

Competition and Coexistence in an Unpredictable World

Alexander Douglas Washburne

A dissertation presented to the faculty of Princeton University in candidacy
for the degree of Doctor of Philosophy

Recommended for acceptance by the department of Quantitative and
Computational Biology

Adviser: Prof. Simon A. Levin

June, 2015

© Copyright by Alexander Douglas Washburne, 2015. All rights reserved.

Abstract

All living things “struggle for existence” as they compete with other organisms over limiting resources. Understanding how the diversity and dynamics of living systems are shaped by competition can help us better understand evolutionary problems of altruism, conservation management of competing species, and even economic policy making to promote productive competition in free markets. This thesis examines competition and its effects on diversity and dynamics in four systems: the slime mold *Dictyostelium discoideum*, predator-prey systems such as wolves in Yellowstone, the human microbiome and the S&P 500. Diversity in slime molds may be maintained despite competition for space in the spore capsules if the natural habitat of slime molds is variable in space and time; resource availability might mediate quorum sensing, and such molecular switches and bet-hedging can be advantageous over competitors without such plasticity. Competition between prey can be mediated by predators, but the ability of predators to stabilize prey communities depends on the size of the community relative to the attack rate of the predator, implying that some predators need especially large reserves to exhibit their full ecological effects. Snapshots of the human microbiome and the S&P 500 might suggest that they could arise from neutral competition, but time-series analysis reveals that many seemingly neutral communities may exhibit non-neutral dynamics. Understanding patterns of diversity and dynamics of adaptive systems requires understanding competition and coexistence in an unpredictable world.

Preface

The world is a finite, unpredictable vat of resources. All living things try to acquire resources and reproduce, but not all organisms will succeed. All of life finds itself in a “struggle for existence” [12] - competing with other members of the same species or with other species in the same trophic guild. When Gause grew mixed cultures of yeast in competition over the same resources, the fitter variants competitively excluded the less fit variants; from these experiments, Gause hypothesized that competition cleared communities of any overgrowth of excessive diversity [13]. Lotka-Volterra style differential equation models of organisms in competition over limiting resources confirmed that species coexistence is tenuous and competitive exclusion may be a likely outcome of competition [14]. This led to the explicit statement of the competitive exclusion principle in 1960 [15]: two species in competition over the same limiting resource cannot coexist.

However, if competitive exclusion is so easily obtained in models, then why do we see such a remarkable diversity of life?! In 1961, just one year after the competitive exclusion principle was stated, Hutchinson [18] pointed out that many species of plankton coexist over just a few limiting resources - how is this possible? If tropical trees compete over just a few limiting resources - space, water, nitrogen, phosphorus, and a few others - how over 220 species of trees coexist in the meager 50 ha plot of Barro Colorado Island? How can 8 large ungulate species coexist in Yellowstone while competing over the same forage and through shared predators? How do hundreds of operational taxonomic units of bacteria coexist on our body? Why have we not seen competitive exclusion in these systems; why is there not one super-tree that dominates Amazonia, a super-moose that dominates Yellowstone, or a super-bug that dominates our large intestines?

In the years since Hutchinson’s paradox of the plankton, ecologists have carved out many mechanisms that maintain coexistence despite competition, mechanisms that range from MacArthur’s niche partitioning in warblers [23, 22] or Janzen and Connell’s suites of distant-dependent specialist predators [19, 6] to Paine and Levin’s rarefaction-like disturbances in intertidal landscapes [21, 25] to spatial structure [1] and spatial and temporal variation in environments [5], and more.

It’s safe to say that the paradox of the plankton is no longer a paradox. Instead of being confused by the diversity of plankton or tropical trees, ecologists now have the opposite problem: given the myriad possible mechanisms for coexistence and mediation of competition, which ones are most important in our study system? How does the suite of stabilizing mechanisms [4] at play in our system affect our management of natural systems? In which other competitive systems are analogous stabilizing mechanisms at play, and how can knowledge of stabilization in ecology inform our management of other competitive systems? For instance, we find a suite of analogous problems of competition and coexistence in human systems - of cultures in sociology or companies in economics: since companies in a market compete over a finite number of customers or a

finite amount of investment capital, fitter companies can expand and advertise and, in theory, competitively exclude less successful companies. How can insights from ecology of non-human life generalize to our understanding of human competition? Such knowledge of human competition would affect our economic policy making and portfolio management - when human resources become limiting, shares might no longer grow exponentially, and so principles of competition and evolution of competitive systems may be needed to accurately price options and other financial derivatives [2]. This thesis examines some of these modern questions of competition and coexistence in a finite, stochastic world in systems ranging from ungulates in yellowstone or arthropods on small islands in the Bahamas, to slime molds, gut bacteria, and large cap companies in the S&P 500.

