated values, and data availability.

Experimental recommendations:

Good data begin with good experimental methods. In the absence of a lag phase, the fastest

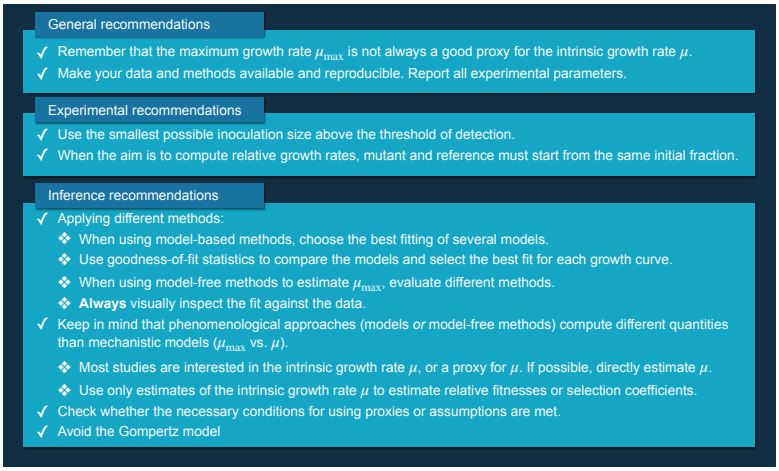


Figure 6: A list of recommendations for better growth rate inference from monoculture growth curves.

growth rate (both for the maximum growth rate $\mu_{\max}$ and for the intrinsic growth rate μ) is observed at the very start of the growth curve. As we show in Figures 3A and 4, the estimates of interest can be highly sensitive to the inoculum size. Also, when a lag phase is suspected, reliable data from the start of the growth curve are necessary to quantify the lag time and growth rate. Therefore, we recommend that experimenters use the smallest inoculation size possible that is still reliably above the threshold of detection (if using a microplate reader, follow the recommendations of Hall et al. 2014).

For accurate estimation of the relative growth rate, we stress that the mutant and reference/wild-type strains should start from the same initial fraction ($N_0^M/K_M = N_0^{WT}/K_{WT}$; Figure 4), which is not necessarily the same absolute size. If the strains have the same carrying capacity ($K_M = K_{WT}$) at stationary phase, then it is possible to either dilute the cultures used for inoculation by the same dilution factor or begin the growth curves at the same absolute inoculum size ($N_0^M = N_0^{WT}$). However, if the carrying capacities differ between strains, then the same absolute inoculum size cannot be used. We found in our literature review that about half of papers use a fixed absolute inoculum size to start their growth curve experiments whereas the other half use a fixed dilution factor, usually without justifying either choice. We recommend that experimenters interested in calculating relative growth rates use a fixed dilution factor. A pilot experiment is ideal to inform on the carrying capacities and corresponding optimal inoculum sizes.

Inference recommendations:
Regarding inference methods, our main recommendation is to try out several different methods on the same data. We recommend that a computational method is used to estimate the growth

rate. (Manual) Fitting of an exponential model should be avoided (Table 3), regardless of whether the exponential model is fit explicitly or implicitly by using the equation $R = (\log_2 N_2 - \log_2 N_1)/(t_2 - t_1)$.

Model-free computational methods tend to be technically easier to use than model-based methods, but they can only estimate $\mu_{\max}$. If a model-free method is used, we recommend to try different programs, ideally based on different algorithms. Model-based methods often require more computational knowledge for fitting, but we still recommend that researchers try to fit more than one model. We recommend this because, by re-analyzing previously published data, we showed that there is not one model that fits best for all data-sets or even one model that best fits every curve within a specific data-set (Figures 5B-C, S1B-C, S2B-C and S3B-C). Therefore, it is important to select the best fitting model for each curve. For those familiar with the R statistical programming language, we recommend the growthrates package, because all of the mechanistic models presented in Table 1, as well as model-free methods for estimating $\mu_{\max}$, can easily be fitted with this package. We recommend that researchers use goodness-of-fit summary statistics (like the RSS, AIC, etc.) to compare models and select the best fit. However, additional visual inspection of the estimated value and the data is essential because goodness-of-fit statistics can be misleading (e.g., Figure S6).

Our work explains and demonstrates that and why the maximum growth rate $\mu_{\max}$ is different from the intrinsic growth rate μ, which is a key point to take away and remember from this paper. We suggest that researchers attempt to estimate μ directly from their data using mechanistic models. However, we have shown that it can be justified to use the phenomenological quantity $\mu_{\max}$ as a proxy for the model parameter μ if certain conditions are met. We recommend that experimenters decide which one makes more sense to use for the experimental question at hand and, based on that decision, select the types of models or model-free approaches to use. If $\mu_{\max}$ is to be inferred, model-free methods or fits of phenomenological models can be used. Model-based methods can either be used to estimate the maximum growth rate $\mu_{\max}$ of phenomenological models, or the intrinsic growth rate μ of mechanistic models. We urge experimenters not to compare estimates obtained from mechanistic and phenomenological models because these different model types estimate different growth rate parameters. Researchers should be aware that confusion between the two quantities is common and different authors/fields use different naming conventions. We hope that this paper provides readers with the necessary conceptual understanding to critically navigate the literature. Finally, we note that only the intrinsic growth rate μ (and not $\mu_{\max}$ or the exponential approximation) should be used for estimating the relative fitness (Figure 4) from monoculture growth curves.

It is important to verify that the necessary conditions are met for the inference method(s) to be used. For example, the exponential approximation should only be used when the following assumptions are met: the inoculum size is much smaller than the carrying capacity ($N_0 \ll K$, e.g., by at least 2 orders of magnitude), only time points at which the population size remains much smaller than the carrying capacity are considered ($N(t) \ll K$), and the lag time is very short or absent (e.g., $h_0 \ll 1$ for growth following a Baranyi model). Given these restrictive assumptions, rather than demonstrating that the conditions for the exponential approximation are met, it may be more feasible for researchers to directly fit one of the mechanistic models listed in Table 1. Another approximation that requires justification is the use of $\mu_{\max}$ as a proxy for μ. This approximation is only valid for small initial population fractions (Figure 3) and short (or absent) lag times. To demonstrate that this is the case,

it is essential that experimenters report the initial population fraction(s) and the lag time(s) when using $\mu_{\max}$ as a proxy for μ.

Finally, we recommend that researchers avoid fitting the Gompertz model for both absolute and relative estimates of μ. In our study, the mechanistic Gompertz model consistently showed wrong estimates for simulated data (Figures 3-4) and unusually large estimates for empirical data (Figure 5).

4 Methods

4.1 Literature review

We quantified the most frequently used methods of analyzing growth curve data for extracting summary statistics. To this end, we used Web of Knowledge and Google Scholar to search for peer-reviewed papers in the evolution and ecology literature that gathered any type of growth curve data proportional to the number of individuals growing in a homogeneous, liquid culture and estimated growth parameters from those data. We searched papers from 1990-2021 and used the following search terms: "Bioscreen C", growth curve, OD, growth rate, batch culture, bacterial OR microbial. Papers that quantified binary presence/absence of growth (e.g. to assay lag time or spore viability) and papers that investigated biofilms were excluded. Most papers were found because they cited one of the following growth curve methods papers, Zwietering et al. (1990); Hall et al. (2014); Sprouffske and Wagner (2016); or Delaney et al. (2013). From each paper, we extracted information about whether the method used to analyze growth curves is explicitly cited or described, what type(s) of growth curve summary statistic was used, whether a model-free or model-based approach was used, whether the growth rate is from a phenomenological or mechanistic approach, which model(s) were fitted (if no equation is given, then the name of the model as reported by the author), whether the growth curves were inoculated from a fixed starting value or as a fraction of the carrying capacity, and whether the growth curve raw data are publicly available or, at least, plotted (summarized in table S1). For all papers in which growth curves were inoculated using a fraction from overnight cultures, the dilution factor was used as an estimator of the initial dilution fraction. If given, we extracted the initial dilution factor and any accompanying information about the inoculum (e.g., length of overnight culture) to indicate whether the dilution factor is a good proxy for the initial population fraction (N_0/K). For papers in which growth curves were inoculated using a fixed absolute number of cells, we report the dilution fraction only if sufficient information about the inoculum size and carrying capacity was provided in the methods. Finally, we categorized different model-free growth rate estimation methods that were applied *ad hoc* as either "Easy Linear" if a consistent method was given for selecting which points to include in the regression (since this is the main feature of popular model-free methods like that of Hall et al. 2014), or as "exponential approximation" if there was no information about which points were included in the regression or as "spline" if pairs of successive measurements were used to estimate the local slope of the curve.

4.2 Analyzing 4 published data-sets

Data-sets appropriate for our analysis were found during our literature review and the data were accessed as indicated in each paper.

We used the following R (version 4.1.1) packages to re-analyze the data: Growthcurver (version 0.3.1), grofit (version 1.1.1-1) and growthrates (version 0.8.2). Each of them was downloaded from the CRAN repository except grofit that we obtained from Kahm et al. (2010). Indeed the latter was found to be removed from the CRAN repository. The package Growthcurver is based on the mechanistic logistic model, whereas grofit includes four phenomenological models (Logistic, Gompertz, modified Gompertz and Richards). The package growthrates provide both model-free methods (Easy Linear and Spline) as well as methods based on mechanistic models (Logistic, Gompertz, Richards, Baranyi and Huang).

We analyzed 143 population growth curves (31 from Adkar et al. 2017; 6 from Ram et al. 2019; 66 from Todd and Selmecki 2020; and 40 from Hammer et al. 2021) using all methods mentioned above. We focused on the maximum growth rate $\mu_{\max}$, because it is the only quantity common to all models and methods. Since the mechanistic models are defined based on the intrinsic growth rate μ, we used Table 1 to calculate the maximum growth rate $\mu_{\max}$ from the respective model.

To test the accuracy of the fits obtained by the model-based methods, we calculated the residual sum of squares (RSS). We used the definition from López et al. (2004):

$$\mathrm{RSS} = \sum_{i=1}^{n} (OD_i - \widehat{OD}_i)^2 \,.$$

Here, n is the number of data points, OD_i is the i^{th} optical density value to be estimated and $\widehat{OD}_i$ is the i^{th} estimated optical density value. Since the models have different numbers of parameters, we also calculated the Akaike's Information Criterion (AIC) for the mechanistic models as given in López et al. (2004) and explained in the supplementary methods.

4.3 Simulations

We generated data representing the dynamics of microbial populations using a Gillespie algorithm for the mechanistic Gompertz, Richards and Logistic models (Gillespie, 1976, 1977). For the Baranyi and Huang models, a modified Next-Reaction algorithm was required since these models have time-dependent growth rates (Anderson, 2007). All simulation code was written in C and is available at https://github.com/LcMrc/GrowthRates. We detail below the algorithms used.

Gillespie algorithm: Let us denote by N the number of individuals. The only elementary event that can happen is division of a microbe, whose rate is denoted by $k_{N\to N+1}$. Let us note that $k_{N\to N+1} = \mu \log(K/N)N$, $k_{N\to N+1} = \mu(1 - N/K)N$ and $k_{N\to N+1} = \mu(1 - (N/K)^{\beta})N$ for the mechanistic Gompertz, Logistic and Richards models, respectively. Simulation steps are as follows:

1. Initialization: The population starts from N_0 microorganisms at time $t = 0$.
2. Time update: The time increment Δt is sampled randomly from an exponential distribution with mean $1/k_{N\to N+1}$ and the time t is updated such that $t \leftarrow t + \Delta t$.
3. Number of individuals update: a division occurs and the population size N increases by one such that $N \leftarrow N + 1$.

4. We go back to Step 2 and iterate until the desired time limit is reached.

Next-Reaction algorithm: Let us denote by N the number of individuals. The only elementary event that can happen is division of a microbe, whose time-dependent rate is denoted by $k_{N\to N+1}(t)$. Let us note that $k_{N\to N+1}(t) = \mu e^{\mu t}(1 - N/K)N/(e^{h_0} - 1 + e^{\mu t})$ and $k_{N\to N+1}(t) = \mu(1 - N/K)N/(1 + e^{\alpha(t-\tau)})$ for the mechanistic Baranyi and Huang models, respectively. In the following, we will denote by P the first firing time and T the internal time.

1. Initialization: The population starts from N_0 microorganisms at time $t = 0$. The first firing time P is sampled from an exponential distribution of mean 1 and the internal time T is set to 0.
2. Time update: The time increment Δt is computed solving $\int_t^{t+\Delta t} k_{N\to N+1}(u)\mathrm{d}u = P - T$ and the time t is updated such that $t \leftarrow t + \Delta t$.
3. Number of individuals update: a division occurs and the population size N increases by one such that $N \leftarrow N + 1$.
4. Internal time update: The internal time T is updated such that $T \leftarrow T + \Delta T$, where $\Delta T = \int_t^{t+\Delta t} k_{N\to N+1}(u)\mathrm{d}u$.
5. First firing time update: The first firing time P is updated such that $P \leftarrow P + \Delta P$, where ΔP is sampled from an exponential distribution of mean 1.
6. We go back to Step 2 and iterate until the desired time limit is reached.

4.4 Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article or Supplemental Material. Annotated C implementations of numerical simulations, annotated code to reproduce all computationally produced graphs, and additional figures and tables reporting the data re-analysis fits and estimates are available at https://github.com/LcMrc/GrowthRates.

5 Author Contributions

All authors designed the study; AHG performed the literature review; LM performed the numerical and analytical work; all authors analyzed and interpreted the data; all authors wrote and edited the manuscript.

6 Acknowledgements

The authors thank the Evolutionary Biology group at IGC for stimulating discussions and the THEE Group at UniBe for feedback on the manuscript. AHG thanks FCT PhD funding grant PD/BD/138215/2018. CB is grateful for funding from ERC Starting Grant 804569 (FIT2GO) and SNSF Project Grant "MiCo4Sys".

References

Tanita Wein and Tal Dagan. The effect of population bottleneck size and selective regime on genetic diversity and evolvability in bacteria. *Genome biology and evolution*, 11(11): 3283–3290, 2019.

Lindi M Wahl and Anna Dai Zhu. Survival probability of beneficial mutations in bacterial batch culture. *Genetics*, 200(1):309–320, 2015. doi: 10.1534/genetics.114.172890.

Thomas E Miller, Jean H Burns, Pablo Munguia, Eric L Walters, Jamie M Kneitel, Paul M Richards, Nicolas Mouquet, and Hannah L Buckley. A critical review of twenty years' use of the resource-ratio theory. *The American Naturalist*, 165(4):439–448, 2005.

Joey R Bernhardt, Pavel Kratina, Aaron Louis Pereira, Manu Tamminen, Mridul K Thomas, and Anita Narwani. The evolution of competitive ability for essential resources. *Philosophical Transactions of the Royal Society B*, 375(1798):20190247, 2020.

J Monod. The growth of bacterial cultures. *Annual Review of Microbiology*, 3(1):371—-394, 1949. doi: 10.1146/annurev.mi.03.100149.002103.

Jeniffer Concepción-Acevedo, Howard N Weiss, Waqas Nasir Chaudhry, and Bruce R Levin. Malthusian parameters as estimators of the fitness of microbes: a cautionary tale about the low side of high throughput. *PloS one*, 10(6):e0126915, 2015.

Yoav Ram, Eynat Dellus-Gur, Maayan Bibi, Kedar Karkare, Uri Obolski, Marcus W. Feldman, Tim F. Cooper, Judith Berman, and Lilach Hadany. Predicting microbial growth in a mixed culture from growth curve data. *Proceedings of the National Academy of Sciences*, 116(29):14698–14707, 2019. ISSN 0027-8424. doi: 10.1073/pnas.1902217116. URL `https://www.pnas.org/content/116/29/14698`.

Michael Knopp and Dan I Andersson. Predictable phenotypes of antibiotic resistance mutations. *MBio*, 9(3):e00770–18, 2018.

Arthur Slator. Ii. the rate of growth of bacteria. *J. Chem. Soc., Trans.*, 109:2–10, 1916. doi: 10.1039/CT9160900002. URL `http://dx.doi.org/10.1039/CT9160900002`.

Tor Carlson. Über geschwindigkeit und grösse der hefevermehrung in würze. *Biochem. Ztschr. Bd.*, 57:313–334, 1913.

Nigel F Delaney, Maria E Kaczmarek, Lewis M Ward, Paige K Swanson, Ming-Chun Lee, and Christopher J Marx. Development of an optimized medium, strain and high-throughput culturing methods for *Methylobacterium extorquens*. *PLoS One*, 8(4):e62957, 2013.

Barry G Hall, Hande Acar, Anna Nandipati, and Miriam Barlow. Growth rates made easy. *Molecular biology and evolution*, 31(1):232–238, 2014.

Keiran Stevenson, Alexander F McVey, Ivan BN Clark, Peter S Swain, and Teuta Pilizota. General calibration of microbial growth in microplate readers. *Scientific reports*, 6(1):1–7, 2016.

Masaomi Kurokawa and Bei-Wen Ying. Precise, high-throughput analysis of bacterial growth. *Journal of Visualized Experiments: JoVE*, 127:e56197, 2017.

M. H. Zwietering, I. Jongenburger, F. M. Rombouts, and K. Van't Riet. Modeling of the bacterial growth curve. *Applied and environmental microbiology*, 56(6):1875–1881, 1990.

József Baranyi and Terry A. Roberts. A dynamic approach to predicting bacterial growth in food. *International Journal of Food Microbiology*, 23(3):277–294, 1994. ISSN 0168-1605. doi: https://doi.org/10.1016/0168-1605(94)90157-0. URL `https://www.sciencedirect.com/science/article/pii/0168160594901570`. Special Issue Predictive Modelling.

Paul P Jung, Nils Christian, Daniel P Kay, Alexander Skupin, and Carole L Linster. Protocols and programs for high-throughput growth and aging phenotyping in yeast. *PloS one*, 10 (3):e0119807, 2015.

Anders Blomberg. Measuring growth rate in high-throughput growth phenotyping. *Current Opinion in Biotechnology*, 22(1):94–102, 2011.

Kathleen Sprouffske and Andreas Wagner. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinformatics*, 17(1):172, 4 2016. ISSN 1471-2105. doi: 10.1186/s12859-016-1016-7. URL `https://doi.org/10.1186/s12859-016-1016-7`.

Thomas Petzoldt. *growthrates: Estimate Growth Rates from Experimental Data*, 2020. URL `https://CRAN.R-project.org/package=growthrates`. R package version 0.8.2; A collection of methods to determine growth rates from experimental data, in particular from batch experiments and plate reader trials.

Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The FAIR guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.

Marcus R Munafò, Brian A Nosek, Dorothy VM Bishop, Katherine S Button, Christopher D Chambers, Nathalie Percie du Sert, Uri Simonsohn, Eric-Jan Wagenmakers, Jennifer J Ware, and John Ioannidis. A manifesto for reproducible science. *Nature human behaviour*, 1(1):1–9, 2017.

Erik Gullberg, Sha Cao, Otto G Berg, Carolina Ilbäck, Linus Sandegren, Diarmaid Hughes, and Dan I Andersson. Selection of resistant bacteria at very low antibiotic concentrations. *PLOS Pathogens*, 7(7):e1002158, 2011.

Sandra Trindade, Ana Sousa, and Isabel Gordo. Antibiotic resistance and stress in the light of fisher's model. *Evolution*, 66(12):3815–3824, 2012.

Luciano Fernandez-Ricaud, Olga Kourtchenko, Martin Zackrisson, Jonas Warringer, and Anders Blomberg. Precog: a tool for automated extraction and visualization of fitness components in microbial growth phenomics. *BMC Bioinformatics*, 17(1):1–15, 2016.

Portia Mira, Miriam Barlow, Juan C Meza, and Barry G Hall. Statistical package for growth rates made easy. *Molecular biology and evolution*, 34(12):3303–3309, 2017. doi: https://doi.org/10.1093/molbev/msx255.

Kazuha Ashino, Kenta Sugano, Toshiyuki Amagasa, and Bei-Wen Ying. Predicting the decision making chemicals used for bacterial growth. *Scientific reports*, 9:7251, 2019. doi: https://doi.org/10.1038/s41598-019-43587-8.

Matthias Kahm, Guido Hasenbrink, Hella Lichtenberg-Fraté, Jost Ludwig, and Maik Kschischo. grofit: Fitting biological growth curves with R. *Journal of Statistical Software*, 33(7):1–21, 2010. doi: 10.18637/jss.v033.i07. URL `https://www.jstatsoft.org/index.php/jss/article/view/v033i07`.

Sarah P Otto and Troy Day. *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press, 2011.

Benjamin M Bolker. *Ecological models and data in R*. Princeton University Press, 2008.

David A Liberles, Ashley I Teufel, Liang Liu, and Tanja Stadler. On the need for mechanistic models in computational genomics and metagenomics. *Genome biology and evolution*, 5 (10):2008–2018, 2013.

Benoît Chezeau and Christophe Vial. *Modeling and Simulation of the Biohydrogen Production Processes*, chapter 19, pages 445–483. Elsevier B.V. All, 2019. doi: 10.1016/C2017-0-03531-0.