The first chapter explores a paradoxical coexistence of many genotypes of the slime mold *Dictyostelium discoideum* in competition for limiting space in the fruiting body. We show that multiple genotypes can coexist in a patchy environment with spatial and temporal variation in the time between resource replenishment; instead of a problem of altruism arising in competition over space in the spore capsule, there may be dispersal-dormancy tradeoffs that can explain the persistence of some so-called loner cells that “choose” not to aggregate into the spore capsule but instead remain dormant; a mosaic of environments that vary in their expected rate of resource arrival can explain many genotypes with different degrees of investment in dispersal vs. dormancy. We finish with a discussion of future empirical investigations that can test our hypothesis and elucidate how exactly slime molds with the same genotype might “choose” to aggregate or to ignore an autoinducing signal from other cells hoping to form a quorum. If such conditional quorum activation does exist, which I hypothesize may be co-mediated by local resource availability, it would have implications for the evolution of quorum sensing in other model organisms such as *Vibrio* species.

The second chapter examines how stabilizing forces such as prey-switching predation - which maintain coexistence in differential equation models - affect small populations prone to fluctuations in population size from variation in fecundity, lifespan, intraguild competitive interactions, and predatory encounters. Our main result is obvious in hindsight: by killing prey, predators reduce mean population sizes of prey, and too high of attack rates - even attack rates by “stabilizing” predators - can drive finite prey populations to rarity and eventually extinction. Because the probability of a species becoming rare is directly related to prey community size, the effect of predation on prey diversity depends on the habitat size and intensity of competition among prey. A particular stabilizing predator may be beneficial for diversity at large spatial scales yet may clobber diversity at small scales. We show that the impact of frequency-dependent killing of prey can be understood through statistical properties of fluctuating prey communities: by suppressing fluctuations in relative abundance, stabilizing forces reduce negative covariance that arises from direct competition and thereby reduce the risk of rarity of prey. High attack rates relative to community size can reduce the mean abundance more than they increase the covariance

in population size; tension between increases in covariance and decreases in the mean abundance of prey populations determines the observed effect of predation on prey diversity. These results change our thinking about competition, stabilizing forces, and coexistence in real study systems where populations are finite, moving us beyond systems of differential equations towards a broader, trophic island-biogeographic perspective that considers the stochastic fluctuations of communities in the presence of stabilizing interactions. This trophic island biogeographic perspective and the scale-dependence of coexistence for a given food web is especially relevant to modern conservation biology where the trophic downgrading [7] and habitat fragmentation [8] are simultaneously changing the stabilizing forces and the community sizes in natural systems.

The third and final chapter makes a solid epistemological stride by using time-series data for hypothesis-testing a particular model of competitive dynamics - the Wright Fisher Process [11, 29], also known to ecologists and population geneticists as Neutral Theory. The Wright-Fisher process has been used as a null model for neutral competition in populations of genes [20], communities of canopy trees [16], and even the market weights of companies in competition over finite investment capital [10, 26], and it's been argued that neutrality may be a common occurrence in natural systems as any large deviation from neutrality may likely result in quick competitive exclusion [17]. From Tajima's D [28] to Fay & Wu's H [9] to tests of species-abundance distributions [24], hypothesis-tests of neutral competition are used to uncover genes that have undergone selection, and similar tests in ecology and economics may reveal niche structure within trophic guilds or economic sectors. Existing hypothesis tests of neutral competition use static snapshots of the frequencies of the different competing species, alleles or companies; the third chapter of my thesis presents tools for the hypothesis testing of neutral competition with time-series datasets. The utility of the tools developed is illustrated first by simulation and then by analyzing data from a time-series dataset of the human microbiome [3] and the day-end prices for 451 companies in the S&P 500 from 2000 to 2005. These tests of real ecological and economic time-series suggest that no one model can account for the stochastic dynamics of real complex-adaptive systems: some communities are difficult to distinguish from a Wright-Fisher process, whereas others such as one male's tongue microbiome and the S&P 500 market weights, exhibit markedly non-neutral dynamics. Focusing on the example of the male's tongue microbiota, I illustrate how the tools provided here for time-series analysis facilitate the construction of suitable alternative models for the competitive dynamics of these non-neutral systems.