Nicolas Rodrigue and Hervé Philippe. Mechanistic revisions of phenomenological modeling strategies in molecular evolution. *Trends in Genetics*, 26(6):248–252, 2010.

A. Tsoularis and J. Wallace. Analysis of logistic growth models. *Mathematical Biosciences*, 179(1):21–55, 2002. ISSN 0025-5564. doi: https://doi.org/10.1016/S0025-5564(02)00096-2. URL `https://www.sciencedirect.com/science/article/pii/S0025556402000962`.

Lihan Huang. A new mechanistic growth model for simultaneous determination of lag phase duration and exponential growth rate and a new bĕlehdrádek-type model for evaluating the effect of temperature on growth rate. *Food Microbiology*, 28(4):770–776, 2011. ISSN 0740-0020. doi: https://doi.org/10.1016/j.fm.2010.05.019. URL `https://www.sciencedirect.com/science/article/pii/S0740002010001255`. Predictive Modeling in Foods.

Stefano Perni, Peter W. Andrew, and Gilbert Shama. Estimating the maximum growth rate from microbial growth curves: definition is everything. *Food Microbiology*, 22(6): 491–495, 2005. ISSN 0740-0020. doi: https://doi.org/10.1016/j.fm.2004.11.014. URL `https://www.sciencedirect.com/science/article/pii/S0740002004001522`.

Ru-Jun Yang, Xiu-Lin Wang, Ying-Ying Zhang, and Yu-Jie Zhan. Influence of cell equivalent spherical diameter on the growth rate and cell density of marine phytoplankton. *Journal of experimental marine biology and ecology*, 331(1):33–40, 2006.

Xueqing Wu, Nathan T Jacobs, Catherine Bozio, Preston Palm, Santiago M Lattar, Christiane R Hanke, David M Watson, Fuminori Sakai, Bruce R Levin, Keith P Klugman, et al. Competitive dominance within biofilm consortia regulates the relative distribution of pneumococcal nasopharyngeal density. *Applied and environmental microbiology*, 83(16): e00953–17, 2017.

Sadia Khan, Tara K Beattie, and Charles W Knapp. The use of minimum selectable concentrations (MSCs) for determining the selection of antimicrobial resistant bacteria. *Ecotoxicology*, 26(2):283–292, 2017.

Bharat V. Adkar, Michael Manhart, Sanchari Bhattacharyya, Jian Tian, Michael Musharbash, and Eugene I. Shakhnovich. Optimization of lag phase shapes the evolution of a bacterial enzyme. *Nature Ecology & Evolution*, 1(6):0149, 4 2017. ISSN 2397-334X. doi: 10.1038/s41559-017-0149. URL https://doi.org/10.1038/s41559-017-0149.

Donatella Ganucci, Simona Guerrini, Silvia Mangani, Massimo Vincenzini, and Lisa Granchi. Quantifying the effects of ethanol and temperature on the fitness advantage of predominant *Saccharomyces cerevisiae* strains occurring in spontaneous wine fermentations. *Frontiers in microbiology*, 9:1563, 2018.

Nigel Delaney. Curve Fitter - software for growth curve fitting and visualization, 2014. http://www.evolvedmicrobe.com/CurveFitter/ [Accessed: 2021-11-23].

R. Kassen. *Experimental Evolution and the Nature of Biodiversity*. Macmillan Learning, 2014. ISBN 9781936221462. URL https://books.google.ch/books?id=LL8YnwEACAAJ.

Prabh Basra, Ahlam Alsaadi, Gabriela Bernal-Astrain, Michael Liam O'Sullivan, Bryn Hazlett, Leah Marie Clarke, Andrew Schoenrock, Sylvain Pitre, and Alex Wong. Fitness tradeoffs of antibiotic resistance in extraintestinal pathogenic *Escherichia coli*. *Genome biology and evolution*, 10(2):667–679, 2018.

Maja Novak, Thomas Pfeiffer, Martin Ackermann, and Sebastian Bonhoeffer. Bacterial growth properties at low optical densities. *Antonie Van Leeuwenhoek*, 96(3):267–274, 2009.

Richard E. Lenski, Michael R. Rose, Suzanne C. Simpson, and Scott C. Tadler. Long-term experimental evolution in escherichia coli. i. adaptation and divergence during 2,000 generations. *The American Naturalist*, 138(6):1315–1341, 1991. URL http://www.jstor.org/stable/2462549. Full publication date: Dec., 1991.

J.F. Crow and M. Kimura. *An Introduction to Population Genetics Theory*. Blackburn Press, 2009. ISBN 9781932846126. URL https://books.google.pn/books?id=VWqKPwAACAAJ.

Luis-Miguel Chevin. On measuring selection in experimental evolution. *Biology Letters*, 7(2):210–213, 2011. doi: 10.1098/rsbl.2010.0580. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsbl.2010.0580.

Tobin J. Hammer, Eli Le, and Nancy A. Moran. Thermal niches of specialized gut symbionts: the case of social bees. *Proceedings of the Royal Society B: Biological Sciences*, 288(1944):20201480, 2021. doi: 10.1098/rspb.2020.1480. URL https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2020.1480.

Robert T Todd and Anna Selmecki. Expandable and reversible copy number amplification drives rapid adaptation to antifungal drugs. *eLife*, 9:e58349, 7 2020. ISSN 2050-084X. doi: 10.7554/eLife.58349. URL https://doi.org/10.7554/eLife.58349.

S. López, M. Prieto, J. Dijkstra, M.S. Dhanoa, and J. France. Statistical evaluation of mathematical models for microbial growth. *International Journal of Food Microbiology*, 96 (3):289–300, 2004. ISSN 0168-1605. doi: https://doi.org/10.1016/j.ijfoodmicro.2004.03.026. URL https://www.sciencedirect.com/science/article/pii/S0168160504002041.

Micha Peleg and Maria G Corradini. Microbial growth curves: what the models tell us and what they cannot. *Critical reviews in food science and nutrition*, 51(10):917–945, 2011.

Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976. ISSN 0021-9991. doi: https://doi.org/10.1016/0021-9991(76)90041-3. URL https://www.sciencedirect.com/science/article/pii/0021999176900413.

Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977. doi: 10.1021/j100540a008. URL https://doi.org/10.1021/j100540a008.

David F. Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *The Journal of Chemical Physics*, 127(21):214107, 2007. doi: 10.1063/1.2799998. URL https://doi.org/10.1063/1.2799998.

Supplement for: Challenges and pitfalls of inferring microbial growth rates from lab cultures

Ana-Hermina Ghenu*[1,2,3], Loïc Marrec†[2,3], and Claudia Bank[1,2,3]

[1]Instituto Gulbenkian de Ciência, Rua da Quinta Grande 6, Oeiras, 2780-156, Portugal
[2]Institut für Ökologie und Evolution, Universität Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland
[3]Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland

October 9, 2023

1 Supplementary methods

As explained in the main text, we re-analyzed the population growth curves from previously published data sets by fitting many different models and calculating the residual sum of squares (RSS). Since the models have different numbers of parameters, we also calculated the Akaike's Information Criterion (AIC) for the mechanistic models giving $N(t)$, which reads

$$\mathrm{AIC} = n \ln\left(\frac{\mathrm{RSS}}{n}\right) + 2(p+1) + \frac{2(p+1)(p+2)}{n-p-2}\,, \tag{1}$$

where p is the number of parameters (López et al. 2004).

However, this equation may not be valid for the phenomenological models because these models are on a logarithmic scale ($y = \ln(N/N_0)$) and therefore the errors around the data are probably not normally distributed as assumed by López et al. 2004.

*Co-first and co-corresponding author: hermina.ghenu@gmail.com
†Co-first and co-corresponding author: loic.marrec@iee.unibe.ch

2 Supplementary figures

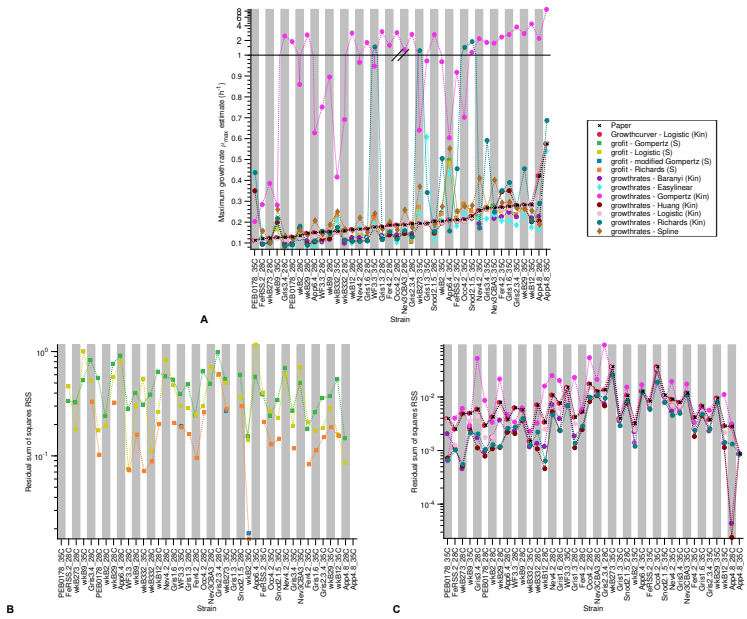


Figure S1: **Analysis of published data sets - Hammer, Le, and Moran 2021**: A) Maximum growth rate $\widehat{\mu}_{\max}$ estimate versus strain. Each growth curve was analyzed using three different R packages including both model-free and model-based methods. The crosses show the values reported in the paper, the circles are obtained by methods based on mechanistic models, the squares by methods based on phenomenological models and the diamonds by model-free methods. B) Residual sum of squares RSS versus strain for the phenomenological models. C) Residual sum of squares RSS versus strain for the mechanistic models. In the figure legend, (Kin) refers to mechanistic models and (S) refers to phenomenological methods.

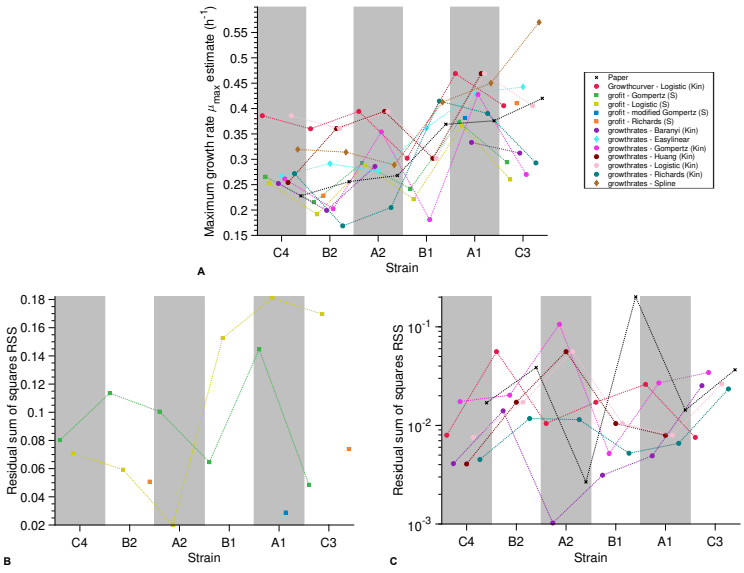


Figure S2: **Analysis of published data sets - Ram2019**: A) Maximum growth rate $\widehat{\mu}_{\max}$ estimate versus strain. Each growth curve was analyzed using three different R packages including both model-free and model-based methods. The crosses show the values reported in the paper, the circles are obtained by methods based on mechanistic models, the squares by methods based on phenomenological models and the diamonds by model-free methods. B) Residual sum of squares RSS versus strain for the phenomenological models. C) Residual sum of squares RSS versus strain for the mechanistic models. In the figure legend, (Kin) refers to mechanistic models and (S) refers to phenomenological methods.

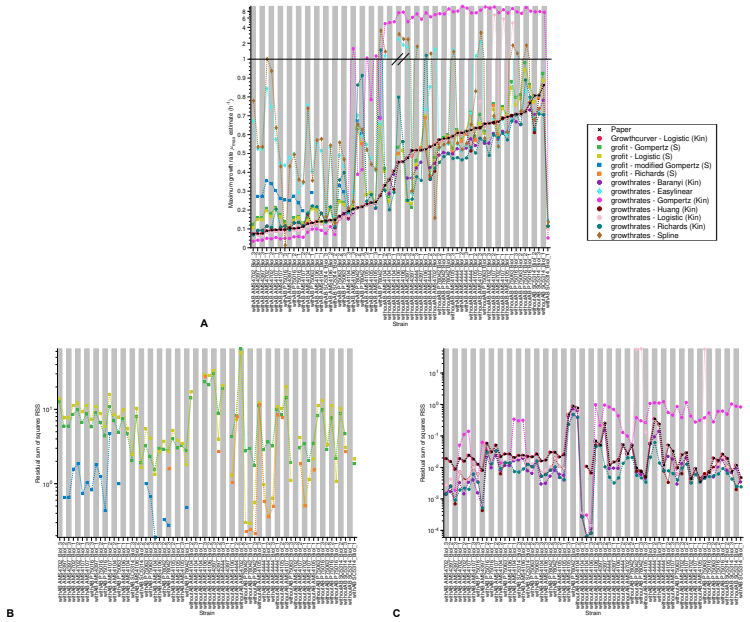


Figure S3: **Analysis of published data sets - Todd and Selmecki 2020**: A) Maximum growth rate $\widehat{\mu}_{\max}$ estimate versus strain. Each growth curve was analyzed using three different R packages including both model-free and model-based methods. The crosses show the values reported in the paper, the circles are obtained by methods based on mechanistic models, the squares by methods based on phenomenological models and the diamonds by model-free methods. B) Residual sum of squares RSS versus strain for the phenomenological models. C) Residual sum of squares RSS versus strain for the mechanistic models. In the figure legend, (Kin) refers to mechanistic models and (S) refers to phenomenological methods.

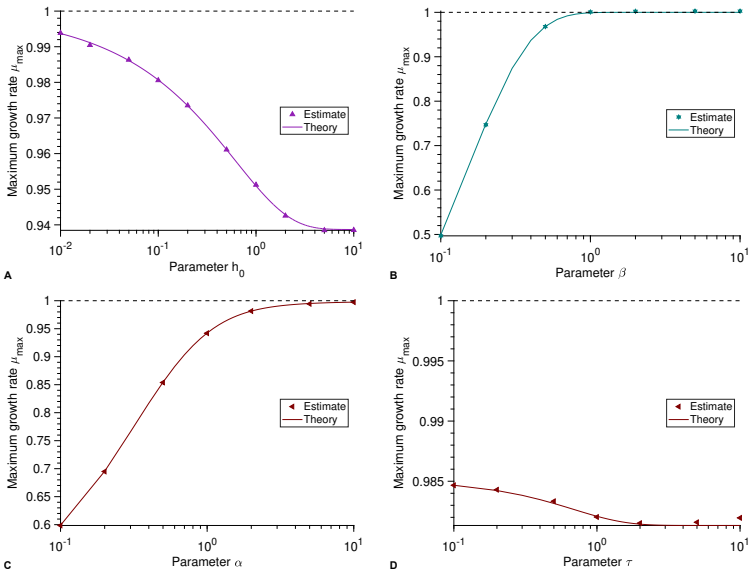


Figure S4: **Maximum growth rate:** A) Maximum growth rate $\mu_{\max}$ versus parameter h_0 for the Baranyi model. B) Maximum growth rate $\mu_{\max}$ versus parameter β for the Richards model. C) Maximum growth rate $\mu_{\max}$ versus parameter α for the Huang model. D) Maximum growth rate $\mu_{\max}$ versus parameter τ for the Huang model. In every panel, each point represents estimated values from simulated data averaged over 10^4 stochastic realizations. The solid lines correspond to the analytical predictions of the maximum growth rate (see Table 2 in the main text). The dashed line shows the intrinsic growth rate value μ. Parameter values: $N_0 = 10^2$, $K = 10^5$, $\mu = 1$, $\alpha = 2$ for panel D and $\tau = 2$ for panel C.

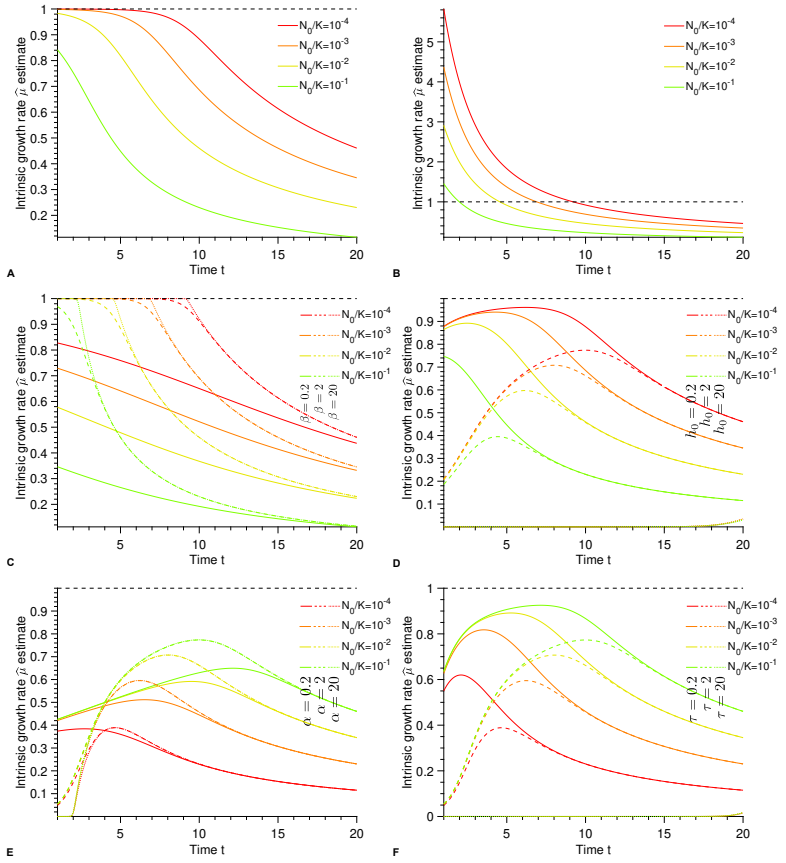


Figure S5: **Intrinsic growth rate:** From A) to F) Intrinsic growth rate $\widehat{\mu}$ estimate versus parameter t for different initial population fractions N_0/K, parameters and mechanistic models (A) Logistic, B) Gompertz, C) Richards, D) Baranyi, E) and F) Huang). The intrinsic growth rate is analytically estimated using the exponential hypothesis $\mu/ = \ln(N(t)/N_0)/t$. The black dashed line shows the real value of μ. Parameter values: $K = 10^5$, μ, $\tau = 2$ for E) and $\alpha = 2$ for F).

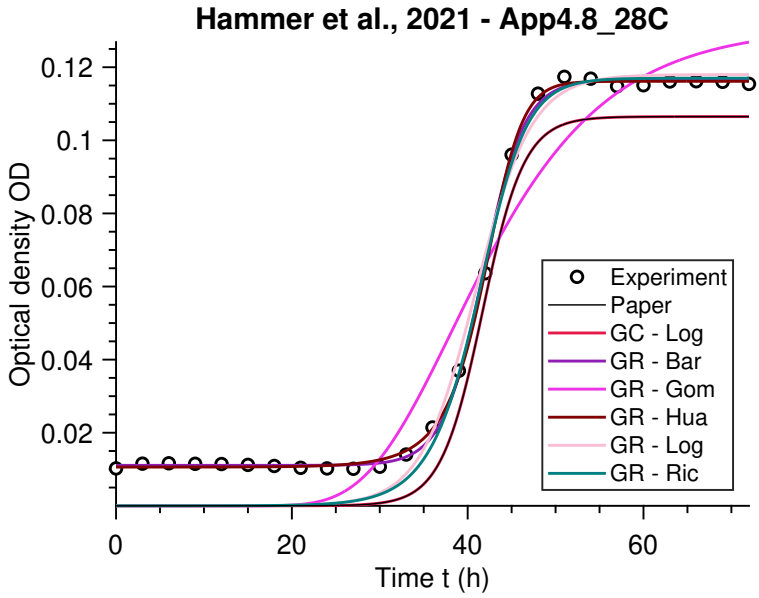
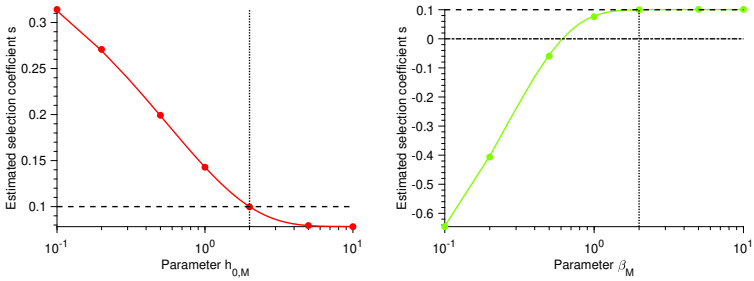


Figure S6: **Analysis of published data:** Optical density OD versus time t. The data points are fitted using different mechanistic models. The black (Paper) and red (Growthcurver - Logistic (Mech)) lines are the same since it was the method applied in the paper.



(a) Baranyi (b) Richards

Figure S7: **Selection coefficient:** a) Estimated selection coefficient $\widehat{s}$ versus parameter h_0 for the Baranyi model. b) Estimated selection coefficient $\widehat{s}$ versus parameter β for the Richards model. In every panel, each point represents estimated values from simulated data averaged over 10^4 stochastic realizations. The solid lines correspond to the analytical predictions of the selection coefficient using those of the maximum growth rate (see Table 2.2.1 in the main text). The dashed line shows the real selection coefficient value s. The dotted line represents the configuration where the parameters of the mutant are equal to those of the wild-type (except their growth rates). The dash-dot line corresponds to the neutral case, i.e. where both the mutant and the wild-type have the same growth rate. Parameter values: $K_W = K_M = 10^5$, $\mu_W = 1$, $\mu_M = 1.1$, $s = 0.1$, $\beta_W = 2$ and $h_{0,W} = 2$. When not specified, β_M and $h_{0,M}$ have the same values as the wild-type.

Chapter 5

Growth traits for predicting antibiotic resistance between environments

[T]he visible effect of a gene substitution depends both on the gene substitution itself and on the genetic complex, or organism, in which this gene substitution is made.
– Ronald Aylmer Fisher, 1930

This chapter is a work in progress. The student's contributions to the work are as follows: Claudia Bank, Isabel Gordo, and the student designed the research. The student gathered the data. The student began analyzing the data with consultation from Claudia Bank and Isabel Gordo. The student wrote the manuscript draft.

Growth traits for predicting antibiotic resistance between environments

Ana-Hermina Ghenu, Isabel Gordo, Claudia Bank

1 Introduction

Antibiotic concentration and temperature are two clinically relevant environmental variables (Savageau, 1983; Andersson and Hughes, 2014; Bell et al., 2014; Durso and Cook, 2014; Blount, 2015) that have, independently, been extensively studied in *E. coli*. Specifically, the dose response curves of many ABR genotypes have been characterised (Mira et al., 2015; Harmand et al., 2017) under optimal temperature conditions ($37^{\circ}C$). Many temperature dose response mechanistic models of Malthusian growth rates have been estimated for *E. coli* (Ratkowsky et al., 1982; Ratkowsky et al., 1983; Zwietering et al., 1990; Daughtry, Davey, and King, 1997), as well as models that predict lag phase (Swinnen et al., 2005) and carrying capacity (Zwietering et al., 1991; Zwietering et al., 1994) or interactions with other environmental variables (Rosso et al., 1995; Salter et al., 2000; Bidlas and Lambert, 2008). Finally, some work has been done looking at effects between high temperature and antibiotic environments. High temperature adapted *E. coli* strains can acquire fitness gains in other environments, like rifampicin ABR by modifying the antibiotic target *rpoB* (Rodríguez-Verdugo et al., 2014). These cross-environment effects are the result of modifications to *rpoB* during high temperature adaptation having pleiotropic effects on metabolism (Bennett and Lenski, 1996) and stress response (Rodríguez-Verdugo, Tenaillon, and Gaut, 2016; Hug and Gaut, 2015).

In Chapter 3 we investigated how genotype-to-fitness landscapes change across environmental gradients. We found that the fitness landscapes became more smooth as the environment deteriorated. In particular, we concluded that non-specific epistasis (i.e., epistasis that is due to reasons *other than* direct interactions, like protein-protein interactions) was responsible for the results that we found. This is because the combinations of mutations we selected are not known the interact with one an-

other. One reason for non-specific epistasis can be that there is a non-linear mapping in the phenotype space (Bank, 2022). In addition, although we detected G$\times$E effects in our previous work, we did not investigate anything more about the nature of this interaction. Often G$\times$E effects can be due to phenotypic plasticity. This means that the same genotype can exhibit multiple phenotypes depending on its environment. For both of these reasons, it is interesting to investigate the phenotype-to-fitness landscape. The goal of this paper is to integrate information from the genotype, phenotype, and fitness in order to better understand how fitness landscapes change across environments.

2 Methods

2.1 Monoculture growth curves

The detailed protocol is shown in figure 1. In summary, the same 24 genotypes under the same environmental conditions were used as for the competition experiments in Chapter 3. Genotypes were grown starting from the same initial inoculum size as for the competition experiments but this time in monoculture. The initial inoculum size was above the threshold of detection for the BioTek H1 microplate reader. All inocula were acclimatized to $37°C$ and the absence of antibiotic, even though the growth curve data were gathered with increased temperatures and/or antibiotic concentrations. Optical density (OD) at 500nm (A500) was gathered every 15 minutes for 48 hours. Then, 10μL of the batch culture was diluted into 190μL of fresh media (1:20 dilution) and growth curves were gathered for an additional 12 hours. Hereafter, this second, diluted growth curve is referred to as the "acclimatized growth curve." Each plate contained all 24 genotypes at all four antibiotic concentrations, therefore any significant batch effect at the level of antibiotic treatment should be easy to normalize. However, different temperatures were run on different plates and on different days. Therefore, any batch effects will be confounded with the temperature treatments. Five replicate growth curves were gathered for each genotype under each of the combinations of environments (in total 24 genotypes $*4$ antibiotic concentrations $*3$ temperatures $*5$ replicates $= 1440$ growth curves).

2.2 Preliminary Analysis

Growth curves were inspected and the best method for baselining the data was found to be subtracting the minimum observed OD of that growth curve from all values (baselined OD = raw OD – min(OD) + 0.01).

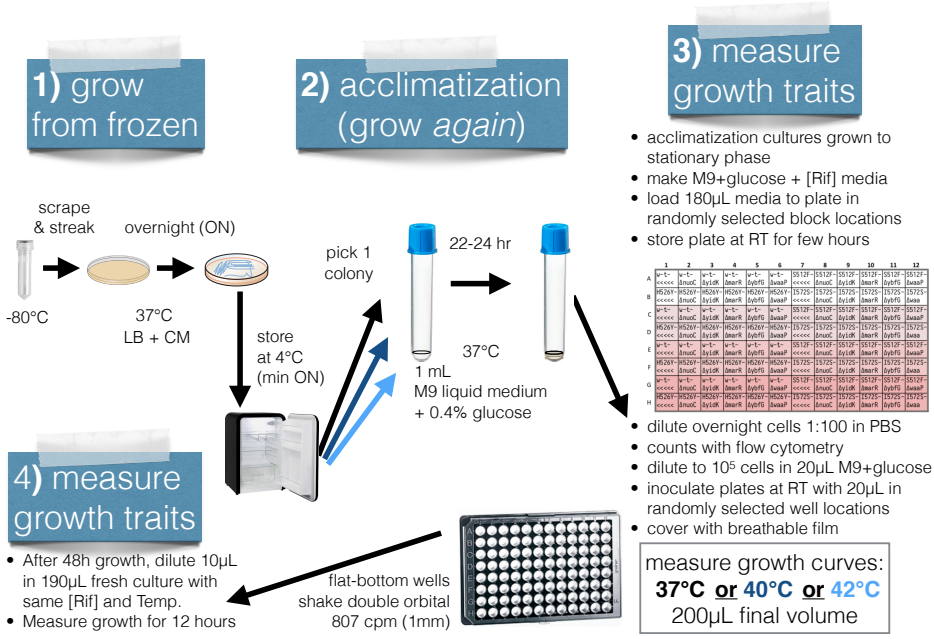


Figure 1: **Summary of the protocol used to estimate growth traits in different antibiotic concentration and temperature environments. 1)** To ameliorate the possibility of *de novo* mutations, frozen glycerol stocks were bottle-necked to single colonies by streaking on agar plates. **2)** After streaking and growth on agar plates, all cells experienced the same temperature acclimatization of at least 2 hours and at most 13 days at 4°C followed by batch culture growth at 37°C for 22-24 hours. **3)** Batch culture competitions took place in 96-well plates with randomized positions of the four blocks of antibiotic concentrations and the 24 samples. Only one plate was incubated at a time, therefore triple arrows show steps in the protocol that had to be repeated for each temperature environment in order to sample a complete block of all samples in all environments. Regardless of the temperature environment where competition occurred, batch culture acclimatization always occurred at 37°C. **4)** After 48 hours, the stationary phase batch culture was diluted 1:20 and grown again for 12 hours. These are referred to as "acclimatized growth curves" in the text.

Preliminary growth traits were extracted from the data. Based on visual inspection, the carrying capacity was defined as the 0.85%ile baselined OD value. The time to reach carrying capacity was defined as the time of occurrence of the 0.85%ile baselined OD value. The area under the curve (AUC) was calculated by summing all baselined OD values up to the time to reach carrying capacity.

ANOVA was used to examine the additive effects of the four different factors on the trait values. The two genotype factors, ABR and KO, were treated as unordered

factors and the two environmental factors, antibiotic concentration and temperature, were treated as ordered factors. I also fitted an ANOVA model with all of the first-order interaction terms. But I did not yet analyze this model for over-fitting so the results are very preliminary.

3 Preliminary Results

Upon visual inspection some growth curves were found to exhibit biphasic growth. It appeared to depend both on the genotype (figure 2) and the environments (figure fig:env). The biphasic growth was stochastic, meaning that it was not always induced, even for the same genotype and environment. The biphasic growth is correlated with the batch (colours in figures 2-3a; colours not shown in figure 3b because the batch and temperature environments are confounded). Biphasic growth was less evident or absent for the acclimatized growth curves (figure 4). I think it is interesting to investigate this biphasic growth. Moreover, as discussed briefly in Chapter 4, biphasic growth will lead to incorrect estimates of the growth rate, especially as this second lag phase always occurs relatively early in the growth curve when replicates displaying monophasic growth are still near the exponential phase. This is why I do not report the growth rate.

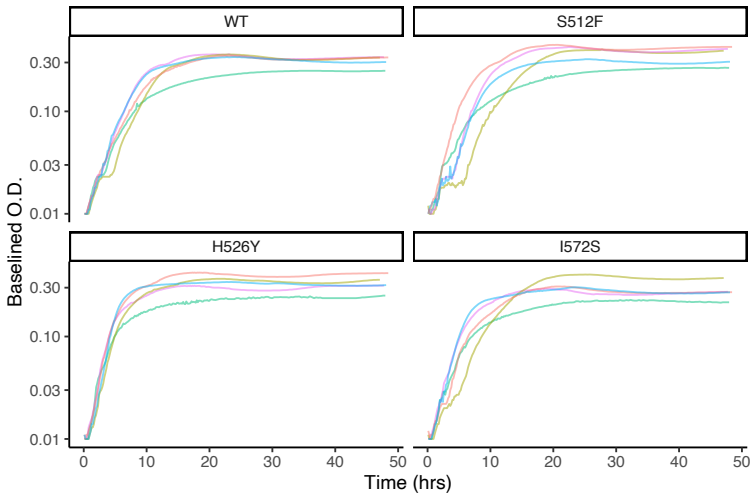
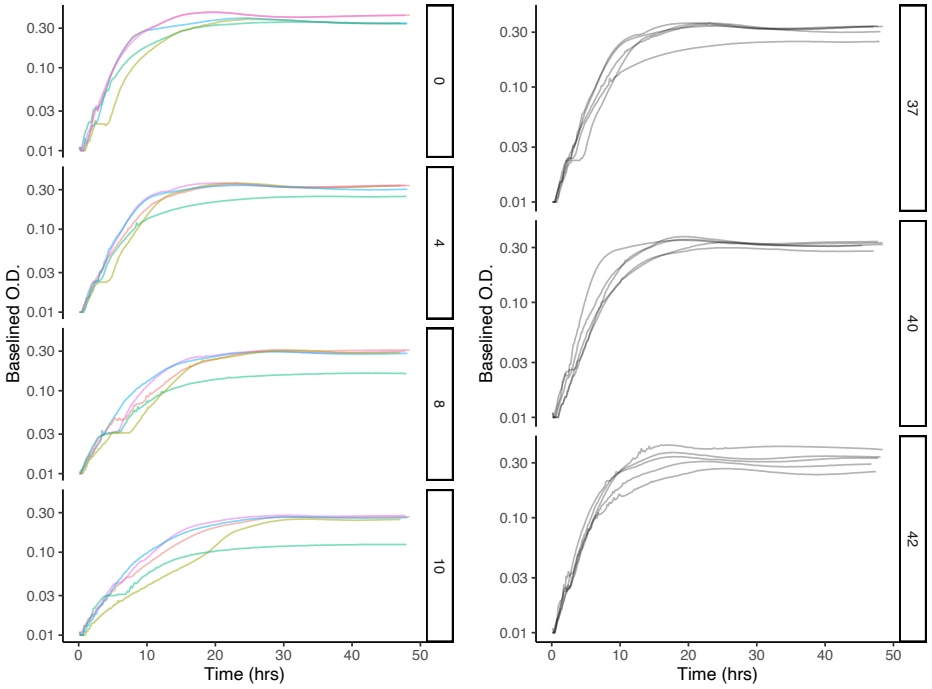


Figure 2: **Biphasic growth depends on the genotype.** Growth curves for different *rpoB* genotypes without any knock-out mutations. The *rpoB* genotype is indicated above each facet. The colours indicate the same batch (but are unlabeled as the batch ID itself is not informative). The environment is 4μg/mL rifampicin and 37°C. Biphasic growth is the shoulder that is seen in the yellow/brown colour for all genotypes except H526Y.



(a) Antibiotic concentration (b) Temperature

Figure 3: **Biphasic growth depends on the environment.** Growth curves for the *rpoB* wild-type (rifampicin-sensitive strain without any knock-out mutations) grown under different environments. The environment is indicated to the right of each facet. *(a)* The environment is variable rifampicin antibiotic concentrations (μg/mL) but fixed 37°C temperature. The colours indicate the same batch. *(b)* The environment is variable temperature (°C) but fixed 4μg/mL rifampicin concentration. The lines are not coloured because batch is confounded with temperature (see methods).

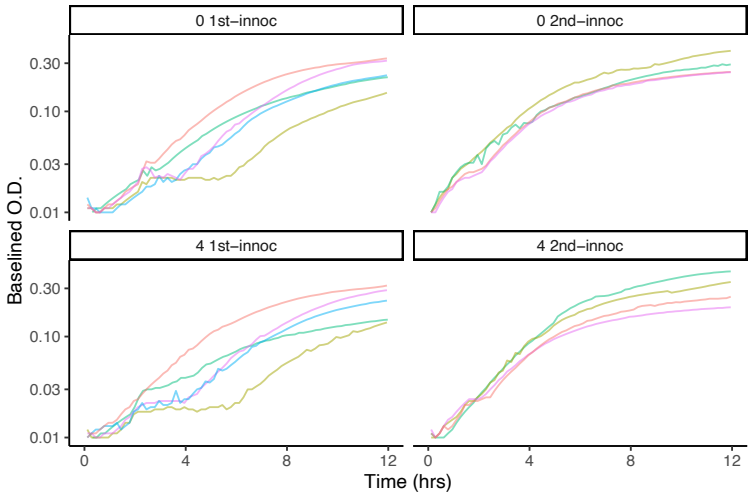


Figure 4: **Biphasic growth depends on acclimatization.** The first 12 hours of growth for the *rpoB* S512F single mutant (i.e., without any knock-out mutations) at 37°C. The left column shows the unacclimatized growth (facets ending in"1st-innoc") and the right column shows the acclimatized growth (facets ending in "2nd-innoc"). The important information from this figure is that the right column shows much less biphasic growth as compared to the left column. The top row shows the growth in the absence of antibiotic (facets starting with "0") and the bottom row shows the growth with 4μg/mL of antibiotic (facets starting with "4"). Note that the inoculate sizes are not standardized in the acclimatized growth; instead, these growth curves are started at a 1:20 dilution of the growth curves on the left after they are well into reach stationary phase at 48 hours (see methods).

The ANOVA for the AUC data with only additive effects is significant ($p < 10^{-15}$) and explains a moderate amount of the variation in the data (adjusted $R^2 = 0.312$). Figure 5 shows the ANOVA results for the AUC data. The main findings from the AUC data are summarized as follows,

- On average across all environments, H526Y grows the best of all *rpoB* genotypes.
- On average across all environments, WT grows the worst of all *rpoB* genotypes.
- On average across all environments, Δ*marR* and Δ*waaP* grow worst of all KO genotypes.
- Antibiotic concentration has a negative effect on the growth of all genotypes on average.
- Higher temperature has a positive effect on the growth of all genotypes on average.
- WT *rpoB* has a negative interaction with antibiotic concentration.
- H526Y grows more poorly than other *rpoB* gentoypes with increasing temperature.
- Δ*waaP* on the WT *rpoB* background grows poorly on average across environments (i.e., it's more susceptible to AB).
- Δ*marR* grows more poorly when in combination with H526Y as compared with other *rpoB* genotypes.
- there is a negative interaction between temperature and antibiotic concentration across all genotypes (but the effect size is pretty small).

The ANOVA for the carrying capacity data with only additive effects is significant ($p < 10^{-15}$; adjusted $R^2 = 0.201$) but explains less of the variation in the data than that for AUC. Figure 6 shows the additive ANOVA results for the carrying capacity data. On the other hand, when first-order interaction terms are added to the model, more of the variation is explained ($p < 10^{-15}$; adjusted $R^2 = 0.361$). Note that we may be over-fitting! Figure 7 shows the ANOVA for the additive terms plus all the first-order interactions. The main findings from the carrying capacity data are summarized as follows,

- WT *rpoB* has a negative interaction with antibiotic concentration.
- Across all environments, S512F grows as well as H536Y, whereas I572S grows the worst for all *rpoB* genotypes.
- No effect of temperature.
- Nevertheless, there is a negative interaction between temperature and antibiotic concentration across all genotypes (small effect size).
- Δ*marR* grows less well than other KO genotypes. In the model with interactions, it shows positive epistasis when combined with the WT *rpoB* genotype.
- Δ*waaP* grows worst of all KO genotypes in the additive model. However, in the model with interactions this is not the case because Δ*waaP* grows fine when in combination with the H526Y *rpoB* genotype. Instead, it shows negative epistatic effect when combined with the WT and S512F *rpoB* genotype but positive epistasis when combined with I572S.
- There is significant epistasis detected among pairs that we have not seen before: Δ*nuoC* with S512F, Δ*ybfG* with I572S. It is unclear if these are real or spurious effects.

The time to reach carrying capacity is a more complex summary statistic of the growth phenotype. Firstly, small values can be the result of either fast growth to an average carrying capacity or that the final carrying capacity was small and therefore was reached quickly even with slow growth. If I really wanted to I could try standardizing the time to reach 85%ile growth by the value of the growth at 85%ile itself but I feel like this standardized metric is basically the same thing as the AUC and I don't want to further complicate things. Honestly, I was hoping to use the time to reach carrying capacity as an inelegant stand-in for the maximum growth rate but it seems that's probably not a great idea. As expected, the ANOVA results for the time to reach carrying capacity (not shown) were very similar to the carrying capacity itself except that the coefficients are swapped in their signs.

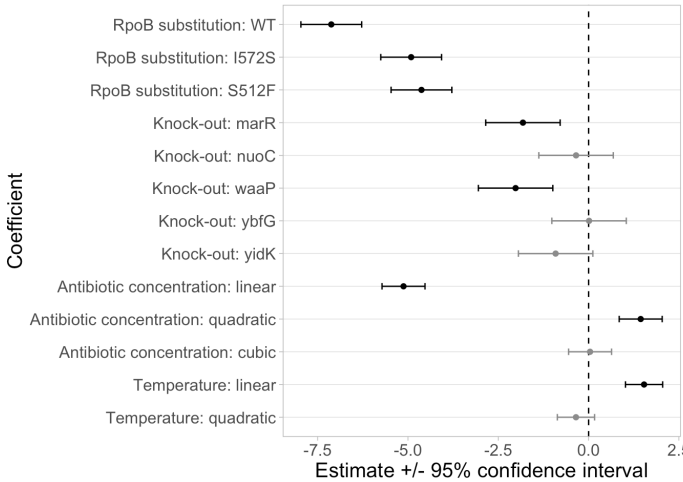


Figure 5: **AUC:** Estimated mean values for the coefficients of the regression model with additive effects (error bars show 95% confidence intervals). The intercept is for *rpoB* with the H526Y substitution and without any knock-out mutation as grown in the 37°C environment without any rifampicin antibiotic. Statistically significant coefficients ($\alpha = 0.05$) are shown in black and have confidence intervals that do not overlap with zero. Coefficients whose estimates are not statistically significant are shown in gray and have confidence intervals that overlap with zero (dashed vertical line).