Taken together, these chapters span a wide spectrum of competitive dynamics - resource competition in a patchy environment, apparent competition of prey with a common generalist predator, and neutral competition in more general competitive systems. At the heart of all of these studies is a limiting resource - food for *Dictyostelium*; food, space, intraguild aggression and a shared, generalist predator for prey; or strict zero-sum competition for space or investment capital - and these studies have examined various consequences for the diversity, dynamics, and empirical studies of the organisms in competition.

Dictyostelium diversity can be stabilized by spatial and temporal variability in resource pulses. Ungulates in Yellowstone can be stabilized by wolves, but whether or not wolves increase the long-term persistence of ungulates depends crucially on the size of habitat - and it's possible that even Yellowstone may not be big enough for wolves, but larger habitats or well-connected blocks of habitat may be required. The failure of introduced anoles on the Bahamas to maintain arthropod diversity could be due to the small size of the islands to which they were introduced [27]. The S&P 500 and the bacteria on the tongue of one particular male are in competition for limiting resources, but evolution has not led to neutrality among competitors in these systems over the timescales investigated; instead, there may be signatures of trust-busting policies or incumbent effects or rare-species advantages in these systems. Understanding the non-neutral competition in these systems has implications ranging from economic policy making to portfolio management and and probiotic pharmacokinetics.

Looking back, these projects are motivated by my interest in the assembly of competitive systems over evolutionary timescales - how do slime molds evolve under the tension between solitary life and cooperative dispersal? How are food webs assembled - does evolution stabilize or destabilize food webs, and, if so, how effective are these stabilizing forces for communities of different size? Do species competitively exclude one another until they are effectively neutral, as Hubbell hypothesized, or are general classes of non-neutral competition underlying patterns of biogeography (species-abundance distributions, species-area relationships, the distribution of times to most-recent common ancestors of species in a region, etc.)?

Looking forward, this thesis is my first step in a long journey to understand how competition structures natural systems, how our understanding of evolution in competitive systems can be utilized to improve the competitiveness of, say, a probiotic designed to establish a stable population and deliver lactase to the host, and how we conceptualize evolution in this world with Russian dolls of self-replicating things, with Dictyostelium straddling the major evolutionary transition between single- and multi-cellularity, with prey species in competition in a trophic guild (subjected to frequency-dependent/apostatic selection from a predator), and with humans - walking titans of 10 trillion cells harboring 100 trillion bacterial cells in their large intestines, each cell containing its own lysogenic viruses or transposable elements that self-replicate - all within a bustling (and currently growing but eventually saturating and limiting) world of human groups (companies, cultures, countries) and symbionts (crops, livestock, dogs) all struggling over limiting resources for survival and replication. Some way, some how, out of this struggle for existence over finite, unpredictable resources, we see not just many species of plankton or tropical trees, but a fractal symphony of symphonies of competing, replicating things. Provided the broad impacts of these ideas gives them some selective advantage in the community of human ideas (in competition over limited publication space and professorships and grants), or provided that I am fortunate in their neutral drift with equally fit competitors, these three chapters will be just the first of many studies on competition and coexistence in an unpredictable world.

Acknowledgments

“The enumeration in the Constitution, of certain rights, shall not be construed to deny or disparage others retained by the people.”

-9th Amendment, U.S. Constitution.

The enumeration in these Acknowledgments of certain people, shall not be construed to deny or disparage gratitude for other, unenumerated people.

I acknowledge support from an NSF Graduate Research Fellowship.