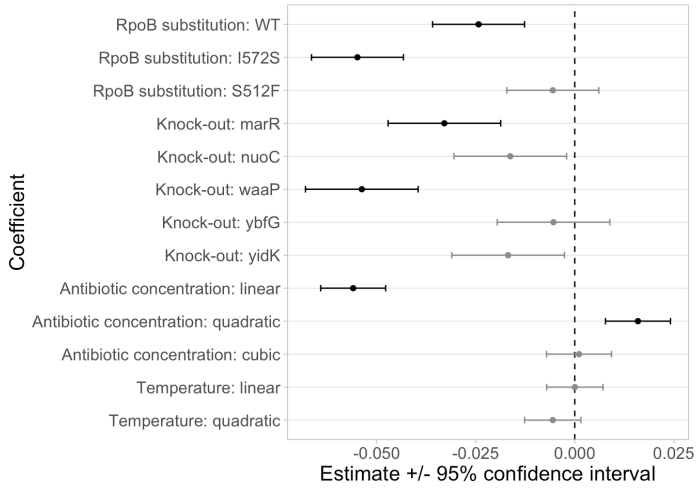


Figure 6: **Carrying capacity with additive effects only:** The style of the figure is the same as for figure 5 above.

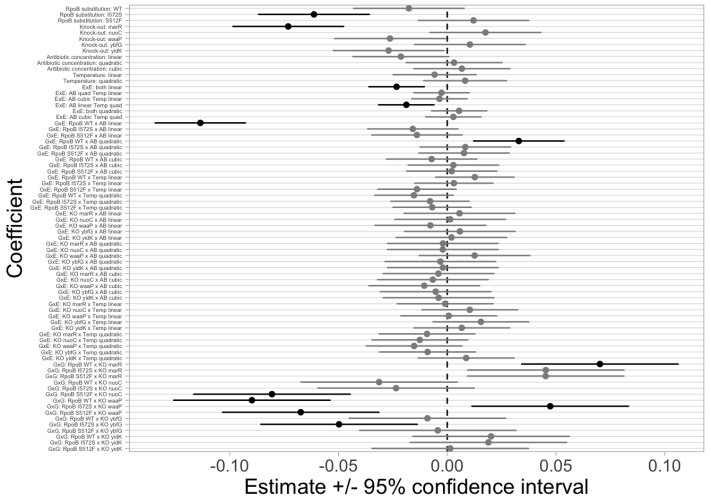


Figure 7: **Carrying capacity with all first-order interaction terms:** The style of the figure is the same as for figure 5 above.

4 Very Preliminary Discussion

In general the AUC results are consistent with our results from the competition data. The AUC finding that Δ*marR* seems to grow as poorly as Δ*waaP* across all environments is surprising. Our competition results found that Δ*marR* grew similarly to the wild-type like genotypes and the other KOs. It could either be that Δ*marR* grows well in the first 20 hours (this is the time when the competitive index was measure) then grows poorly after that or it could be that flow cytometry and the OD measurements have different results. The carrying capacity data is in agreement with the AUC data, however. I think this discrepancy between the trait information and the fitness data could indicate that the Δ*marR* genotype may exhibit a smaller cell size even at stationary phase but the same total cell numbers.

I expected the results for the carrying capacity to be similar with AUC because the time when the carrying capacity was reached is used as the cut-off for the AUC. So I thought that the two should be somehow correlated but, while the results are somewhat similar, there are many differences.

Future work on this project will use more reproducible methods (like publicly available code or programs) to extract the growth traits from the growth curve data. I would like to use a phenomenological method, AMiGA (Midani, Collins, and Britton, 2021), because it is very flexible and would be able to quantify the biphasic growth. As discussed in Chapter 4, it would probably be better to use a kinetic model. Unfortunately I am not yet very familiar with biphasic growth models of this type.

The reason why we initially gathered the “acclimatized growth curves” is because we assumed that the growth rate would be constant between the first and second inocula. Using this assumption and the known dilution size, we could then calculate the fraction of live cells that were present in the stationary phase culture. The death rate is an important life history trait but our OD data does not provide us with any way to readily quantify this trait. However, we now know that it is not reasonable to assume that the growth rate in the acclimatized growth curves (i.e., second inoculum) would be the same as that of the unacclimatized growth curves (i.e., first inoculum; figure 4). Instead I think it would be interesting to try to use this acclimatized growth curve data in order to quantify phenotypic plasticity. Acclimatization is a type of phenotypic plasticity (Bennett and Lenski, 1997). Phenotypic plasticity is important in bacteria; for example it can confer resistance to stressful environments (Goodson and Rowbury, 1989). Specifically, *rpoB* mutations are known to impact phenotypic plasticity (Hug and Gaut, 2015).

More importantly, future work will try to tie together the genotype-to-fitness mapping produced in Chapter 3 in order to better understand how fitness landscapes change across environments.

References

Andersson, D. I and D. Hughes (2014). "Microbiological effects of sublethal levels of antibiotics". In: *Nat. Rev. Microbiol.* 12.7, pp. 465–478. URL: https://dx.doi.org/10.1038/nrmicro3270.

Bank, C. (2022). "Epistasis and Adaptation on Fitness Landscapes". In: *Annu. Rev. Ecol. Evol. Syst.* 53, pp. 457–479. URL: https://doi.org/10.1146/annurev-ecolsys-102320-112153.

Bell, B. G. et al. (2014). "A systematic review and meta-analysis of the effects of antibiotic consumption on antibiotic resistance". In: *BMC Infect. Dis.* 14.1, pp. 1–25. URL: https://doi.org/10.1186/1471-2334-14-13.

Bennett, A. F. and R. E. Lenski (1996). "Evolutionary adaptation to temperature. V. Adaptive mechanisms and correlated responses in experimental lines of *Escherichia coli*". In: *Evolution* 50.2, pp. 493–503. URL: https://doi.org/10.1111/j.1558-5646.1996.tb03862.x.

— (1997). "Evolutionary adaptation to temperature. VI. Phenotypic acclimation and its evolution in *Escherichia coli*". In: *Evolution* 51.1, pp. 36–44. URL: https://doi.org/10.1111/j.1558-5646.1997.tb02386.x.

Bidlas, E. and R. J. W. Lambert (2008). "Quantification of hurdles: predicting the combination of effects—interaction vs. non-interaction". In: *Int. J. Food Microbiol.* 128.1, pp. 78–88. URL: https://doi.org/10.1016/j.ijfoodmicro.2008.06.030.

Blount, Z. D. (2015). "The unexhausted potential of *E. coli*". In: *eLife* 4, e05826. URL: https://doi.org/10.7554/eLife.05826.

Daughtry, B. J., K. R. Davey, and K. D. King (1997). "Temperature dependence of growth kinetics of food bacteria". In: *Food Microbiol.* 14.1, pp. 21–30. URL: https://doi.org/10.1006/fmic.1996.0064.

Durso, L. M. and K. L. Cook (2014). "Impacts of antibiotic use in agriculture: what are the benefits and risks?" In: *Curr. Opin. Microbiol.* 19, pp. 37–44. URL: http://dx.doi.org/10.1016/j.mib.2014.05.019.

Goodson, M. and R. J. Rowbury (1989). "Habituation to normally lethal acidity by prior growth of *Escherichia coli* at a sub-lethal acid pH value". In: *Lett. Appl. Microbiol.* 8.2, pp. 77–79. URL: https://doi.org/10.1111/j.1472-765X.1989.tb00227.x.

Harmand, N. et al. (2017). "Fisher's geometrical model and the mutational patterns of antibiotic resistance across dose gradients". In: *Evolution* 71.1, pp. 23–37. URL: https://doi.org/10.1111/evo.13111.

Hug, S. M. and B. S. Gaut (2015). "The phenotypic signature of adaptation to thermal stress in *Escherichia coli*". In: *BMC Evol. Biol.* 15, p. 177. URL: https://doi.org/10.1186/s12862-015-0457-3.

Midani, F. S., J. Collins, and R. A. Britton (2021). "AMiGA: software for automated analysis of microbial growth assays". In: *mSystems* 6.4, e00508–21. URL: https://doi.org/10.1128/mSystems.00508-21.

Mira, P. M. et al. (2015). "Adaptive landscapes of resistance genes change as antibiotic concentrations change". In: *Mol. Biol. Evol.* 32.10, pp. 2707–2715. URL: https://doi.org/10.1093/molbev/msv146.

Ratkowsky, D. A. et al. (1982). "Relationship between temperature and growth rate of bacterial cultures". In: *J. Bacteriol.* 149.1, pp. 1–5. URL: https://doi.org/10.1128/jb.149.1.1-5.1982.

Ratkowsky, D. A. et al. (1983). "Model for bacterial culture growth rate throughout the entire biokinetic temperature range". In: *J. Bacteriol.* 154.3, pp. 1222–1226. URL: https://doi.org/10.1128/jb.154.3.1222-1226.1983.

Rodríguez-Verdugo, A., O. Tenaillon, and B. S. Gaut (2016). "First-step mutations during adaptation restore the expression of hundreds of genes". In: *Mol. Biol. Evol.* 33.1, pp. 25–39. URL: https://doi.org/10.1093/molbev/msv228.

Rodríguez-Verdugo, A. et al. (2014). "Different tradeoffs result from alternate genetic adaptations to a common environment". In: *PNAS* 111.33, pp. 12121–12126. URL: https://doi.org/10.1073/pnas.1406886111.

Rosso, L. et al. (1995). "Convenient model to describe the combined effects of temperature and pH on microbial growth". In: *Appl. Env. Microbiol.* 61.2, pp. 610–616. URL: https://doi.org/10.1128/aem.61.2.610-616.1995.

Salter, M. A. et al. (2000). "Modelling the combined temperature and salt (NaCl) limits for growth of a pathogenic *Escherichia coli* strain using nonlinear logistic regression". In: *Int. J. Food Microbiol.* 61.2-3, pp. 159–167. URL: https://doi.org/10.1016/S0168-1605(00)00352-4.

Savageau, M. A. (1983). "*Escherichia coli* habitats, cell types, and molecular mechanisms of gene control". In: *Am. Nat.* 122.6, pp. 732–744. URL: https://doi.org/10.1086/284168.

Swinnen, I. A. M. et al. (2005). "Quantifying microbial lag phenomena due to a sudden rise in temperature: a systematic macroscopic study". In: *Int. J. Food Microbiol.* 100.1-3, pp. 85–96. URL: https://doi.org/10.1016/j.ijfoodmicro.2004.10.022.

Zwietering, M. H. et al. (1990). "Modeling of the bacterial growth curve". In: *Appl. Environ. Microbiol.* 56.6, pp. 1875–1881. URL: https://doi.org/10.1128/aem.56.6.1875-1881.1990.

Zwietering, M. H. et al. (1991). "Modeling of bacterial growth as a function of temperature". In: *Appl. Environ. Microbiol.* 57.4, pp. 1094–1101. URL: https://doi.org/10.1128/aem.57.4.1094-1101.1991.

Zwietering, M. H. et al. (1994). "Evaluation of data transformations and validation of a model for the effect of temperature on bacterial growth". In: *Appl. Environ. Microbiol.* 60.1, pp. 195–203. URL: https://doi.org/10.1128/aem.60.1.195-203.1994.

Chapter 6

Discussion

Dr. Ian Malcolm: *Gee, the lack of humility before nature that's being displayed here, um, staggers me.*

[. . .]

Dr. Ian Malcolm: *You stood on the shoulders of geniuses, uh, to accomplish something as fast as you could. And before you even knew what you had, you, you patented it, and packaged it, and slapped it on a plastic lunch box, and now you're selling it! You wanna sell it! Well–*

John Hammond: *I, I don't think you're giving us our due credit. Our scientists have done things which nobody's* ever *done before.*

Dr. Ian Malcolm: *Yeah. Yeah, but your scientists were so preoccupied with whether or not they could to even stop and think about whether they should!*

– Jurrassic Park

Contents of Discussion Chapter

In the introduction I posed six questions that tie together the work presented in my PhD thesis. I will answer the four smaller, scientific questions in turn and then answer the main scientific question of my thesis. Finally, at the very end, I turn to the philosophical and ethical considerations of predicting evolution.

1 Can we build fitness landscape models that capture the complex fitness effects observed in natural populations?

In Chapter 2, we built a fitness landscape that captures the hybridization-facilitating and hybridization-averse fitness effects exhibited in a case study natural population. A ridged and holey fitness landscape captures the main pattern observed in the natural ant populations. Moreover, we could approximate the model results for lethal incompatibilities to many loci to describe what happens in the biologically more realistic scenario where multiple loci underlie incompatibilities.

Our model failed to capture some of the aspects of the data. Figures S7c and S8c show the change in frequency of the different genotypes as a result of selection. The real data show much stronger viability selection (i.e., the absolute change in genotype frequencies is much larger for the real data) as compared to our statistical model predictions. Fitting the model with mate choice to the data would probably not solve the problem. Maybe having more pairs of incompatible loci in the model might solve the problem, that is just a guess. But, even only just with mate choice and certainly with more incompatibilities, we were reaching the end of how complex a model (number of parameters) we can fit to the little available data. More data on mating preferences, for example from the lab, could help us focus on a narrower range of mate choice parameters. More importantly, better assignment of introgressed individuals to their genotype categories coupled with their frequencies before and after viability selection would give us more data for fitting the model. Multiple barrier loci (putative BDMIs) have since been identified in males of this hybrid population (Kulmuni et al., 2020). A similar model with more pairs of incompatible loci could be fit to this future data (but see below).

Overall, our model was pretty good and captured the general aspects of the data (figures S7a-b and S8a-b), like the unusual skew in the population where *F. aquilonia*-like parental genotypes out-number the *F. polyctena*-like parental genotypes. This skew in the data is probably what led our model to predict that one of the compromise scenarios (either asymmetric co-existence or single-locus polymorphism) are more likely future outcomes for the population (figure S9). We rejected the long-term outcome of either the complete loss of polymorphism scenario (exclusion) or the complete maintenance of polymorphism scenario (symmetric co-existence). But that was pretty obvious to happen anyway given the skew in the data and especially because of our strong modelling assumption that the population is at equilibrium.

Recent results on the demographic history of the parental species show that these species diverged with asymmetrical gene flow from *F. aquilonia* into *F. polyctena* until recently ($\sim 2 \cdot 10^5$ generations ago, or 2-19 KYA), when gene flow stopped (Portinha et al., 2022). The age of the hybrid population is still unknown, but it is definitely younger than the species divergence time. The relative age of the hybrid zone is important as ancestral polymorphism leading to incomplete lineage sorting and even differential sorting of incompatibility loci could be more of an issue with the hybrid population. In addition to the historical gene flow from *F. aquilonia* into *F. polyctena*, the *F. aquilonia* population is larger than *F. polyctena*, suggesting that the skewed population ratio of the hybrids could simply be the result of more gene flow from the larger parental population. The hybrid population is probably not at equilibrium both because of its young age, possible ongoing differential gene flow from the parental populations, and, most importantly, because of changes in selection due to changes in the environment. Sampling of allele frequencies for the male barrier loci in 2004 and again 10 years later found that selection against hybrid males in 2004 had shifted to selection in favour of hybrid males by 2014 (Kulmuni et al., 2020). Our model was built to describe the earlier result. This change in selection is thought to be due to the different temperature tolerances of the parental species (Martin-Roy et al., 2021). Looking at the survivorship of different genotypes under different environmental temperatures over time, the researchers predict that climate change will lead to northernly range shifts for both parental species but that the hybrid population may be better equipped than either parental species to deal with extreme climate events.

The main take-home message for me from Chapter 2 is that yes, we can build fitness landscape models that capture the complex fitness effects observed in natural populations. However, those fitness landscape models may not be useful for making future evolutionary predictions about the population (Shaw, 2019). Our model with one pair of incompatible loci [I hesitate to call it simple because its structure of having both holes and a ridge is relatively complex; although its few loci greatly simplify the underlying biology] captured the general trend in the 2004 data and allowed us to better examine how different evolutionary forces play out within a system. The limitations of the fitness landscape model highlighted above (few loci, equilibrium dynamic assumptions, and changing environments) can be somewhat overcome by including the recent findings. We could fit newer, more complex fitness landscapes using that information. However, this underscores that the fitness landscape models estimated to have shaped the past history of genomes are not the same fitness landscapes models that shape that same population's future evolution. Instead, historical fitness landscapes reflect

the past demographic history and environments that shaped those genomes over sufficiently many generations to arrive at current allele frequencies, linkage disequilibrium across the genome, site-frequency spectrum, etc. In order to better understand the evolutionary past, we need to combine the fitness landscape of interest with the demographic history. On the other hand, in order to understand the evolutionary future of a population, the demographic history and present are only somewhat important. They are still relevant because the demographic history determines the currently segregating variation in the population, while the demographic present determines the efficacy of selection. But, more importantly, we need to better characterize the evolutionary forces that are currently acting on the population, as done by Kulmuni et al., 2020; Martin-Roy et al., 2021 even without using fitness landscape models.

Another conclusion that I take from Chapter 2 is that natural systems are very complex, have many different evolutionary processes happening at the same time, and that the selective environment is changing all the time. This makes it difficult to make predictions about the future. It is also difficult to gather the necessary data, like genotype-specific fitness data. We turn below to the study of a more tractable evolving system: using laboratory experimental evolution methods to study ABR in *E. coli*.

2 What do empirical fitness landscapes of untested genetic combinations look like?

There are many different models for theoretical fitness landscapes (Gavrilets, 2019; Fragata et al., 2019) but it seems that empirical fitness landscapes have some different features as compared to the theory (De Visser and Krug, 2014; Bank, 2022). Fitness landscapes look different depending on the types of mutations investigated. We chose to look at intergenic epistasis between SNP and KO mutations for novel combinations of mutations. We already knew from previous studies of looking at mutations along an adaptive walk that it's less likely to observe epistasis between mutations fixed during an adaptive walk, even between KO's (Gorter et al., 2018). So we thought it would be interesting to look at between-gene, between-category novel (i.e., evolutionarily untested) mutation combinations. I expected there would be more epistasis because random mutations are probably bad in combination and because *rpoB* is highly pleiotropic. As detailed in the paper, we found about as much epistasis as found in other studies that looked at random mutations.

When considering between-gene fitness landscapes, it would be interesting to compare the fitness landscapes of genes with more or less pleiotropy. The big-picture

question would be along the lines of, Is it possible for a gene to be inherently more or less prone to epistasis? Genes with more pleiotropic interactions probably need to evolve more robustness and therefore fewer G×G interactions as compared to less pleiotropic genes. My co-supervisor, Isabel Gordo, hypothesizes that environmental specialization can be used as a proxy in *E. coli* for gene pleiotropy. This is because gene mutations (especially KOs) that are impacted by only one environment probably correspond with genes that have specialized functions and are less well connected in metabolic gene networks. Conversely, when genes that are well connected in metabolic gene networks are knocked out, these KOs should impact growth in many different environments. This hypothesis rests on the assumption that gene expression in *E. coli* is primarily regulated through metabolism (e.g., Strom and Kaasen, 1993). I really like this hypothesis of Gordo and mention it now as an example of a larger methodology that she employs called getting "a feeling for the environment," which I will return to later.

We tried to address the question of how gene pleiotropy impacts epistasis in the work presented in Chapter 3 but we could not. The initial approach we tried for identifying KO candidates was fruitless in our case but it may prove worthwhile if done a little differently. In our case, using the database from Nichols et al., 2011 (and then Kritikos et al., 2017) to try to identify KO's specialized to each environment did not work because we had selected the rifampicin antibiotic and high temperature environments *a priori* based on the expected G×E between *rpoB* mutations and high temperature (Trindade, Sousa, and Gordo, 2012; Rodríguez-Verdugo, Gaut, and Tenaillon, 2013; more in next section). Unfortunately, we were unable to find KO mutations that varied in their specialization to the different environments. The main problem that we had was that the growth predictions from Nichols et al., 2011 were not consistent with our results in liquid media, not even for LB. Although the microbiologist 'folk wisdom' warns that solid and liquid media yield different effects, it was frustrating that we could not capitalize on the available, published database. Moreover, it is puzzling, at best (and maddening at worst), that the solid media predictions for could not provide absolutely *any* information about performance in liquid culture. I could understand it if the KOs on the Keio collection background gave inconsistent results with the KOs on the Gordo lab WT backgrounds, since the Keio collection WT-background is biofilm-forming (Tenorio et al., 2003) while the WT used in the Gordo lab is not. But the same genotypes grown in an ostensibly very similar environment gave different results than available in the literature.

If future researchers decide to investigate the correlation between gene pleiotropy

and epistasis using performance across many environments as a proxy for pleiotropy, here is my advice. If you want to use data from a previously published high-throughput screens done in a different lab than the one you are working in (i.e., as we tried to do), I suggest selecting both the KO's *and the environments* based on genes that tend to either be sensitive to many environments (putative highly pleiotropic genes) and invariable across most environments but still responsive to one or a few (putative low pleiotropy genes). Putative highly pleiotropic genes should include both genes whose KO tends to lead to lack of growth across many environments as well as increased growth, or both decreased and increased growth. This is because high pleiotropy genes may exhibit different functional roles but the variable of interest here is pleiotropy itself, not function. Based on our experience using the databases of Nichols et al., 2011; Kritikos et al., 2017, I advise that it would be even better if the researchers perform the high-throughput screen in their own lab. This is both because results obtained in another lab environment may not be robust (organisms are very sensitive to their environments) but also because the relevant data may not be available in that database (e.g., Nichols et al., 2011; Kritikos et al., 2017 only report z-scores but not the absolute growths standardized to the Keio collection WT).