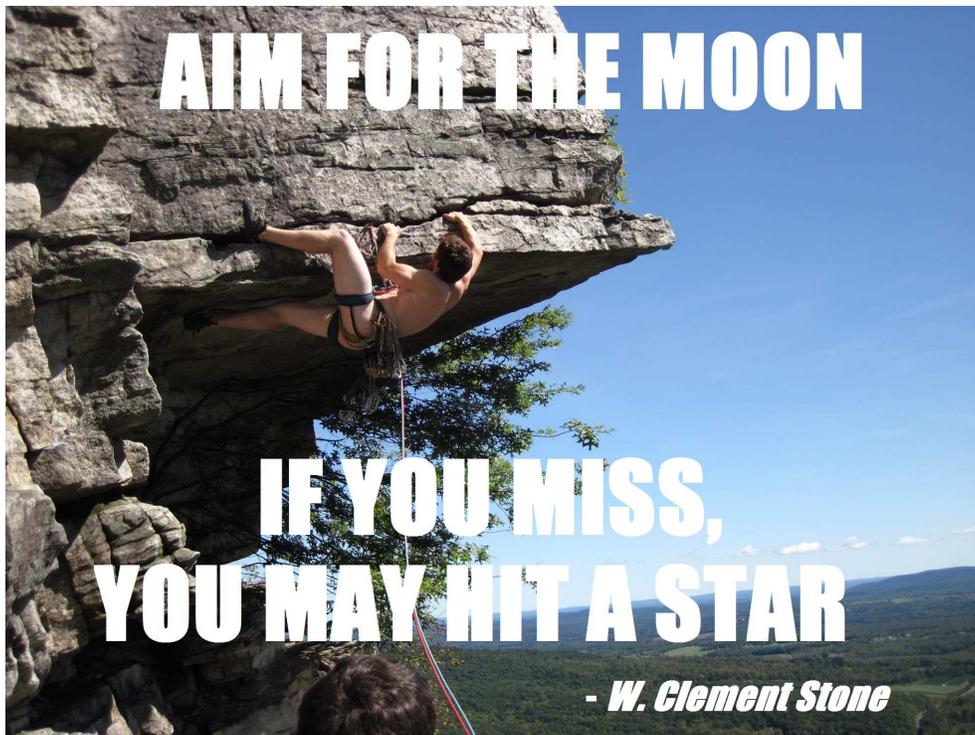
Dr. Alison E. Gammie was my Ms. Miyagi - this thesis would not have been possible without her mentoring and both professional and personal support. Likewise, Prof. Henry Horn was my Mr. Miyagi, and a role model for how to succeed as an academic generalist. My adviser, Prof. Simon Levin cultivated an excellent community of collaborators and gave me full freedom to pursue authentic academic work by saying, “Do what you want to do, I’ll defend your work.” Those words gave me a personal responsibility for my productivity and cultivated a stable, profound love of the science I’m doing. Prof. Ned Wingreen was an invaluable member of my committee. Professor Corina Tarnita and soon-to-be-Dr. Jacob B. Socolar are stellar collaborators and co-first-authors - the time spent working with them added years of fuel to the fire of my love of science.

The Princeton University Mountaineering Club deserves acknowledgment for introducing me to many essential collaborators, especially its founder Dr. Ben Safdi whose discussions about detecting dark matter while warming up on a V6 epitomized the academically productive culture of the club. Dan Lacker taught me everything I know about stochastic calculus during our drives to the Gunks and the New River Gorge. Emmanuel Shertzer helped out with parameter estimation for the Wright-Fisher Process. Jono Squire kept my sights (and psyche) high, motivating the use of Mathematica during our ascents of Moonlight Buttress, Half Dome, and other classic climbs. Josh Burby’s work in Hamiltonian mechanics was essential for finding variance-stabilizing transformations for the Wright-Fisher Process. Thomas van Boeckel - also known as “Papa Thom”, “Brother Thom”, “Mr. Thom”, “Admiral Thom”, and “El CapiThom” - fell from the sky and joined the PUMC. We climbed together, drank together, explored the West together, and solved damn near all of life’s problems together. We started a group called the IAS - the Institute for Actual Science - that met every Sunday over bourbon to discuss where science will actually help solve the world’s problems, helping us carve out our beliefs on the role of science in society. I will be forever grateful for these exceptional collaborators and life-long friends met through the PUMC.

Finally, I acknowledge an infinite wellspring of support from my family. My mom, Dr. Margaret Werner-Washburne is a molecular biologist and life-coach - I wish everyone in grad school could have my mom as their mother, mentor, and best friend. My dad, Bruce Washburne, taught me the importance of

saying “howdy” and being unconditionally nice to everyone, no matter what New Jersey says. My brother, Gabe Washburne, has been a life-long role model and a contagious source of enrichment and charisma. My Uncle Paul Sullivan taught me the importance of loving everything - “Love is everywhere you see it. You are all that you can be, so be it!” Last, but certainly not least, is my dog, Dr. Jack Washburne, who supported me every step of the way and taught me that all problems in life can be solved with a sufficiently long walk through the woods.

And finally, to future students who feel doubted and are told they can’t get a PhD with the amount of time they devote to living a good life outside of work, I acknowledge you, support you in your own authentic career, and encourage you to follow your heart (as my mom always told me) to find your own calling. I leave you with an inspirational picture from an adventure with colleagues/friends/partners-in-crime Jono Squire (photographer) and Dan Lacker (belaying in the foreground).



References

- [1] Priyanga Amarasekare. Competitive coexistence in spatially structured environments: a synthesis. *Ecology Letters*, 6(12):1109–1122, 2003.

- [2] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.
- [3] J Gregory Caporaso, Christian L Lauber, Elizabeth K Costello, Donna Berg-Lyons, Antonio Gonzalez, Jesse Stombaugh, Dan Knights, Pawel Gajer, Jacques Ravel, Noah Fierer, et al. Moving pictures of the human microbiome. *Genome Biol*, 12(5):R50, 2011.
- [4] Peter Chesson. Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, pages 343–366, 2000.
- [5] Peter L Chesson. Coexistence of competitors in spatially and temporally varying environments: a look at the combined effects of different sorts of variability. *Theoretical Population Biology*, 28(3):263–287, 1985.
- [6] Joseph H. Connell. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In P. J. den Boer and G. Gradwell, editors, *Dynamics of Populations*. Pudoc, Washington, 1971.
- [7] James A Estes, John Terborgh, Justin S Brashares, Mary E Power, Joel Berger, William J Bond, Stephen R Carpenter, Timothy E Essington, Robert D Holt, Jeremy BC Jackson, et al. Trophic downgrading of planet earth. *science*, 333(6040):301–306, 2011.
- [8] Lenore Fahrig. Effects of habitat fragmentation on biodiversity. *Annual review of ecology, evolution, and systematics*, pages 487–515, 2003.
- [9] Justin C Fay and Chung-I Wu. Hitchhiking under positive darwinian selection. *Genetics*, 155(3):1405–1413, 2000.
- [10] Robert Fernholz and Ioannis Karatzas. Relative arbitrage in volatility-stabilized markets. *Annals of Finance*, 1(2):149–177, 2005.
- [11] Ronald Aylmer Fisher. *The genetical theory of natural selection*. Clarendon, 1930.
- [12] George Francis Gause. *The struggle for existence*. Courier Corporation, 2003.
- [13] Georgii Frantsevich Gause. Experimental studies on the struggle for existence i. mixed population of two species of yeast. *Journal of experimental biology*, 9(4):389–402, 1932.
- [14] GF Gause and AA Witt. Behavior of mixed populations and the problem of natural selection. *American Naturalist*, pages 596–609, 1935.
- [15] Garrett Hardin et al. The competitive exclusion principle. *science*, 131(3409):1292–1297, 1960.

- [16] Stephen P Hubbell. *The unified neutral theory of biodiversity and biogeography (MPB-32)*, volume 32. Princeton University Press, 2001.
- [17] Stephen P Hubbell. Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional ecology*, 19(1):166–172, 2005.
- [18] G Evelyn Hutchinson. The paradox of the plankton. *American Naturalist*, pages 137–145, 1961.
- [19] Daniel H. Janzen. Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940):501–528, 1970.
- [20] Motoo Kimura. *The neutral theory of molecular evolution*. Cambridge University Press, 1985.
- [21] Simon A Levin and Robert T Paine. Disturbance, patch formation, and community structure. *Proceedings of the National Academy of Sciences*, 71(7):2744–2747, 1974.
- [22] Robert MacArthur and Richard Levins. Competition, habitat selection, and character displacement in a patchy environment. *Proceedings of the National Academy of Sciences of the United States of America*, 51(6):1207, 1964.
- [23] Robert H MacArthur. Population ecology of some warblers of northeastern coniferous forests. *Ecology*, 39(4):599–619, 1958.
- [24] Thomas J Matthews and Robert J Whittaker. Neutral theory and the species abundance distribution: recent developments and prospects for unifying niche and neutral perspectives. *Ecology and evolution*, 4(11):2263–2277, 2014.
- [25] Robert T Paine and Simon A Levin. Intertidal landscapes: disturbance and the dynamics of pattern. *Ecological monographs*, 51(2):145–178, 1981.
- [26] Soumik Pal. Analysis of market weights under volatility-stabilized market models. *The Annals of Applied Probability*, 21(3):1180–1213, 2011.
- [27] David A Spiller and Thomas W Schoener. An experimental study of the effect of lizards on web-spider communities. *Ecological Monographs*, pages 58–77, 1988.
- [28] Fumio Tajima. Statistical method for testing the neutral mutation hypothesis by dna polymorphism. *Genetics*, 123(3):585–595, 1989.
- [29] Sewall Wright. Evolution in mendelian populations. *Genetics*, 16(2):97, 1931.