Another way to test more directly how empirical fitness landscapes change between high pleiotropy and lower pleiotropy genes would be to start with engineered genotypes that have only the minimal functional genome (e.g., Martínez-García and Lorenzo, 2016). As all the metabolic pathways and interacting partners are known for these synthetic genomes, it should be easier to identify gene candidates which are more or less pleiotropic. In this case, as all genes are required in the genome, it would only be possible to look at betweeen-gene interactions among mutated amino-acids, not KOs. Moreover, experimental evolution of these synthetic genomes under different demographic conditions and environmental stressors could give us some idea of how fitness landscapes change as a result of evolution (e.g., evolution of robustness).

I would guess that empirical fitness landscapes display more 'mutational robustness' than the fitness landscape models usually considered in theory. This is a vague term that is difficult to explain clearly but has been hypothesized to evolve in response to environmental stochasticity and be related to epistasis (reviewed in Siegal and Leu, 2014). This seems an exciting possibility that I wish I could have spent more time on in Chapter 3 but I can't think of any way to directly test this hypothesis with the empirical data we gathered. Real fitness landscapes are probably shaped both by the functional properties of the underlying biology (e.g., proteins or gene regulatory networks) and by the evolutionary history of the population (e.g., demography, niche

specialist versus generalist, pangenome size). I think it could be interesting to compare species with different pangenomes. *E. coli* has a large pangenome so maybe its core genes have evolved to be more robust. Comparing between different species may not be the most elegant approach as different species vary in other ways than just the size of their pangenomes.

3 How do empirical fitness landscapes of untested genetic combinations change across environments?

Although previous work has extended some fitness landscape models across environments when the optimum is assumed to be known for each environment (Martin and Lenormand, 2015), much work is still required to understand how fitness optima of different environments are related (Harmand et al., 2017). In reviewing the literature, we noticed that there were two ways that epistasis between environments had been approached. The first approach was to expose the same fitness landscape to unrelated environments and see how the landscapes changed (e.g., Bank et al., 2016). This approach is interesting because it investigated fitness landscape properties directly as the focus. The main limitation we saw, however, was that there was no *a priori* hypothesis for how the investigated environments could impact the landscape. For example, even a naive ranking of environments based on a concentration gradient or dose response curve could allow for more meaningful conclusions than just that the same fitness landscape looks different in a different environment. The second approach was to investigate the epistasis of ABR mutations under increasing concentrations of antibiotic(s). In this case, investigators were motivated to understand G×E for some limited mutations but across a variety of well described environmental gradients, sometimes even combinations of gradients to look at environmental trade-offs and synergisms (e.g., Brochado et al., 2018). Where the first approach was poor in its description of the environment, the second approach was rich in the environments chosen and the potential mechanistic underpinnings of the observed G×E (J, 2014). Conversely, where the second approach lacked a deeper investigation of epistasis for fitness (often not estimating fitness at all and only focusing on the minimum inhibitory concentration; reviewed in Bank, 2022; see next section below), the first approach was was rich in its descriptions of the fitness landscapes. Above all, we felt that a combination of these two approaches would be most amenable to applying the existing theory on fitness landscapes between environments. For example, we wanted to extend on the work of Trindade, Sousa, and Gordo, 2012 to estimate how the location of the optimum under Fisher's geometric model shifts

across environmental gradients.

Previous work in the Gordo lab had focused on streptomycin and rifampicin ABR, including trade-offs between these ABR mechanisms (Durão et al., 2015; Sousa et al., 2017). Isabel Gordo encouraged me to move away from clinical environments and instead look at a combination of environments that would be relevant for understanding the evolution and maintenance of ABR mutations out in the natural world. As such, I selected one environmental axis to be antibiotic and the other axis to be physiologically relevant temperatures, as had previously been investigated by Trindade, Sousa, and Gordo, 2012. I also chose this combination of environments because of the previous results of Tenaillon et al., 2012. In this experiment, evolution to high temperature was replicated over 100 times and epistasis was found to play an important role in the mechanism of *de novo* adaptation. Depending on mutational stochasticisty, evolution could either proceed through mutations to *rpoB*, including the evolution of ABR to rifampicin (Rodríguez-Verdugo, Gaut, and Tenaillon, 2013), or by mutations to the *rho* gene. From this work and others (Jin and Gross, 1989), the feeling we were getting for the environment was that there should be a rich combination of G×E interactions along both environmental axes for the *rpoB* mutations and, as discussed in the section above, for the KO mutations. In the end, I was disappointed that the M9 environment we selected did not display the expected G×E interactions along the temperature axis. We selected the M9 minimal medium because it is a clear and described medium. If I had more time I would have re-done the competition experiments in a rich medium, like LB.

Since I began work on my PhD thesis, several studies have been published that have extended our knowledge of how empirical fitness landscapes change across environments (e.g., Hall et al., 2019; Flynn et al., 2020). The most important works combine mechanistic models with empirical data (e.g., Guerrero et al., 2019; Das et al., 2020; Pinheiro et al., 2021). I wish we could have done that. I hope that the growth curve data from the unpublished chapter can be analyzed in such a way in order to integrate genotype-to-phenotype-to-fitness mappings between environments. I guess it's okay even if we did not get a chance to fit models to my data because we need to quantify what the world looks like before we can come up with models. I wish I could have made a contribution to this area but I could not figure out how (yet???).

4 How do we measure fitness?

Fitness is a complex phenotype that describes the "dynamics of frequency change of mutations" in a population (Chevin, 2011). Any trait that leads to different individuals leaving a different number of progeny (e.g. growth, survival, and reproduction) can be a component of fitness (Orr, 2009). Different fields of microbiology consider different proxies for fitness, from competitive index (Chevin, 2011; Sousa et al., 2017) to maximum growth rate in monoculture (Hall et al., 2014; Concepción-Acevedo et al., 2015) to the minimum inhibitory concentration of an antibiotic (Knöppel, Näsvall, and Andersson, 2017). In order to relate the fitness of an individual across different environments, it is important to define and quantify the different growth traits that contribute to fitness.

For larger organisms, it is well accepted that fitness proxies (e.g., number of viable gametes, number of offspring, survivorship) are used as a substitute for fitness (e.g., Kulmuni and Pamilo, 2014). For microbial populations, it is possible to estimate the population size over many generations in a short amount of time. At the end of the day, we are probably never going to be able to measure fitness directly because of inclusive fitness, G×E, and environmental fluctuations. The paper presented in Chapter 4 was a side-project and it took us a long time to get this paper into review (we had maybe 7 desk rejections) before we just sent it off to *PLOS One* (where it is in review at time of writing). I guess our paper is not a novel contribution to the literature on estimating fitness but that's okay too because scientific contributions can be important without being completely novel. Our meta-analysis shows that researchers need to be more careful and the novel contribution is that we show that the common intuition of using relative growth rates can lead to a change in sign of the estimated fitness value / selection coefficient. The point is that accurate fitness measures are important for parameterizing models that will be used for predicting evolution. I think it's better to use an approach like that used in Chapter 3 (as opposed to those discussed in Chapter 4-5) to measure fitness.

The work from Chapter 4 emerged as a side-project because I wanted to know what the best way would be to analyze the data that I presented in Chapter 5. Evolution is most predictable for traits that are closely tied to fitness: life history traits. We were hoping that we could use the growth curve data to fit mechanistic batch culture growth curve models in order to extract the life history traits, for example growth rate (exponential phase) versus death rate (stationary phase). However, density-dependent effects impact these life history traits (Hilau et al., 2022) and so we cannot assume a constant death rate, for example. Using the growth curves allowed us to non-destructively gather

a lot of temporal data. On the other hand, we could have used flow cytometry with live/dead staining at several time points to estimate the viability, proliferation, and death rate of the populations more directly (by setting up a larger batch, distributing it over several smaller replicates, incubating, and destructively sampling the replicates at different time points during the growth curve). The benefit would be more accurate estimates of life history traits but it would be more labour-intensive and therefore we would need to have a shorter total duration, fewer time points, and potentially fewer total replicates. (Micro-)chemostats are great because direct measurement of the net growth rate, meeting common model assumptions of density independence because the density is kept constant. But we do not have chemostats or microfluidic devices available in the Gordo lab. At any rate, neither chemostats nor microfluidic devices seem very biologically realistic environments, as even the mammalian gut (which can be likened to a chemostat as it receives regular input of nutrients and output of waste) experiences feast and famine circadian rhythms that impact the microbial ecology (Deaver, Eum, and Toborek, 2018). This paragraph is a bit directionless but I think the main conclusion I draw is that it is important to use the organism to get a feeling for the environment. This means that in order to predict evolution in complex environments, we must measure fitness directly in the complex environments of interest. This is technically challenging but it is important for creating useful models and future evolutionary predictions.

5 Is it possible to predict evolution on fitness landscapes in different environments?

Although we are far away from reaching this goal, my thesis work makes me confident that it will one day be possible to predict the evolution of some biological systems, despite the challenge of G×G, G×E, G×G×E, and higher order interactions. Perhaps one reason why I am so optimistic about the possibility of predicting evolution is because my work focuses on segregating variation (Shaw, 2019). Chapter 2 asks about maintenance or loss of existing variation, Chapters 3 & 5 map the shape of fitness landscapes for a finite, constructed set of mutations, and Chapter 4 asks how we can best measure the life history traits of existing genetic variants. Evolution from segregating variation may be more predictable than waiting around for evolution from *de novo* mutations. Nevertheless, the law of large numbers means that evolution can sometimes be more predictable at longer time-scales (and large population sizes, of course) such that enough time passes for many possible mutations to segregate in the population (Good

et al., 2017; Burford Reiskind et al., 2021). Instead of asking whether or not prediction is possible, I think it is more important to discuss which biological systems would be most amenable to predictions.

In the Introduction Chapter I made a distinction between predicting evolution with the goal of improved scientific explanation versus the goal of developing evolutionary prediction technologies. I am optimistic that both are possible but as the evolutionary prediction technologies seem further beyond our reach, I will focus my discussion on those. It seems possible to define some biological systems that would be more amenable to applying evolutionary prediction and control technologies.

Prediction and control of evolution seems most feasible for controlled microbial systems, like bioreactors. In particular, systems that use constant, described cultures for inoculation and have tightly controlled environments would be preferred. Sexual species are preferred, whereas clonal interference may pose a challenge for prediction in asexual species because evolution will tend to depend more on the stochastic order in which mutations appear and compete with one another. The issue of mutational stochasticity can be ameliorated by sequencing in real-time and then updating the model predictions by conditioning on the observed mutations' allele frequencies. Unfortunately, most commercial bioreactor systems (e.g., alcohol fermentation, cheese-making, water filtration) are not yet as sufficiently monitored or controlled as would be necessary for implementing technologies of evolutionary control (e.g., fluctuating temperature and nutrient concentrations, immigration from environmental strains). Moreover, some systems are often already quite optimized and evolutionary technologies may not be able to provide significant gains, especially when assuming additive interactions (Durão et al., 2021 but see Adrian-Kalchhauser and Bank, 2022). Control of the environment is a key advantage of these systems. Domesticated populations of larger (i.e., non-microbial) organisms may be a better choice for evolutionary predictions than wild populations. Similar to bioreactors, there is already more environmental monitoring, control, and data available for domesticated species.

Microbial populations are more amenable to prediction than larger, longer-lived organisms. This is both because of their shorter generation times and larger population sizes, as has been abundantly remarked in the literature, but I think also because of their biology. As already alluded to, organisms with large per site mutational input to the population (i.e., large per site mutation rates and large population sizes) have very low waiting times for new mutations, even when asexual. Organisms with smaller genome sizes are more amenable to prediction because the potential number of mutational targets is smaller. Therefore the task of laboriously mapping genotype-to-fitness

landscapes through deep mutational scanning will yield better returns because a larger fraction of the potential mutational space of the organism can be empirically described. Similarly, it will be easier to empirically map the phenotype-to-fitness, and perhaps the complete genotype-to-phenotype-to-fitness, space for organisms that have simpler life histories (e.g., occupy the same niche throughout their lifecycle, gene expression and function is more uniform across the lifespan, adapted to living in a small set of environments as opposed to generalists). This is not to say that all microbes share these biological features, but microbes tend to share at least some of these features. The microbes that do share all of these features may be among the easiest to predict and control. For example, these traits are present in the few viral populations whose evolution we have already been able to predict (Agor and Özaltın, 2018; Feder, Pennings, and Petrov, 2021). The main challenge of microbial populations probably remains their interactions with their environments, which often include biotic interactions with host organisms.

For wild populations (i.e., those that are not under human management), a large amount of data would need to be gathered about the populations and their environments in order to make predictions. This would have to involve narrowing down the relevant environmental factors for monitoring and control. Indeed, this appears to be the largest challenge for predicting the evolution of microbial parasite populations. Wild populations of larger (i.e., non-microbial) organisms seem the least amenable to prediction because of their longer generation times, relatively smaller population sizes, and uncontrollable environmental conditions. This category of organism includes both charismatic species, whose evolution we may want to predict and manage as part of conservation efforts, as well as our own species. Even when predictions can be made, as we did for the Finnish hybrid ant population in Chapter 2, the (uncontrollable) environments of these populations may change within a few generations, as discussed above.

I appreciate that everyone is very excited about predicting evolution (it is an exciting prospect!) and I understand that predicting evolution can be a lucrative proposal on a funding grant. In particular, as alluded to in the Introduction Chapter, predicting evolution promises to turn the relatively esoteric and unremunerative field of evolution science into a gainful, maybe even profitable, application. This is exactly the type of innovation that motivates public funding for basic research. For this funding reason alone, I can understand why there is so much enthusiasm in the field about the possibility of predicting evolution. It is *not* my goal to dispute a concept or line of work that brings more funding to evolution research. However, I think that we need to reflect

about whether these technologies are the best use of our knowledge – and resources, not to mention time! – in terms of addressing the urgent problems that we believe are important to address (e.g., climate change, biodiversity loss, ABR). On the one hand, I am aware that we tend to write funding grants in a cynical fashion, promising to shoot for the moon when we know we can, at best, only deliver the stars. On the other hand, it is exasperating to me that we as a community of intelligent, resourceful, forward-thinking people choose to prioritize solutions that are decades away when we have much less time than that to address issues that impact the very survival of our civilization and our species(!). Moreover, I consider skepticism about the value of scientific predictive technologies to be a rational emotional reaction in the face of current (and past) events: experts had been predicting a global pandemic like COVID-19 for at least a decade before it happened; the science on climate change began to emerge around 50 years ago but we took few efforts to curb our emissions; ABR evolution as a result of high antibiotic concentrations was discovered concomitant with antibiotics themselves; push-back against capitalist exploitation and alienation began alongside the opening of the first factories during the Industrial Revolution; etc. And now it is clear that there are political concerns that are creeping into my scientific judgements.

Before moving on to address the political/philosophical concerns head-on, I want to address the practical challenge of predicting evolution that was described in the Introductory Chapter. How useful would it actually be to predict evolution? This depends on the type of population that we want to predict the evolution of. I am most sympathetic to prediction and control over the evolution of microbial parasites. I hope that this is because these biological systems are more amenable to control, as argued above, but I am aware that the feeling is also because I sympathize least with human pathogenic organisms. Nevertheless, our collective, recent experiences of intense push-back against modifying human behaviour to impact the ecology (and perhaps evolution) of a wild microbe (i.e., SARS-CoV-2) make me feel less optimistic about the prospect of controlling, or even predicting, the evolution of microbial parasites. I am most skeptical of attempts to predict and control evolution as a technological solution to anthropogenic environmental deterioration. Thompson et al., 2021 provides a pointed example of this perspective. The authors specifically mention improving "ecosystem resilience," but it seems odd to me to believe that naturally evolved ecosystems are inherently less resilient than can be practically achieved through human rational control. I think the returns on investment of predicting evolution are far worse than applying general conservation knowledge about the importance of minimizing human encroachment into the species' range and general evolutionary knowledge about the dangers of inbreeding

depression / adaptive benefits of genetic variation (although Shaw, 2019 makes an appealing argument about predicting the evolutionary impact of human mediated gene-flow). Most of all, I believe that political solutions may sometimes be far superior to technological solutions. I refuse to live in a world where the scientific consensus believes that it is more likely to rationally re-engineer the resilience of our entire global ecosystems than to give legal rights to Indigenous peoples over their traditional lands (Garnett et al., 2018; IPBES, 2019).

Sometimes in evolution the ostensibly scientific controversies are not actually about the science. Instead they are political, moral, and philosophical concerns that find themselves hotly debated in our science because of our moment in history. Often these controversies fade and science tends to pick one side, but I think the winning idea often receives significant help from historical and political influences in addition to scientific evidence. Two examples that I glossed over in my Introduction Chapter include the argument about hard versus soft inheritance (which is a philosophical argument about the role of our inherent being versus our environment in shaping who we become) and the argument about the gene's-eye view versus multi-level selection (which I see as an extension of the kin selection versus individual selection argument, both of which are political arguments about individualist versus communitarian ways of organizing our lives). I'm not an historian, nor a philosopher for that matter, so I'm neither skilled nor knowledgeable enough to examine the historical context that precipitated these specific concerns. But I do notice that these controversial questions are not resolved by the science: instead, one side becomes orthodoxy as the chief proponents for the other side leave the field (by choice or by age) or shift their interests.

Although its history is obscure because eugenics is a difficult topic, predicting evolution is another example of a long-standing controversy in evolutionary biology that continues because of political/philosophical concerns. (Following the definition of eugenics given in the Introduction Chapter, I herein refer specifically to eugenics as a historical political movement from the mid-19th to the mid-20th century. As such, I consider the field of predicting evolution to have first emerged around the time of Darwin/Galton, as eugenics, and to continue into the present day.) On the one hand, proponents of predicting evolution focus on its utilitarian applications and the ambition of a project that would unite all fields of biology. Their argument about the predictability of evolution rests on a modernist understanding of the world as mechanistically describable and statistically quantifiable (Lloyd, 1993; Leiss, 1994; Keller, 1996). On the other hand, opponents of predicting evolution focus on the complexity,

interdependence, and serendipity of the world.[1] While the story of predicting evolution told here has focused on population genetics at short-time scales, which is therefore bolstered by results from experimental evolution, the predicting evolution debate is prevalent even in fields like paleontology, which cannot take an experimental or $\Delta t \to 0$ time-scales approach to evolution. Therefore, the question of predicting evolution is not merely a matter of setting our time-scale to be sufficiently small, as arguments from population genetics alone tend to lead us to believe. It is also not just a matter of experimentalists wanting to do experiments on the world, as the deep philosophical tensions that have existed between population genetics and paleontology might lead some to believe (ref?). As argued in the Introductory Chapter, and depicted by the only figure in that chapter, I believe that predicting evolution is a genuine feature that is made possible both by the mechanisms/processes that operate in our world (ontology) and our understanding of them (epistemology). This part is not controversial for me. What I believe makes predicting evolution controversial today, even across disciplines of evolution, is that there is a fundamental difference in world-view that is espoused by supporters in each camp.

In order to fast-forward through the types of drawn-out scientific conflicts that might result from conflating political/philosophical arguments with scientific ones (e.g., soft versus hard inheritance, gene's-eye view versus multi-level selection), I think it is important to devote some time to the philosophical considerations of predicting evolution.

6 Should we predict evolution? Or: Towards a cross-disciplinary discourse on the ethics of evolutionary predictions

As demonstrated implicitly in the Introductory Chapter, where the history and ideas of predicting evolution were shown to have descended from the (British) eugenics movement, predicting evolution poses an inherent ethical challenge. But what exactly is that challenge and what can our eugenic past teach us about the present of predicting evolution? After giving some background (subsection 6.1), in subsection 6.2 I identify the moral dilemma of technologies that predict (and control) evolution and I explain that

[1] I personally prefer to see the world as sacred, full of complex causalities and unexpected possibilities. But my preference for an embodied experience of such a world says nothing about the true nature of the world. Nor does the nature of the world say anything about the felt-sense embodiment that I choose to experience of the world.

eugenics was a special-case of this larger moral dilemma. The resolution that we, as a collective, have decided upon for this dilemma was achieved by assuming a humanist ethics, which prioritizes the agency of all human individuals as inherently valuable. Using ideas from moral relativism, I investigate how the dilemma of predicting evolution can be seen under different ethical systems. I use ideas from naturalistic philosophy (Godfrey-Smith, 2003) to argue in favour of building a culture of science that allows a diversity of perspectives and worldviews to co-exist in order to engage in dialogue about scientific and social issues. Then, I use the emerging scientific consensus that the health of humans is intrinsically tied to the health of other organisms and ecosystems (One Health; Destoumieux-Garzón et al., 2018), to motivate a posthumanist ethics. Finally, and more broadly, predicting evolution can be construed from a science-skeptic, anti-positivist lens (which is prevalent in some parts of the social studies and humanities) as part of the agenda of technoscientific control over the natural world that modern Western science embarked on during the Scientific Revolution and continues to perpetuate today (Pattberg, 2007; Shapin and Schaffer, 2011; Fichman, 2021). In subsection 6.3 I provide a defense of science by arguing that many humans engage in science as a way of feeling deeply connected to the natural world. In this way, I maintain that science is and should continue to be a profoundly human endeavour.