Chapter 1: Spores and non-aggregating cells can explain coexistence of genotypes in *Dictyostelium discoideum*

Corina E. Tarnita^{a,1}, Alex D. Washburne^{b,1}, Allyson E. Sgro^c, Simon A. Levin^a

a - Department of Ecology and Evolutionary Biology

b - Lewis-Sigler Institute for Integrative Genomics

c - Joseph Henry Laboratories of Physics

1 - Joint first authors

Keywords: slime molds, cooperation, altruism, dispersal, genetic diversity, multiple habitats, coexistence, chimeras, variable environments.

An expanded version was published as:

Tarnita, C. E., Washburne, A., Martinez-Garcia, R., Sgro, A. E., & Levin, S. A. (2015). Fitness tradeoffs between spores and nonaggregating cells can explain the coexistence of diverse genotypes in cellular slime molds. Proceedings of the National Academy of Sciences, 201424242.

Abstract

Dictyostelium discoideum is a well-studied amoeba with a complex life cycle that includes both a single-cellular and a multicellular stage. To achieve the multicellular phase, individual amoebae aggregate upon starvation to form a fruiting body made of stalk and spores. The amoebae that contribute to stalk formation undergo apoptosis while those that become spores survive and are able to reproduce, a behavior that has been described as altruism. When amoebae aggregate they do not discriminate kin from non-kin, which sometimes leads to chimeric fruiting bodies. Within chimeras, a linear hierarchy of winning and losing genotypes has been documented, which suggests that there should eventually be only one surviving genotype. This contradicts however the great diversity of *D. discoideum* genotypes found in nature. Here we suggest that a little-studied component of *D. discoideum* fitness – the loner cells that stay behind and do not participate in the aggregate – could be selected for and achieve different abundances in environments with different food recovery properties, and that this could account for the apparent contradiction. We show that if at least two environments exist that are sufficiently different in their food recovery characteristics, and if in addition they are connected via weak-to-moderate dispersal, then coexistence of genotypes can occur. We further argue that the consideration of loners as a component of *D. discoideum* fitness makes it hard to define altruistic behavior, winners or losers, without a clear description of the ecology encountered by each genotype.

Significance Statement

Dictyostelium discoideum is an amoeba whose life cycle includes both single-cellular and multicellular stages, the latter achieved when individual amoebae aggregate upon starvation. Because amoebae do not discriminate kin from non-kin, chimeric aggregates can arise. Within chimeras, there is asymmetric contribution to reproduction and, based on this, the co-existing genotypes can be arrayed in a linear dominance hierarchy. This implies that only one genotype can survive, contradicting the great diversity of strains found in nature. We suggest that cells that fail to aggregate provide an additional fitness component that, in a rich ecology with multiple environments connected via dispersal, resolves the contradiction. The study of non-aggregators further sheds light on the discussion of altruism in *D. discoideum*.

Introduction

Dictyostelium discoideum is a social amoeba (or cellular slime mold) native to the upper layers of soil and leaf litter, where it feeds on bacteria (1-5). Upon exhausting their local supply of food, the solitary starving amoebae initiate a developmental program, joining with nearby amoebae to form an aggregate that