6.1 Good science is aware of its social context

I must apologize to the reader for introducing new ideas and jargon so late in the game (with only 16 pages left, I too thought the heavy lifting was finished!), but please indulge me this informal explanation. The main purpose of this section is to introduce some key ideas from philosophy of science and science studies. Cross-disciplinary dialogue is only possible when we speak a common language. The jargon of philosophy is difficult so I intentionally use colloquial language below.

6.1.1 Background and new jargon from philosophy

Evolutionary biology, like all authoritative/academic natural sciences, is positioned in a materialist ontological stance and a positivist epistemological stance. We believe that the world contains material objects that are "out there," perceptible to our senses (ontological stance) and that we can know true things about the world by empirically observing it and using the scientific process to discover its laws (epistemological stance). The words ontology and epistemology are generally unfamiliar to contemporary scientists because we have a strong consensus of positivist materialism within our working

culture. Don't worry, reader, I'm *mostly* a positivist materialist,[2] too – this is still a PhD thesis in the sciences after all. I'm merely introducing some language to help us all get on the same page. Considering and explicitly stating a researcher's ontological/epistemological stance helps clarify how meaningful knowledge can be created inside of that world-view. Moreover, as all knowledge-making endeavors are social and situated within a larger culture (Kuhn, 1996; Latour and Woolgar, 1987), stating the assumed ontological/epistemological stance makes cross-disciplinary and cross-cultural discussions possible. This is why I didn't have to begin my PhD thesis by stating that I'm a positivist materialist: within our scientific working culture, we take this as a given (Park, Konge, and Artino, 2020). But this is also why I have to now disclose my stance, in order to bridge disciplines (and probably cultures) in my attempt at creating ethical knowledge. The nature of the world (i.e., ontology) together with how we can know things about it (i.e., epistemology) only in part determine how we can create good knowledge. The other part is determined by our methodology and our axiology. Methodology is a familiar term for us, thankfully requiring no change or update of its definition. But axiology may be unfamiliar. Axiology refers to what is good and valuable (Peers, 2018). In most cultures axiology includes ideas about the sacred or spiritual, but normative Western sciences tend to discount the sacred/spiritual as inherently valueless because spirituality is understood to contradict a strict materialist ontology (but see TallBear, 2013). This highlights that one's ontological and epistemological stance will impact their axiological stance. Similarly, methodology will always be informed by the other three. Axiology is particularly important for ethical discussions because ethics is a type of decision-making that optimizes goodness or value.

The scientific process in general, and the evolutionary algorithm in particular, is philosophically interesting because it describes a methodology for gradually updating our understanding of the nature of the world, essentially to pull ourselves up by our bootstraps into a new ontological stance (Dennett, 1995). It's somehow amusing that we hardly use any of this language in science even though we are in the business of ontology – some may argue that we are the only legitimate business of ontology. The ontological stance of the natural sciences has wide-ranging social and political consequences, especially as the ruling cultures (as well as many living beings on earth,

[2]I say that I'm mostly a materialist because I believe the world also contains ideas, which are immaterial. So I am a dualist because I believe the world of ideas is fundamentally different from the world of matter. But I think the world of ideas and the world of matter mutually interact with one another. I say that I'm mostly a positivist because I believe that we can observe phenomena out in the world. But I believe the neuro-scientific evidence that our observations and experiences are impacted by our mental states, including social contexts and beliefs. Overall, my ontological and epistemological stance is very similar to that of most contemporary scientists.

depending on our climate change policies) are dependent on technological innovation driven by scientific progress. This underscores the importance of researchers to be able to articulate and, when relevant, disclose their positions.[3] Moreover (and, thankfully, as this theoretical diversion has dragged on), this brings us to the last piece of jargon: political position.

Politics refers to how humans govern and organize society, as well as beliefs about how they *ought to* do this. Although (party) politics have long been recognized as playing out at the municipal to global scales, starting with Simone de Beauvoir and others, feminist scholars have argued that the personal is also political. This concept summarizes the idea that adverse personal experiences often have larger, structural causes and serves as a rallying cry for people to organize collectively in order to address issues that are not (currently) part of party politics (Hanisch, 2000). The role of the 'personal,' political positions of scientists in the creation of knowledge about the world has been debated for over 50 years, sometimes viciously (Lloyd, 1993; Godfrey-Smith, 2003). Although in the social sciences it is now relatively well established that the positionality/position of the investigator impacts the content of the knowledge they create (Rouse, 1996; Peers, 2018), the natural sciences remain particularly resistant to this idea (Rose, 1983). Philosophers, historians, sociologists, and others continue to debate what is the role of the observer for informing the knowledge claims made about the natural world. The debates were so intense that academic conversation became impossible and splintered the new discipline of science studies from the traditional discipline of (history and) philosophy of science. Scientists in general have tended to stay in their lane [4] and so have offered little input in this debate, although this is

[3]And sometimes it may be relevant for scientists to disclose their positionality. Positioning refers to one's chosen philosophical stance about ontology, epistemology, and axiology (Bacevic, 2021). Positionality refers to one's assigned roles and power in society (Jafar, 2018). I know I have been relatively sloppy with the jargon of positioning, positionality, and political position but I was told not to *waste* too much time on this part of my thesis (even though, I hope I demonstrated in the sections above that it is not possible for me to do the science without doing the philosophy and politics, too). A friend who will soon defend her PhD in a relevant subject area read the draft of this chapter and said that the ideas are solid. Also this paper is nice too: Kittay, 2009.

[4]I think this is for many, diverse, sometimes personal, sometimes political reasons but will only enumerate one for each limb. On the left hand, these postmodernist/poststructuralist/postcolonial/posthumanist/etc literatures are technically difficult to understand even for scientists whose political positions would make them sympathetic to challenging authority. We cannot make useful, substantive contributions to a discourse without first understanding what it's about (but see the work of Gould). On the right hand, the most prominent examples of scientists who engage in this type of debate in recent years have done so to ostensibly rescue science from the perceived dangers of a nefarious postmodernist, Marxist, queer agenda (e.g., Jerry Coyne, Richard Dawkins, Jordan Peterson). Try as I might, it's hard for me to ignore their microaggressions and dispel the impression that they are arguing in bad faith. The point is that these prominent examples are rife with controversy and most working scientists do not welcome that kind of attention. On one foot, most

perhaps changing (e.g., Kamath et al., 2022). Of all the ideas introduced above, the most important is that what we deem good/valuable (i.e., our axiological stance) plays a large part in determining what we should do (i.e., our ethics).

6.1.2 The role of politics in the natural sciences

The moral that is often read by scientists from the eugenics movement is that scientists should avoid politically motivated science: good science is always politically dispassionate and all politically motivated science is, by definition, bad science. This claim arises out of the patriarchical, modernist foundations of Western science (Lloyd, 1993; Keller, 1996; Jaggar, 2015). The 'good science is apolitical' value-judgement leads some people, perhaps naively, to the conclusion that Fisher's eugenically-motivated scientific contributions should be treated with suspicion and scrutinized as bad science. I think we can all agree that Fisher's contribution to evolutionary theory is exceptional science that we would all dream to emulate. I hope that my exposition of Fisher's work in the Introductory Chapter showed that Fisher's science was also politically motivated. So where does that leave the 'good science is apolitical' claim? Most people who hold this view are unable to reconcile it with Fisher's politics. This is why many scientists simply ignore his politics, pretend that Fisher was somehow deluded about his own stated motivations (Grafen, 2003), or may even believe that eugenics is an apolitical position and the real culprit is racism (Bodmer et al., 2021). The 'good science is apolitical' claim can easily be rescued if Fisher's motivations can be construed as purely scientific. However, recent shifts in the popularity and social power of anti-racist ideas have made it harder to ignore that Fisher opens *The Genetical Theory of Natural Selection* with the statement that the work is eugenically motivated. The example of Fisher suggests that things are not so simple – good science can sometimes be politically motivated!

---

academic scientists are too busy trying to secure their spot in a fiercely competitive academic system. We cannot afford to 'waste' our time and emotional energy on difficult cross-disciplinary work. And, finally, on the other foot, certain demographics (i.e., positionalities) have historically tended to be over-represented in academic science and continue to do so at higher positions of authority (e.g., living in the global north, cis-gender, English native or near native speaking, white, conventionally attractive, able bodied at least during early career stages, masculine, financially privileged, monogamous heterosexual, from families of origin that prize academic success and meritocratic individualistic beliefs, etc.). As demography and positionality are well correlated (and maybe position is too?) this influences the axiological assumptions and so what questions are considered scientifically worthwhile. Put colloquially: people who might have more skin in the game (i.e., those from marginalized backgrounds) tend to not be able to secure the permanent or otherwise relevant positions of authority where they can change the axiological assumptions. Things will shift as more demographically diverse individuals attain positions of authority, thanks in large part from active intervention on the part of existing faculty to diversify their ranks. However, regardless of the demographics of those in positions of authority, it does not benefit the individuals in power to foster anti-authoritarian ways of approaching the world.

We can examine the anti-thesis to 'good science is apolitical' in order to use the dialectical method to arrive at a more nuanced opinion (Levins and Lewontin, 1985). Let us consider the anti-thetical claim, 'political motivation produces good science.' Surely there are examples in support of this, like novel biological insights gleaned from focusing research efforts on neglected tropical diseases (ref?), a better understanding of human migrations (TallBear, 2013), a better understanding of female reproductive biology (ref?), etc. On the other hand, we can also find examples that contradict this claim. The tragedy of biology in the USSR under Trofim Denisovich Lysenko, the director of the Institute of Genetics in the Academy of Sciences of the Soviet Union, is an often-cited example of where external, political ideologies limited the capacity of science because Lysenkoism was strongly invested in arriving at particular research outcomes (Soifer, 1994). This example is relevant to evolution because it cut short the promising work of another father of the modern synthesis, Sergei Sergeevich Chetverikov (Levit, 2015; Rispoli, 2021). I would argue that the lesson from Lysenkoism is not about political motivation, but about which part of the scientific process is impacted. Researchers motivated by assuring their own career progression have sometimes been strongly invested in supporting a particular hypothesis and this has led them to misconstrue or outright falsify their data (refs?). Quickly completing the dialectic method to arrive at an more nuanced perspective, I would argue that 'politically motivated science is axiologically neutral,' neither good nor bad. When researchers use their politics to motivate the questions they ask, either as a result of external efforts to provide more funding towards the problems experienced by marginalized groups or as a result of the internal, personal motivations of the scientists themselves, politically motivated science can undoubtedly be good science. When researchers use their politics to determine the research outcomes, either as a result of party politics or personal politics, the value of that science is suspect. This is not an issue of politics, however, this is an issue of scientific methodology. Any science that is done with too much pressure for obtaining a particular conclusion is suspicious, regardless of whether that pressure is as a result of party politics, corporate funding (e.g., Gould, 1991), or personal career advancement.[5]

To summarize: at this point we have gotten away from the normative, patriarchal, modernist view of science ('good science is politically neutral') and we have also rejected some of the hard-line alternative stances ('political motivation produces good science'). We were focusing above on the political motivation of science and concluded that science is agnostic towards politics. I think we can understand something more by examining

[5]The feminist question that arises is why so many scientists find themselves in adverse personal circumstances where the impact factor of their work determines their career stability. Of course, this too is a political issue.

a slightly different claim, that 'good science is aware of its social context.' In other words, science can be motivated by political causes, or its findings may have political implications, but the important thing is that we should try our best to be aware of these politics, instead of claiming an apolitical perspective. This claim also accommodates the possibility that some science truly is apolitical. But, instead of assuming in our scientific culture that the null hypothesis (H_0) is one of apolitical science, we can instead hypothesize that H_0 : all science is impacted by politics and then try to examine the instances where H_A : science is apolitical. In other words, we should try to *a priori* assume that our science is motivated by our (political) perspectives, funded by political entities for their own motives, and will have political implications for society at large. Therefore, we should work to be aware of the social context of the scientific knowledge and technologies that we produce.

This means that researchers in predicting evolution should be aware that the history of the subject is related to eugenics movements. Moreover, they should understand that when predicting evolution is wielded as a technology (i.e., for preparation or control) and the population being forecast is human, the goal of this technology is exactly the same as that of eugenics movements (see below). Science is agnostic to politics – not indifferent. So we need to try to be more aware of what is our political and otherwise embodied perception of the world (e.g., ontology, epistemology, axiology, political opinions), how that might impact our science, and how our science may impact the world.

6.2 Beyond a humanist understanding of predicting evolution

In biology, we usually use humanist ethics to guide our studies. This means that humans are prioritized in our ethical considerations. This is how we justify, for example, sacrificing some minimum number of animal (and other) lives in service of learning more about saving and improving human lives. Humanist ethics has its roots in the European Renaissance and Age of Enlightenment but it features throughout contemporary ideologies (Wikipedia, 2023a). Humanism advocates for the agency and potential of all humans and human social groups. The goal of science and technology under humanist ethics is the betterment of human life.

The 1951 United Nations Educational, Scientific and Cultural Organization (UNESCO) Declaration on Race (along with the earlier United Nations Universal Declaration of Human Rights) is regarded by historians as an important humanist document that marks the end of the eugenics movements (Kevles, 1985; Roll-Hansen, 2010). Although advances in genetics had motivated most scientists to abandon the 'mainline'

eugenics movements some years before the start of World War II, political support for eugenics by the general public continued and only began to decline after the genocides committed by the Nazis were publicized. UNESCO created the Declaration on Race in order to combat racism, as stipulated in its constitution. The UNESCO Declaration on Race was first written in consultation with anthropologists and later revised with input from leading geneticists of the time, including Haldane and Dobzhansky (Fisher's comments are included but he did not draft the statement). Although this was a progressive document in its time, it is dated as compared to our current understanding of human history and biology (e.g., the separate African, Asian, and European "major groups"; comparison of intelligence test results for people of different "levels of civilisation"). The 1951 Declaration on Race has been developed and updated by UNESCO over the years and has been important for dismantling institutional racism and establishing laws that protect racial or cultural minorities. In addition to the genocides committed during World War II, the forced sterilization of criminal or 'feeble-minded' people (that disproportionately targeted Indigenous, Black, Latinx, and other racialized groups) by the government of the USA is another outcome of the eugenics movements. This came into the public spotlight in the 1970's and is still being publicly reckoned with. I mention these two examples specifically because they are the most well known examples and most scientists know about the eugenics movements from general cultural knowledge (i.e., the history of eugenics is not part of the knowledge that we consider as mandatory for a biologist to know). The point of this paragraph is to establish that a particular kind of humanist ethics guide our understanding of eugenics. These outcomes are probably why people tend to associate eugenics with racism, forgetting that eugenics movements also targeted disabled people and other outsiders.

The moral that is read by scientists from the public reckoning about the Nazi genocides committed during WWII, as well as the USA domestic investigations of forced sterilization of racialized people, is that prediction and control of entire human populations is ethically wrong. This is a broad generalization about the beliefs of a large group of people, but my point is that the ethical implications are understood in a universalist humanist framework that emphasizes anti-racism but makes little mention of disability. Individual people's reproductive choices and right to life cannot be taken away from them.

The central ethical issue of eugenics is that it took away people's inherent right to self-determination. The humanist perspective values human agency as a key part of its axiology, and so, eugenics is ethically wrong because it removes human agency. In particular, eugenics puts the people who are doing the predictions in the driver's seat

of evolution, controlling the evolutionary fate of the human population of interest. In other words, it alienated those humans from their own evolutionary potential, agency, and right to life. In the case of the Nazi genocides, it put Nazi doctors and scientists in the driver's seat of the evolutionary fate of Jewish, Romani, disabled people, and other populations.

Now that we have the ethical issue under humanism clearly stated, we can examine the relationship between predicting evolution and eugenics. As I have implied throughout the Introduction and Discussion Chapters, the literature on predicting evolution never mentions predicting human evolution. While this may be partially due to scientific issues (e.g., the long generation time of human populations) and a lack of clear discrimination between the science versus the technology of evolutionary prediction, I think it's also a deliberate choice to avoid conflation of the contemporary field of predicting evolution with historical eugenics (even though, as I have argued, the two fields are inseparable so the conflation would not be spurious). Biology assumes a humanist ethics, therefore normative biologists know that using the tools of predicting evolution on human populations would be ethically wrong (and will not be funded by UN member states). The moral dilemma that I identify is that, 'the technology for evolutionary prediction alienates the population from its evolutionary process.' When the population is humans and authoritarian political processes are used to deploy the technology, the moral dilemma of evolutionary prediction technology is exactly equal to that found in historical eugenics movements.

But what does it mean to alienate a population from its own evolutionary process? Lewontin argued that our conceptual understanding of organisms under Darwinian evolutionary theory alienates the organism from the things that make it evolve: mutations happen randomly and are selected by the environment (Lewontin, 1983). This analysis motivated him to come up with his niche construction model, where the organism and the environment mutually shape one another (Tanaka, Godfrey-Smith, and Kerr, 2020). My insight about predicting evolution is drawn from his work directly. While he argues that the explanatory framework of Darwinian evolution epistemologically alienates the organism from its evolutionary process (i.e., organisms out-there in the world are not changed, only our ideas and understanding of them are), I argue that predicting evolution physically (i.e., ontologically) alienates the population from its evolutionary process. In other words, I am assuming that the environment is exerting some selection force and that this will lead to adaptive evolution in the direction of the phenotype space that will be useful for future populations. Predicting evolution does not allow the population of organisms to participate in this co-sculpting of itself and its environ-

ment that is proposed by niche construction. Instead, the population is sculpted by the individual(s) wielding the evolutionary prediction technology.

One problem with humanism is that it makes implicit assumptions about who gets to be granted human privileges and agency. There are many humanist historical documents that did not grant full personhood to all members of the *Homo sapiens* species. We know from human genomic data that all human cultures (and genders) are part of the human species. However, activists working for the liberation of women, racialized people, queer people, and other groups fought hard for their full personhood to be recognized and granted equal ethical consideration as white men. Some groups in the West/global north, like people with disabilities, transgender people, and refugees, are still fighting to gain access to their 'unassailable rights' under humanist ethics. The difficulty of humanism is that not everyone is treated as fully human by institutions (and on a personal level). This is why reprogenetics generally continues to be regarded as a beneficial and ethical practice in contemporary societies, even though it contains ethical dilemmas about the value and agency of disabled people (Valentine, 2021).

The reason why I dislike technology for evolutionary prediction is because my personal ethical position is some flavour of posthumanism. **Post**humanism is an ethical position that tries to move beyond the limitations of humanism by "extending the circle of moral concern" (Wikipedia, 2023b) to include other species. From a humanist perspective the ethical dilemma posed by predicting human evolution (i.e., eugenics) is solved by strictly delineating humans as the only agents worthy of controlling their own evolutionary fate. However, evolution is not anthropocentric. As Darwin himself struggled to reconcile in his moral understanding of the world, evolution teaches us that humans are just another evolving species. If we relax the assumption of which species are granted agency of their own evolutionary fate and take seriously the Darwinian evolutionary position that humans are not at the center, then predicting evolution *in general* becomes ethically dubious as it alienates populations from their own evolutionary process. Moreover, the emerging scientific consensus seems to be that human health, as well as our evolutionary fate, is inextricably linked to that of other organisms, ecosystems, and the planet (Destoumieux-Garzón et al., 2018). For me, this motivates the adoption of a posthumanist ethics.

It is worthwhile to examine how the claim 'the technology for evolutionary prediction alienates the population from its own evolutionary process' is understood under different ethical positions. I only give a few examples but future work can consider this more fully. The humanist position has commonalities with the posthumanist position because both believe that science should have utility for someone: humanists believe

it should have utility for humans, posthumanists believe it should have utility for the more-than-human world. But not all ethical positions assume that science should have utility for anybody. We can take the position that science and technology are intrinsically worthwhile for curiosity's sake. In this case technology for evolutionary prediction poses no ethical dilemma because it's just another fun problem to work on and explore. Similarly, if we take the nihilistic position that species or populations have no inherent value, then there is again no ethical dilemma. On the other hand, a transhumanist utilitarian position might consider eugenics a moral imperative. **Trans**humanism is the position that humans should use technology to enhance human lives. From this perspective the concern about alienation of an abstract evolutionary potential and intrinsically valuable inter-dependency between organisms and their environments are not very meaningful. Instead, transhumanists would argue that *not* applying evolutionary prediction technologies to human populations would be more alienating than applying them (why should the environment be in the driver's seat when humans themselves can be). Of course not all transhumanists would approve of eugenics (e.g., Haraway, 2013), but those transhumanists who value strong utilitarianism (i.e., greatest total good for the greatest number even if there is a cost to some) might.

It is valuable to consider our own ethical positions and those of other, specific people. For example, working with historians, we can also try to analyze the claim that 'technology for evolutionary prediction alienates the population from its evolutionary process' from the positions of Galton, Pearson, or Fisher. This would be speculative, of course, but I think it could help us to better understand and re-contextualize their work. Fisher's math is important and, while my conviction from the work I have created here is that everyone who uses Fisher's math must also read Fisher's writing to better understand his ideas, I think it is unkind to force people to read Fisher. For those of us belonging to marginalized groups, it is, simply put, emotionally difficult to read Fisher (and even more difficult to read Pearson or Galton!). Historians do valuable work and scientists can work with them to help us tell a better history that can be read by contemporary audiences, especially when writing textbooks or designing lesson plans for students. I am sure that there will always be a role to be played by scientists in writing the history of our own field. But historians have a specialized knowledge that complements our own.

I hope that this discussion has illustrated more clearly why scrutinizing our own position (as well as positionalities) and considering a larger diversity of positions can help us do better science. Some have characterized this work on positioning (and positionality) as "navel gazing" and, to be sure, there is some of that involved. However,

there is no position from nowhere. There is only the socially normalized perception of a default point of view. Science is currently funded through a system that requires us to justify the utility of our work so that governments or corporations can give us funding to do that work. We are already considering our own work from different positions when we write our funding grants to explain why our work is useful to society. From this perspective, I understand why my supervisors are excited about developing technologies for predicting evolution: this is a utilitarian application of an otherwise abstract field. Predicting evolution makes evolution fundable! Personally, I have always found the grant-writing part of my job very demoralizing. But now, with a better understanding of different positions, I think I would feel less frustrated by it. By taking our work and deliberately scrutinizing its value from a different point of view, we can better understand the aspects of our work that have lasting, impactful value. The true test of good science is the test of time. By viewing our work from different positions we, of course, cannot be sure that it will stand the test of time. But we can at least try to see if our work can withstand the scrutiny of a different way of seeing the world.

Finally, considering our own positionality makes science a more inclusive place. Folk wisdom (and maybe even academic research now (refs?)) tells us that a more diverse team creates better, more original work. I wonder if this is because a more diverse teams benefits from being able to situate its work in more diverse positionalities. This would suggest that the strength of diversity doesn't come from tokenism (i.e., it's not about collecting a variety of body abilities, genders, skin colours, etc. of people who all embody the same positions) but from our unique and situated ways of seeing the world (which will always include people of different body abilities, genders, and skin colours but what I'm saying is that these people can also be allowed to have different ideas, too). Moreover, this implies that diversity is best fostered in an environment where we can each voice our differences and allow our different experiences to shape how we do science. Perhaps, by fostering the intellectual and positional diversity of scientists we can climb the fitness landscape of scientific knowledge faster, together.

6.3 A feeling for the organism, a feeling for the environment

While investigating my discomfort with the premise of predicting evolution, the predominant thoughts and emotions I experienced in reaction to the subject matter was that I do not want to be controlled, I do not want the magic of the natural world to be explained away. This led me to learn more about what academics outside of science think about the relationship that science has with control. As alluded to on occasion in the Introduction Chapter and here, science and technology can be seen as means for

gaining greater power in the world and trying to dominate over nature (Lloyd, 1993; Leiss, 1994; Keller, 1996). This depicts scientists as controlling or manipulative. I think it is normal to want to have security and predictability in a world that is complex and dangerous. Therefore, I think there is truth in the statement that science is, as a whole, and that scientists are, as people, motivated by increasing our ability to control the natural world. On the other hand, this is a sweeping generalization that dehumanizes a variety of disciplines and individuals.

We are drawn to the natural sciences because we want to learn something more about the world. Our motivations are include: curiosity to understand subjects different from ourselves (Keller, 1984; Keller, 1996), curiosity to understand how the world works, solving an interesting puzzle (Kuhn, 1996), understanding why we are here and why the world is as it is, what may happen in the future, applications for bettering human life, etc. One way to order these motivations is to arrange them along a spectrum from control to attunement. On one extreme end is control. Explanation would be more on the control side of this spectrum and puzzle-solving would go roughly in the middle of the spectrum. On the other extreme end is immersion in and a feeling of connection with the world of our study subject(s). Barbara McClintock, the geneticist who won the Nobel prize for her discovery of transposable elements, was motivated to create scientific knowledge by gaining a "feeling for the organism" (see Keller, 1984; Keller, 1996). As humans we are intrinsically motivated to connect with other (human) subjects and our neurobiology allows us to 'attune' or harmonize our emotional states with those around us, like our parents, partners, and children (Kolk, 2014). In the attachment theory of psychology, this process of intersubjectivly coordinating our mental states and coming into human connection is called attunement.

McClintock's biographer, Evelyn Fox Keller, explains the unusual relationship McClintock had with her subject of study as a way of empathizing and connecting with the organism in order to tell its story. Keller compares the classical Greek scientific writings of Plato with Francis Bacon's modern works and those of contemporary sciences and concludes that the goal of science has shifted from communion with nature into domination over nature (Keller, 1996). Keller then goes on to argue that the logical extension of the feminist slogan that "the personal is political" is to see "the scientific as personal." It is with this political motivation that Keller discusses McClintock's experience of having a feeling for the organism.

My hypothesis that scientists seek attunement with their study systems draws directly from Keller's work. Instead of seeing this as an exceptional motivation for some few individuals, I argue that it is, at least in part, a common motivation for

all scientists. For example, I see an intriguing similarity between McClintock's feeling for the organism and the way that my PhD co-supervisor, Isabel Gordo, approaches the study of *E. coli* adapting to different host environments. Gordo describes this approach as "getting a feeling for the environment" through the evolutionary response of the bacteria. She is empathizing with the subjectivity of the bacterial evolutionary process in order to gain a foothold into understanding the evolutionary stressors that are found in the environment (see sections above). I argue that this empathetic attunement with our study systems is a common but unacknowledged motivation of our science. Moreover, I maintain that it is valuable and good science.

Many biologists learning to manipulate living matter complain that the lifecycle of their model system has come to dictate their life, like parents studying and matching the rhythms of their new baby (Kolk, 2014; Kain and Terrell, 2018). Many theoretical scientists feel more comfortable engaging in conversations about explicit abstractions, numbers, and data than they do engaging in the subtle and contextual world of emotions or politics [not that the first is in any way less subtle or less contextual]. These examples along with a myriad of other experiences of attunement are critical motivations and deciding factors that scientists use when choosing to pursue scientific research, their specific field of study, and their research interests. Moreover, these experiences of attunement are intrinsically valuable or meaningful to us. Experiences of attunement feature in our axiological judgements, which define what is worthwhile, good, or valuable, and will therefore feature in our knowledge-making (Peers, 2018; Branch et al., 2022). But these experiences rarely feature in the dominant stories we tell about the scientific process and are not acknowledged in the scientific articles we publish. Although we are rationally aware that we study non-human subjects with whom we cannot share a human-to-human intersubjective connection, the evolution and development of our cognition is based in social connection (Kolk, 2014; Kain and Terrell, 2018; Haines, 2019), and so even in the natural sciences we seek a feeling of deepening connection with our subject(s) and the world at large.

Although we study non-human subjects – sometimes non-living, inanimate, or theoretical subjects – all natural scientists who I have met are motivated, to varying degree, by the feeling of coming into connection with the world of their study subjects (i.e., attunement). For researchers whose subjects are humans and/or non-human apes, it is well established that the judgements of the observer impact their observations (Hrdy, 1999). Consequently, methodologies have been developed to make observerational judgements explicit (e.g., Laverty, 2003). Moreover, when the subject of study is human or near-human, their subjectivity cannot be ignored and so there is an ethical re-

quirement for the researcher to co-create knowledge that is meaningful and worthwhile for the studied community (Peers, 2018). Evolutionary theory and the natural sciences in general have not been impacted much by such methodological efforts (Godfrey-Smith, 2003), despite a long history of critiques on the capitalist (Kropotkin, 1890; Gross and Averill, 2003), racist (e.g., Kampourakis and Peterson, 2023), sexist (Hubbard, 2003), and more recently ableist (Branch et al., 2022) ideological assumptions of evolution.

The positivist ideology that is dominant in the natural sciences (i.e., that we are objective discoverers of natural laws; reviewed in Park, Konge, and Artino, 2020) tends to reduce the variety of legitimate motivations for doing science to those at the explanation and control end of the spectrum. In our contemporary technocratic, capitalist social context, the task of science is to discover natural laws in order to exact further mastery over nature for the ease and betterment of human life. This dominant story is a vast oversimplification of the subjective experiences and motivations of people working in basic science research. But this dominant story features heavily in how we see ourselves, the topics we research, how we phrase our questions, what research gets funded, and what we regard as valuable research.

A stated reason why we do science for grant applications is bettering human life. Much like the world of the evolutionary biologists who supported the eugenics movements, our world is currently in a perceived moment of deterioration. Anthropogenic climate change, loss of biodiversity, susceptibility to antibiotic resistant infection, zoonotic infections, threat of job security from technological advancements, and many other causes of environmental deterioration are the quotidian background of our lives. The natural world itself, which is often a source of anxiety-relief, is now for many a source of anxiety. Technologies for prediction and control of the natural world are touted as our only hope. However, technology is not limited to the industrial, commercial, and instrumental. Biologists, like all scientists, want to connect with their study subjects and to connect with other humans through their connection with the natural world. By focusing on our desire to connect with the world and using this as a legitimate way of doing science and a legitimate goal of the technologies that we build, we can create a better world that people are more likely to engage and connect with.

7 Concluding remarks

The topic of my thesis was shaped primarily by bureaucratic considerations. Already over time and without any submitted manuscripts from my ABR work, I was asked to add the Finnish wood ant paper (Chapter 2) to my thesis in order to meet the

university minimum requirements of one accepted paper. The only theme that I could think of to coherently tie together these subject areas was not one that I like. Although my supervisors are enthusiastic about predicting evolution, I have always felt wary of the subject. I knew that my concerns were related to the eugenic history of predicting evolution but I did not want to make my life difficult and I tried as best I could to ignore that aspect. Early drafts felt gushingly optimistic about the possibility of predicting evolution (it *is* exciting!) and so I would cry continuously while writing but I didn't yet understand why. Starting from the work of Darwin, science tells us that emotions are not random: emotions convey summarized information that can be essential for your survival. In *Origin* and more fully in *The Descent of Man*, Darwin argued that human morality itself is likely an evolved adaptation. The evolutionary data of that claim hasn't convinced me yet, but I believe that the same capacity for morality and intelligence is shared across humans from all populations, cultures, body shapes and colours, thinking styles, and even knowledge-making disciplines (like the humanities and social studies). It felt important that I was crying about the eugenics history of evolutionary theory while writing about its current project of evolutionary prediction. We want to do things that we are proud of, to work on important topics. This is important. The community of working, academic scientists sometimes embodies strict hierarchies and believes itself to be the only legitimate knowledge-making discipline (although in my experience the supremacy of science, or "scientism," is perpetuated more by lay supporters of science). Historians have discussed the eugenics history of evolution for over 50 years but we still cannot make sense of it ourselves within the field. Predicting evolution teaches us that historical contingency influences evolutionary outcomes. But where we start on the adaptive landscape is not the only determinant for where we will end up. By considering many different view-points, and maybe even different cultural starting conditions, I hope that we can build towards a better and more human(e) way of doing good science.

I have done my best to follow the normative format of a PhD thesis, by tying my three and a bit papers together under one question, even though I am critical of the myth-building that science does about itself (e.g., that a good scientist's research program is driven by a coherent question *versus* there are many ways to be a good scientist and do good science but the research programs we value most as a community are those that tell a coherent story). I am sure that this is not the thesis that my supervisors imagined when they told me 18 months ago that I should add the Finnish wood ant paper to my thesis. Someone at a party recently suggested that I could have just declared my thesis topic to be "ANTibiotic" resistance and, honestly, I agree with their

implication that science does not function according to the myths that we tell ourselves. I grieve that, contrary to the normative career expectations for graduate students, the completion of my PhD most likely marks the end of my creative independence within evolutionary biology. I keenly hope that the work presented in the Introduction and Discussion Chapters will make a lasting contribution to our field.

I believe there is value in doing what you're told and trusting the process (I'm sure it surprises my supervisors to hear that! It sounds strange to say it in the document that will certify me as a scientist and therefore forever legitimize my iconoclastic commitment to skepticism and empiricism). I have found value in stretching my knowledge and abilities in order to package a paper about hybrid speciation in ants together with two+ papers about ABR. Working under a time constraint, my main supervisor, Claudia Bank, encouraged me to be less perfectionist and, by implication, more vulnerable and more confident in my scientific merit. Writing this thesis has been difficult, especially because this is my second time around. But writing this thesis has also been immensely rewarding. The most important thing I learned is that I am allowed to have my own opinions and to express them in the manner that I judge to be most appropriate (even if that sometimes means slipping off the authoritative mask of scientific language). I took a second shot at a PhD because, over my many years in academia, I had seen how the thesis writing process itself seemed to transform my peers from shaky, self-doubting graduate students, like me, into confident, independent thinkers. The thesis did not make them brilliant, they were already brilliant in their own ways. But the thesis writing process seemed to reflect their light back at them so that they could finally glimpse it for themselves. I am proud to have been given the chance to craft my own mirror. I am grateful to my supervisors, the IGC, and the scientific community as a whole for giving me a second go. And I am grateful to you (the reader) for reading and intellectually engaging with my ideas – that attentiveness alone is a gift that I cannot find the words to express!

References

Adrian-Kalchhauser, I. and C. Bank (2022). *MiCo4Sys - Microbial Community Composition and Colonization in Compartmentalized Aquaculture Systems*. SNSF grant.

Agor, J. K. and O. Y. Özaltın (2018). "Models for predicting the evolution of influenza to inform vaccine strain selection". In: *Hum. Vaccines Immunother.* 14.3, pp. 678–683. URL: https://doi.org/10.1080/21645515.2017.1423152.

Bacevic, J. (2021). "Epistemic injustice and epistemic positioning: towards an intersectional political economy". In: *Curr. Sociol.*, p. 00113921211057609. URL: https://doi.org/10.1177/00113921211057609.

Bank, C. (2022). "Epistasis and Adaptation on Fitness Landscapes". In: *Annu. Rev. Ecol. Evol. Syst.* 53, pp. 457–479. URL: https://doi.org/10.1146/annurev-ecolsys-102320-112153.

Bank, C. et al. (2016). "On the (un)predictability of a large intragenic fitness landscape". In: *Proceedings of the National Academy of Sciences* 113.49, pp. 14085–14090. URL: https://doi.org/10.1073/pnas.1612676113.

Beauvoir, S. de (2009). *The Second Sex*. Trans. by C. Borde and S. Malovany-Chevallier. London, UK: Jonathan Cape.

Bodmer, W. et al. (2021). "The outstanding scientist, R. A. Fisher: his views on eugenics and race". In: *Heredity* 126.4, pp. 565–576. URL: https://doi.org/10.1038/s41437-020-00394-6.

Branch, H. A. et al. (2022). "Discussions of the "not so fit": how ableism limits diverse thought and investigative potential in evolutionary biology". In: *Am. Nat.* 200.1, pp. 101–113. URL: https://doi.org/10.1086/720003.

Brochado, A. R. et al. (2018). "Species-specific activity of antibacterial drug combinations". In: *Nature* 559.7713, pp. 259–263. URL: https://doi.org/10.1038/s41586-018-0278-9.

Burford Reiskind, M. O. et al. (2021). "Nothing in evolution makes sense except in the light of biology". In: *Bioscience* 71.4, pp. 370–382. URL: https://doi.org/10.1093/biosci/biaa170.

Chevin, L. M. (2011). "On measuring selection in experimental evolution". In: *Biol. Lett.* 7.2, pp. 210–213. URL: https://doi.org/10.1098/rsbl.2010.0580.

Concepción-Acevedo, J. et al. (2015). "Malthusian parameters as estimators of the fitness of microbes: a cautionary tale about the low side of high throughput". In: *PlOS One* 10.6, e0126915. URL: https://doi.org/10.1371/journal.pone.0126915.

Das, S. G. et al. (2020). "Predictable properties of fitness landscapes induced by adaptational tradeoffs". In: *eLife* 9, e55155. URL: https://doi.org/10.7554/eLife.55155.

De Visser, J. A. G. M. and J. Krug (2014). "Empirical fitness landscapes and the predictability of evolution". In: *Nat. Rev. Genet.* 15.7, pp. 480–490. URL: https://doi.org/10.1038/nrg3744.

Deaver, J. A., S. Y. Eum, and M. Toborek (2018). "Circadian disruption changes gut microbiome taxa and functional gene composition". In: *Front. Microbiol.*, p. 737. URL: https://doi.org/10.3389/fmicb.2018.00737.

Dennett, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. London, England: Penguin Books Ltd.

Destoumieux-Garzón, D. et al. (2018). "The one health concept: 10 years old and a long road ahead". In: *Front. Vet. Sci.*, p. 14. URL: https://doi.org/10.3389/fvets.2018.00014.

Durão, P. et al. (2015). "Multiple resistance at no cost: rifampicin and streptomycin a dangerous liaison in the spread of antibiotic resistance". In: *Mol. Eiol. Evol.* 32.10, pp. 2675–2680. URL: https://doi.org/10.1093/molbev/msv143.

Durão, P. et al. (2021). "Competition dynamics in long-term propagations of *Schizosaccharomyces pombe* strain communities". In: *Ecology and Evolution* 11.21, pp. 15085–15097. URL: https://doi.org/10.1002/ece3.8191.

Feder, A. F, P. S. Pennings, and D. A. Petrov (2021). "The clarifying role of time series data in the population genetics of HIV". In: *PLOS Genet.* 17.1, e1009050. URL: https://doi.org/10.1371/journal.pgen.1009050.

Fichman, M. (2021). "Technoscientific Control of Nature: The Ultimate Paradox". In: *Tech. Sci.* 34, pp. 123–456. URL: https://doi.org/10.1111/tosc.12345.

Flynn, J. M. et al. (2020). "Comprehensive fitness maps of Hsp90 show widespread environmental dependence". In: *eLife* 9, e53810. URL: https://doi.org/10.7554/eLife.53810.

Fragata, I. et al. (2019). "Evolution in the light of fitness landscape theory". In: *Trends Ecol. Evol.* 34.1, pp. 69–82. URL: https://doi.org/10.1016/j.tree.2018.10.009.

Garnett, S. T. et al. (2018). "A spatial overview of the global importance of Indigenous lands for conservation". In: *Nat. Sustain.* 1.7, pp. 369–374. URL: https://doi.org/10.1038/s41893-018-0100-6.

Gavrilets, S. (2019). *Fitness Landscapes and the Origin of Species*. Princeton, MA: Princeton University Press.

Godfrey-Smith, P. (2003). *Theory and Reality: an introduction to the philosophy of science*. 1st. The University of Chicago Press.

Good, B. H. et al. (2017). "The dynamics of molecular evolution over 60,000 generations". In: *Nature* 551.7678, pp. 45–50. URL: https://doi.org/10.1038/nature24287.

Gorter, F. A. et al. (2018). "Local fitness landscapes predict yeast evolutionary dynamics in directionally changing environments". In: *Genetics* 208, pp. 307–322. URL: https://doi.org/10.1534/genetics.117.300519.

Gould, S. J. (1991). "The smoking gun of eugenics". In: *Nat. Hist.* 12, pp. 8–17.

Grafen, A. (2003). "Fisher the evolutionary biologist". In: *J. R. Stat. Soc. D (The Statistician)* 52.3, pp. 319–329. URL: https://doi.org/10.1111/1467-9884.00362.

Gross, M. and M. B. Averill (2003). "Discovering Reality". In: ed. by S. Harding and M. B. Hintikka. 2nd. Springer Science+Business Media. Chap. Evolution and the patriarchal myths scarcity and competition.

Guerrero, R. F. et al. (2019). "Proteostasis environment shapes higher-order epistasis operating on antibiotic resistance". In: *Genetics* 212, pp. 565–575. URL: https://doi.org/10.1534/genetics.119.302138.

Haines, S. K. (2019). *The Politics of Trauma: Somatics, Healing, and Social Justice*. Berkeley, CA: North Atlantic Books.

Hall, A. E. et al. (2019). "Environment changes epistasis to alter trade-offs along alternative evolutionary paths". In: *Evolution* 73, pp. 2094–2105. URL: https://doi.org/10.1111/evo.13825.

Hall, B. G. et al. (2014). "Growth rates made easy". In: *Mol. Biol. Evol.* 31.1, pp. 232–238. URL: https://doi.org/10.1093/molbev/mst187.

Hanisch, C. (2000). "Radical Feminism: A Documentary Reader". In: ed. by B. A. Crow. New York, New York: New York University Press. Chap. The personal is political. URL: https://books.google.si/books?id=qG27dckXXOgC&lpg=PA113&ots=tI78BAdNAp&dq=The%20Personal%20is%20Political&lr.