allows for facile transport to an area with more bacteria. Under some environmental conditions, the aggregate initially forms a slug that can migrate toward attractants such as light and heat that indicate a better future environment; but the culmination of development, whether or not a slug is formed, is ultimately a fruiting body. In this fruiting body, roughly 20% of the cells sacrifice themselves, vacuolating and forming a cellulose wall that provides a sturdy stalk to hold the remaining dormant spore cells off the ground so that they can more easily be dispersed to a new area. Since the stalk cells die for the benefit of the spore cells, this process is often discussed in terms of altruism and cooperation (reviewed in (5)). In nature, there is significant diversity and coexistence of multiple strains of *D. discoideum* (6). Moreover, chimeras (aggregates consisting of at least two genotypes) occur naturally (6-8), which implies that *D. discoideum* amoebae do not discriminate perfectly in the process of aggregation. These chimeras are functional and viable in the sense that their aggregation results in a fruiting body in which the multiple genotypes participate both in stalk formation and in spore production. However, the genotypes are not necessarily equally represented in the stalk and spores of the chimera (7). Through the mixing of different genotypes together into chimeras in the laboratory, linear hierarchies of competitors have emerged (9,10), in which a stronger competitor is a genotype that is disproportionately represented in the spores. From this perspective, not all strains are equivalent, and hence chimeras are considered suboptimal assemblages (10). In the absence of any additional information on possible frequency-dependent processes that might maintain coexistence, the existence of a linear competitive hierarchy suggests that there should be only one winner. This raises a paradox – if there are clear winners within chimeras, why is there so much diversity and coexistence among strains in nature (9)? One possible resolution used to explain the absence of certain types of cheaters in nature relies on spatial structure leading locally to a high degree of relatedness that helps select against cheating genotypes (8, 11); however, this is only the case for cheaters that are very harmful to group productivity (including their own when clonal), and not for more ‘mild’ cheaters that do take advantage of others in chimeras but are still able to act cooperatively to a certain extent, especially when clonal (8). The genotypes studied (10) fit better with the latter category since they all contribute in some amount to both spores and stalk, and therefore the high genetic relatedness mediated by spatial structure is not an explanation (in this straightforward form, without additional information about possible frequency dependent interactions between strains) for the observed coexistence.

Another explanation for coexistence that was suggested but unexplored both for *D. discoideum* (9) and for other cellular slime molds (12) is that strains that are at a disadvantage in chimeras have an advantage at a different stage in their life cycle, i.e. that there is a tradeoff between sporulation efficiency and other fitness-related traits. To date, however, analyses of genotypical fitness have focused on spore contribution as the sole fitness indicator. During the social phase however, not all cells aggregate; some cells stay behind. We will refer to them as non-aggregators or loners. Loner cells have been generally ignored so

far; because *D. discoideum* do not make microcysts, it was assumed that non-aggregating cells would simply die, which would be “an unlikely strategy” (9). But recent results show that loner cells are viable, meaning that they can eat and divide if food is replenished in the environment, and that therefore the loners can be an important component of *D. discoideum* fitness (13). However, the authors also found that epigenetic differences may lead to differing numbers of aggregating and non-aggregating cells in a clonal population (13). We therefore performed experiments to confirm that there were no longer-term effects of starvation or an epigenetic effect that prevented these non-aggregating “loner” cells from aggregating in the future under starvation conditions. To test if such an effect exists, we used cells left behind during the aggregation of starving *D. discoideum* from a naturally isolated, clonal population that were allowed to form fruiting bodies on non-nutrient agar (see SI for experimental details). The fruiting bodies were then removed and fresh bacteria were added to the food source facilitating the non-aggregating cells to re-grow and deplete the bacteria as expected from (13) and re-aggregate and go on to form normal fruiting bodies, leaving behind a population of non-aggregating cells themselves (see SI Fig A.1). Therefore, we confirm that the loners can be an important component of *D. discoideum* fitness. Loner cells can act as a form of exploitation strategy: certain environments may become advantageous quickly and, unlike spores that take time to germinate, loner cells can begin to eat and divide instantaneously. Thus, leaving behind some loner cells can give a genotype a head start in their home environment. Then, environments where food replenishes faster (henceforth fast-recovery environments) will select for genotypes that are more likely to invest in loners, while environments where food replenishes slower (henceforth slow recovery environments) will select for genotypes that are more likely to invest in spores. Stochastic (unpredictable) environments potentially can select for mixed strategies (13). We propose that, furthermore, if instead of focusing on one environment we consider multiple environments connected via dispersal, the loners and spores can play additional roles. The loners – an exploitation strategy – can also be seen as fulfilling the role of local dispersal. By contrast, the spores – an insurance against prolonged starvation – also fulfill the role of global dispersal since they are more likely to be dispersed to other environments. The spores will thus either survive until food is reintroduced, or they will get dispersed to new environments, some of which can be food-rich. Since environments with different food-replenishment characteristics select for different investments in spores versus loners, we hypothesize that weak-to-moderate dispersal between fast-recovery and slow-recovery environments can allow for the coexistence of multiple genotypes. We explore the feasibility of this notion through a model of resource competition that considers both loners and spores as part of *D. discoideum* fitness, in one environment, and in two environments connected by various degrees of dispersal.