Haraway, D. (2013). "A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century". In: *The transgender studies reader*. Routledge, pp. 103–118.

Harmand, N. et al. (2017). "Fisher's geometrical model and the mutational patterns of antibiotic resistance across dose gradients". In: *Evolution* 71.1, pp. 23–37. URL: https://doi.org/10.1111/evo.13111.

Hilau, S. et al. (2022). "Density-dependent effects are the main determinants of variation in growth dynamics between closely related bacterial strains". In: *PLOS Comp. Biol.* 18.10, e1010565. URL: https://doi.org/10.1371/journal.pcbi.1010565.

Hrdy, S. B. (1999). *The Woman That Never Evolved*. Cambridge, MA: Harvard University Press.

Hubbard, R. (2003). "Discovering Reality". In: ed. by S. Harding and M. B. Hintikka. 2nd. Springer Science+Business Media. Chap. Have Only Men Evolved?

IPBES (2019). *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Ed. by E. S. Brondizio et al. Bonn, Germany: IPBES secretariat. URL: https://doi.org/10.5281/zenodo.3831673.

J, Engelstädter (2014). "Fitness landscapes emerging from pharmacodynamic functions in the evolution of multidrug resistance". In: *J. Evol. Biol.* 27, pp. 840–853. URL: https://doi.org/10.1111/jeb.12355.

Jafar, A. J. N. (2018). *What is positionality and should it be expressed in quantitative studies?* URL: http://dx.doi.org/10.1136/emermed-2017-207158.

Jaggar, A. M. (2015). "Love and knowledge: Emotion in feminist epistemology". In: *Women, knowledge, and reality*. Routledge, pp. 166–190.

Jin, D. J. and C. A. Gross (1989). "Characterization of the pleiotropic phenotypes of rifampin-resistant *rpoB* mutants of *Escherichia coli*". In: *J. Bacteriol.* 171, pp. 5229–5231. URL: https://doi.org/10.1128/jb.171.9.5229-5231.1989.

Kain, K. L. and S. J. Terrell (2018). *Nurturing Resilience: Helping Clients Move Forward from Developmental Trauma*. Berkeley, CA: North Atlantic Books.

Kamath, A. et al. (2022). "Nature, data, and power: how hegemonies shaped this special section". In: *Am. Nat.* 200.1, pp. 81–88. URL: https://doi.org/10.1086/720001.

Kampourakis, K. and E. L. Peterson (2023). "The racist origins, racialist connotations, and purity assumptions of the concept of "admixture" in human evolutionary genetics". In: *Genetics*, iyad002. URL: https://doi.org/10.1093/genetics/iyad002.

Keller, E. F. (1984). *A Feeling for the Organism, 10th Aniversary Edition: The Life and Work of Barbara McClintock*. 2nd. Times Books.

— (1996). *Reflections on gender and science*. 2nd. Yale University Press.

Kevles, D. J. (1985). *In the name of eugenics: genetics and the uses of human heredity*. 1st. California: University of California Press.

Kittay, E. F. (2009). "The personal is philosophical is political: A philosopher and mother of a cognitively disabled person sends notes from the battlefield". In: *Metaphilosophy* 40.3-4, pp. 606–627. URL: https://doi.org/10.1111/j.1467-9973.2009.01600.x.

Knöppel, A., J. Näsvall, and D. I. Andersson (2017). "Evolution of antibiotic resistance without antibiotic exposure". In: *Antimicrob. Agents Chemother.* 61.11, e01495–17. URL: https://doi.org/10.1128/AAC.01495-17.

Kolk, B. A. Van der (2014). *The body keeps the score: Brain, mind, and body in the healing of trauma*. New York, NY: Penguin Group.

Kritikos, G. et al. (2017). "A tool named Iris for versatile high-throughput phenotyping in microorganisms". In: *Nat. Microbiol.* 2, p. 17014. URL: https://doi.org/10.1038/nmicrobiol.2017.14.

Kropotkin, P. (1890). "Mutual Aid Among Animals". In: *The Nineteenth century and after: a monthly review* 28.163, pp. 337–354.

Kuhn, T. S. (1996). *The Structure of Scientific Revolutions*. 3rd. University of Chicago Press.

Kulmuni, J. and P. Pamilo (2014). "Introgression in hybrid ants is favored in females but selected against in males". In: *PNAS* 111.35, pp. 12805–12810. URL: https://doi.org/10.1073/pnas.1323045111.

Kulmuni, J. et al. (2020). "Instability of natural selection at candidate barrier loci underlying speciation in wood ants". In: *Mol. Ecol.* 29.20, pp. 3988–3999. URL: https://doi.org/10.1111/mec.15606.

Latour, B. and S. Woolgar (1987). *Laboratory Life: The Construction of Scientific Facts*. Princeton, New Jersey: Princeton University Press.

Laverty, S. M. (2003). "Hermeneutic phenomenology and phenomenology: A comparison of historical and methodological considerations". In: *Int. J. Qualitat. Meth.* 2.3, pp. 21–35.

Leiss, W. (1994). *Domination over nature*. 2nd. University of Toronto Press.

Levins, R. and R. Lewontin (1985). *The dialectical biologist*. Delhi, India: Aakar Books.

Levit, G. S. (2015). "eLS". In: Chichester: John Wiley & Sons, Ltd. Chap. Sergei Chetverikov (1880–1959). URL: 10.1002/9780470015902.a0025066.

Lewontin, R. C (1983). "The organism as the subject and object of evolution". In: *Scientia* 77.18.

Lloyd, G. (1993). *The Man of Reason: 'Male' and 'Female' in Western Philosophy*. 2nd. Routledge.

Martin, G. and T. Lenormand (2015). "The fitness effect of mutations across environments: Fisher's geometrical model with multiple optima". In: *Evolution* 69, pp. 1433–1447. URL: https://doi.org/10.1111/evo.12671.

Martin-Roy, R. et al. (2021). "Differences in thermal tolerance between parental species could fuel thermal adaptation in hybrid wood ants". In: *Am. Nat.* 198.2, pp. 278–294. URL: https://doi.org/10.1086/715012.

Martínez-García, E. and V. de Lorenzo (2016). "The quest for the minimal bacterial genome". In: *Curr. Opin. Biotechnol.* 42, pp. 216–224. URL: https://doi.org/10.1016/j.copbio.2016.09.001.

Nichols, R. J. et al. (2011). "Phenotypic landscape of a bacterial cell". In: *Cell* 144, pp. 143–156. URL: https://doi.org/10.1016/j.cell.2010.11.052.

Orr, H. A. (2009). "Fitness and its role in evolutionary genetics". In: *Nat. Rev. Genet.* 10.8, pp. 531–539. URL: https://doi.org/10.1038/nrg2603.

Park, Y. S., L. Konge, and A. R. Artino (2020). "The positivism paradigm of research". In: *Acad. Med.* 95.5, pp. 690–694. URL: https://doi.org/10.1097/ACM.0000000000003093.

Pattberg, Philipp (2007). "Conquest, domination and control: Europe's mastery of nature in historic perspective". In: *J. Polit. Ecol.* 14, pp. 1–9. URL: https://doi.org/10.2458/v14i1.21681.

Peers, D. (2018). "Engaging axiology: enabling meaningful transdisciplinary collaboration in adapted physical activity". In: *Adapt. Phys. Activ. Q.* 35.3, pp. 267–284. URL: https://doi.org/10.1123/apaq.2017-0095.

Pinheiro, F. et al. (2021). "Metabolic fitness landscapes predict the evolution of antibiotic resistance". In: *Nat. Ecol. Evol.* 5, pp. 677–687. URL: https://doi.org/10.1038/s41559-021-01397-0.

Portinha, B. et al. (2022). "Whole-genome analysis of multiple wood ant population pairs supports similar speciation histories, but different degrees of gene flow, across their European ranges". In: *Mol. Ecol.* 31.12, pp. 3416–3431. URL: https://doi.org/10.1111/mec.16481.

Rispoli, G. (2021). "Evolutionary Developmental Biology: A Reference Guide". In: ed. by L. N. de la Rosa and G. B. Müller. Springer Nature Switzerland. Chap. Sergey S. Chetverikov (1880–1959).

Rodríguez-Verdugo, A., B. S. Gaut, and O. Tenaillon (2013). "Evolution of *Escherichia coli* rifampicin resistance in an antibiotic-free environment during thermal stress". In: *BMC Evol. Biol.* 13.1, pp. 1–11. URL: https://doi.org/10.1186/1471-2148-13-50.

Roll-Hansen, N. (2010). "The Oxford Handbook of the History of Eugenics". In: ed. by A Bashford and P. Levine. Oxford University Press. Chap. Eugenics and the Science of Genetics.

Rose, H. (1983). "Hand, brain, and heart: A feminist epistemology for the natural sciences". In: *Signs* 9.1, pp. 73–90. URL: https://doi.org/10.1086/494025.

Rouse, J. (1996). "Feminism, Science, and the Philosophy of Science". In: ed. by L. H. Nelson and J. Nelson. Springer. Chap. Feminism and the social construction of scientific knowledge.

Shapin, S. and S. Schaffer (2011). *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life*. 2nd. Princeton University Press.

Shaw, R. G. (2019). "From the past to the future: considering the value and limits of evolutionary prediction". In: *Am. Nat.* 193.1, pp. 1–10. URL: https://doi.org/10.1086/700565.

Siegal, M. L. and J. Y. Leu (2014). "On the nature and evolutionary impact of phenotypic robustness mechanisms". In: *Annu. Rev. Ecol. Evol. Syst.* 45, pp. 495–517. URL: https://doi.org/10.1146/annurev-ecolsys-120213-091705.

Soifer, V. (1994). *Lysenko and the tragedy of Soviet science*. New Brunswick, N.J.: Rutgers University Press. URL: https://archive.org/details/lysenkotragedyof00soif.

Sousa, J. M. et al. (2017). "Multidrug-resistant bacteria compensate for the epistasis between resistances". In: *PLOS Biol.* 15.4, e2001741.

Strom, A. R. and I. Kaasen (1993). "Trehalose metabolism in *Escherichia coli*: stress protection and stress regulation of gene expression". In: *Mol. Microbiol.* 8.2, pp. 205–210. URL: https://doi.org/10.1111/j.1365-2958.1993.tb01564.x.

TallBear, K. (2013). *Native American DNA: tribal belonging and the false promise of genetic science*. Minneapolis, MN: University of Minnesota Press.

Tanaka, M. M., P. Godfrey-Smith, and B. Kerr (2020). "The dual landscape model of adaptation and niche construction". In: *Phil. Sci.* 87.3, pp. 478–498. URL: https://doi.org/10.1086/708692.

Tenaillon, O. et al. (2012). "The molecular diversity of adaptive convergence". In: *Science* 335, pp. 457–461. URL: https://doi.org/10.1126/science.1212986.

Tenorio, E. et al. (2003). "Systematic characterization of *Escherichia coli* genes/ORFs affecting biofilm formation". In: *FEMS Microbiol. Lett.* 225.1, pp. 107–114. URL: https://doi.org/10.1016/S0378-1097(03)00507-X.

Thompson, C. L. et al. (2021). "Back to the future: Reintegrating biology to understand how past eco-evolutionary change can predict future outcomes". In: *Integr. Comp. Biol.* 61.6, pp. 2218–2232. URL: https://doi.org/10.1093/icb/icab068.

Trindade, S., A. Sousa, and I. Gordo (2012). "Antibiotic resistance and stress in the light of Fisher's model". In: *Evolution* 66, pp. 3815–3824. URL: https://doi.org/10.1111/j.1558-5646.2012.01722.x.

UNESCO (1978). *Declaration on Race and Racial Prejudice*. Accessed on January 10, 2023. URL: https://en.unesco.org/about-us/legal-affairs/declaration-race-and-racial-prejudice.

Valentine, D. (2021). "Technologies of Reproduction: Race, Disability, and Neoliberal Eugenics". In: *J. Phil. Disability*. URL: https://doi.org/10.5840/jpd20217122.

Wikipedia (2023a). *Humanism*. Accessed on February 26, 2023. URL: https://en.wikipedia.org/wiki/Humanism.

— (2023b). *Postumanism*. Accessed on February 26, 2023. URL: https://en.wikipedia.org/wiki/Posthumanism.

Acknowledgements

Acknowledgements are given in the reverse order of "the nested model of attachment and trauma" by psychotherapist Jessica Fern in *Poly*secure: *Attachment, Trauma, and Consensual Nonmonogamy*.

Global and Collective:

Thank you to the Earth for lovingly holding all of us, nurishing us, and taking all our garbage, thanklessly and all of the time. In particular, thank you to the ocean view from the train between IGC and Lisboa for buoying my spirits, the walk through the Estação Agronómica and the Carcavelos vineyard for drying my tears, and the little forest next to my apartment in Bern for inspiring my thesis writing.

This thesis was accomplished during a global pandemic that had been predicted by experts for more than a decade but still caught us unprepared. Although the nervous breakdown I had was a helluva ride, thank you for giving me an embodied understanding that science and technology should be in service of all. Ditto times one thousand for climate change.

Societal:

Thank you to science as a whole social enterprise. Like any human endeavour, it's got a lot of problems that we need to work on improving. But it's also an incredible place full of incredible people working together for a better world.

I acknowledge the knowledge-makers whose work was critical in shaping my world-view between 2017-2023 but that I couldn't find a way to cite directly (listed by category not to constrain or box-in, but to acknowledge each situated, collective movement).
Indigenous Sovereignty: All My Relations podcast, Ejército Zapatista de Liberación Nacional, Seeding Sovereignty, Indigenizing Arts Education, Native Movement, and others.
Black Liberation Movement / Black Lives Matter: adrienne maree brown, Prentis Hemphill (the Embodiment Institute), Generative Somatics, Tricia Hersey (the Nap Ministry), Bayo Akomolafe, Andrewism, the Emergent Strategy Ideation Institute, Sonya Renee Taylor, Black Anarchist Federation, Audre Lorde, bell hooks, and others.
(Dis)Ability Justice: Sunaura Taylor, JD Davids (the Cranky Queer), Jesse Meadows (Sluggish), Marta Rose (Divergent Design Studios), Mia Mingus, Nick Walker, Carmo Gê Pereira, and others.
Queerness: Kai Cheng Thom, Abigail Thorn (PhilosophyTube), Margeaux Feldman, Cedar Wren McCloud's *Numinous Tarot*, Clementine Morrigan, Leslie Feinberg, Inês Rolo, and others.
Consent: *Yes Means Yes: Visions of Female Sexual Power and a World Without Rape* edited by Jaclyn Friedman and Jessica Valenti, Sarah Schulman, A. V. Flox, Emily Nagoski, Betty Martin, and others.
Anti-Hegemony: Stefania Barca, Karl Marx, Silvia Federici, Cristy C. Road's *Next World Tarot*, For The Wild podcast, Free Radicals activist collective,

Upstream podcast, Working Class History podcast, David Graeber, and others.
Spirituality and Compassion: Taoism as explained by Benjamin Hoff in *The Tao of Pooh*, Jack Kornfield, the Center for Mindful Self-Compassion, Thích Nhat Hanh and his Order of Interbeing, the Way Out Is In podcast, and others.

Local Communities and Culture:

Thank you IGC! I am incredibly sad to be leaving, like everyone I know who has left before me. From my 3-month exchange in 2012, to my 9-months as a lab manager, and my PhD 2017-2023, you inspire me to imagine a joyful and interactive way to do great science. For me, you have always been a special place to be.

Thanks to the Institute for Ecology and Evolution at Universität Bern for taking me in during the last year and a half. I know that my heartbreak over leaving IGC has made it difficult for me to give you the chance you deserve. We've got time.

Thank you to the inspiration and support of the SMBE Inclusion, Diversity, Equity and Access (IDEA) Initiative. It's very cool to be part of a like-minded group of molecular and evolutionary biologists.

Thank you so much to Portugal and the Portuguese people. A language, culture, and recent past so similar to my own coutry of origin that taught me lessons that I was missing about community, kindness, and humanity. Thank you also to Southern India, where I attended a seminar on "Living Matter" in 2018 at the International Centre for Theoretical Sciences, and where I learned about devotion.

Thank you to the spiritual support of the Lisbon Quakers, Central European Quakers, and Disgraça grupo de autodefesa feminista run by Samuel MacDowell. You showed me that the grace and magic of the universe can always be felt in stillness.

Thank you to the Portuguese collectives and groups that I had the privilege to be a part of: Climáximo (especially their 2020 climate camp), Festival Feminista de Lisboa 2019, PolyPT, Consent Gathering, and Expressar Consentimento. I wish I could've been better and given back at least as much as I learned from you.

Academic Home:

Thank you to my main advisor, Claudia Bank. I could make a long paragraph here listing everything I'm grateful to have learned from you, but you've also taught me that sometimes it's just faster and more satisfying to do things in person.

To the **EvilDynamics / THEE** past, present, and affiliated group members: Inês Fragata, Andreia Teixeira, Alexandre Blanckaert, Mark Schmitz, André Amado, Loïc Marrec, Dragan Stajić, Marco Louro, Emma Berdan, Massimo Amicone, Ana Morales-Arce, Li Juan, Christian Diwo, Davide Cusseddu,

Mado Kapopoulou, Dave McLeod, Jérôme Stäheli, Loraine Hablützel and others. It was so much fun working with all of you!

Thank you to my co-advisor, Isabel Gordo. I learned so much from you about how to do good science by being brave and having a bold, coherent vision of evolution.

To the **Evolutionary Biology** past, present, and affiliated group members: Dani Güleresi, Paulo Durão, Roberto Balbontín, Ricardo Ramiro, Jorge Sousa, Luis Cardoso, Anastasia Kottara, Anke Konrad, Tanja Đapa, Francisco Cerqueira, Beatriz Abreu, Catarina Pinto, Jernej Jurić, Ana Margarida Sousa, Nelson Frazão, Hugo Barreto, and others. Thanks for teaching me how to work with microbes and como ser um dos fixes.

Thank you to members of the PREPARE consortium for your insightful discussions and help with design of the ABR fitness landscapes: Thomas Batallion, Alex Wong, Aaron Hinz, Rees Kassen, and others.

Thank you to my thesis committee, Lounès Chikhi and Erida Gjini, for striking just the right balance of friendliness and formality.

Thank you to the IBB 2017 program and to my peers: Sahar Tehrani, Camila Ramos, Catarina Carmo, Anastasiia Lozovska, Tânia Paulo, Anton Kermanov, Temitope Etibor, Margarida Araújo, Sebastiaan van den Berg, André Dias, and André Madaleno. It was an amazing – and, with COVID, an absurd – experience that I'm so glad I had the chance to share with you all, even if I could not always be as present as I wish. Ana Aranda, thank you for managing my thesis submission process, my paperwork in general these past years, and the rest of the IBB program all by yourself these past weeks.

Relationships:

Thank you to Inês Amaro for creating the witchy cover art, as well as for suggesting the wonderful *Jurassic Park* quote.

Thank you to Leila Shirai for translating my thesis abstract from English to Portuguese. Even though it feels like a failure that I did not have time to do the translation myself, it was such a gift to re-connect with you, my dear.

Thank you, as well, to Juliana Kraeva Jolly for being my thesis-writing buddy, and for encouraging me to include ideas from existential therapy in my Discussion Chapter.

Thank you to all my friends and partners/lovers/crushes who listened to me go on about my thesis, supported me, and helped me to grow: Dave Bapst, Diogo Monte-Mór, Jurij Smrke, Izidor Barši, Sinan Eden, Flavia Schlichta, Zachary M. Bailey, Bea Moira, Klara Chulechuga Wi, Atus Mariqueo-Russell, Tara Sadoway, Francisco Maiochi, Cameron Drysdale, Lindsay Keegan, Luís Blumel, Davide Rizzo, Mariana Silva, and others. Finally, thank you to Natália WuWei for inspiring me to unabashedly be a queer, neurodivergent scientist.

Thank you to my family that I was born into and that I made along the way: Jessica Thompson, Jake Szamosi, Nazia Darvesh, Karen Lau, David Jolly, Vladimir Crețulescu, Šmajgl family, Noemie Prévost, Mihai Ghenu, Nina

Ghenu, Șerban Ghenu, and others. Finally, thank you to my partner, Samo Šmajgl, for making everything possible.

Personal and Self:

I acknowledge that this PhD thesis was accomplished during various personal issues, including me wanting to die. Unfortunately, I know many people like me, both inside and outside of academia. I would like to thank my therapists, teachers, and peer support workers for helping me survive and eventually flourish.

Sometimes we joke that our PhD is our baby. If my time spent working towards a PhD was a literal baby, it would be a tween. I don't know anyone like me. I have my perseverance and scientific passion to thank for this gift of finally completing a PhD thesis.

Last of all, thank you to my love of writing, even though it sometimes feels impossible; to my need for opennes and my bravery to say things, even when they're hard; and, finally, to my big-picture thinking, even when it leads me down seemingly erratic rabbit-holes.