Model and results

In our model, a *D. discoideum* genotype is characterized by a scalar α , which represents the fraction of cells that aggregate; the remaining $1 - \alpha$ is the fraction of non-aggregating cells, or loners. Thus, a monoculture of genotype $\alpha = 0$ does not undergo an aggregation phase; a monoculture of $\alpha = 1$ only produces aggregates and leaves no loner cells behind. Intermediate α values represent a mixed strategy, where some cells aggregate and others do not. Out of the cells that aggregate, only a fraction will become spores (the remaining cells either commit suicide and contribute to the formation of stalk, or get shed from the slug during migration). Since cells that are shed during migration have also been proved to be viable (14), they may similarly contribute to *D. discoideum* fitness; however, the number of cells left in the trail is probably influenced by how far individual slugs migrate in different environments, which has not been quantified in the wild. Similarly, different genotypes have different investments in stalk versus spores (10); however, in this paper we are not concerned with the selective forces that shape the stalk-to-spore investment ratio. For these reasons and in order to simplify our analysis, we will assume that the cell loss during migration and the investment in the stalk are fixed and identical for all genotypes so that we can focus our analysis solely on the fraction of spores versus loner cells. The model we describe will depend on the ecology that influences the lifecycle of *D. discoideum*. For simplicity, we are not concerned with soil type, light or moisture and assume those to be the same across environments; the property of interest is the ability of food to replenish in a given environment, or in other words, the starvation times experienced in that environment. One environment (patch). In the first part of the model we assume that the world is comprised of a single environment, in which food replenishment can be either deterministic (certain) or stochastic (uncertain). This model is similar to (13) but it is more general in that it explicitly accounts for resource competition among different genotypes. The amoebae consume resources, reproduce freely and grow at a rate governed by Michaelis-Menten kinetics (15). The different genotypes compete indirectly through the existing resources, but we assume no other frequency-dependence between types (i.e. we assume that intrinsic parameters such as the growth and aggregation rates of different genotypes are independent of the composition of the population). Since there is little-to-no information in the literature about growth rates of different genotypes, we make the simplifying assumption that different genotypes grow according to the same saturating functional form $cR/(R_{1/2} + R)$ where R is the amount of available resources, c is the consumption rate and $R_{1/2}$ is the concentration of resources at which the reaction rate is at half its maximum. We further assume that amoebae die at rate μ .

We describe the dynamics of active amoebae of all genotypes α existing in the same environment as resources are depleted by the growing amoebae until they are no longer able to sustain growth, after which the amoebae enter a starvation phase. During this phase, a fraction α of cells of genotype α aggregate with the purpose of forming spores, while the remaining $1 - \alpha$ stay as

starving loners. The non-aggregating (loner) cells stop consuming resources, stop reproducing, and decay at rate δ until the next resource pulse. Of the aggregating cells, a fraction s become viable spores; we assume that spores are very resistant to environmental stress, but that they nevertheless incur a small decay rate δ ; therefore, we assume that $\delta < \mu$. When the starvation period is over, food is reintroduced in one resource pulse, R_0 . Then the surviving loner cells start consuming resources and reproducing immediately, while spores undergo a delay period τ , which is the time required to activate the metabolic machinery necessary for resource consumption. The longer the delay τ , the more cost will be incurred by spores in an environment where loner cells are already consuming the resource while the spores undergo the germination process. Therefore genotypes that can leave behind some loners can have a head start and be favored. If resources get depleted before the germination period is over, we assume that spores return to dormancy, without incurring any cost associated with the abortion of the germination process. In reality, in addition to the costly delay of germination, *D. discoideum* also experiences a costly delay of sporulation: after only 6 hours, just as individual amoeba are beginning to aggregate, they are irreversibly committed to continuing with sporulation for the remaining 18 hours of the process (16). Since we are trying to show that in certain environments loners can be selected for, an additional cost for the spores will only make the selection for loners stronger and reinforce our results. Therefore, for simplicity, we do not include in our model the additional cost due to the irreversibility of the sporulation process. The dynamic equations describing the spores and loners are presented in the Supplementary Information. In a single environment as described above, with competing *D. discoideum* genotypes and instantaneous resource pulses arriving at random times, we explore how the lengths of the starvation periods that a genotype experiences affect its investment in spores and loners. Our results agree with (13). In a deterministic environment (i.e. when the starvation times are always of the same length) there is selection for one of the pure strategies: we find a critical threshold starvation time T_{cr} such that for $T < T_{cr}$ the winning genotype is one that never produces any aggregates ($\alpha = 0$) while for $T > T_{cr}$ the winning genotype is one that always aggregates to produce spores ($\alpha = 1$). When the environment is stochastic such that successive starvation times are independent and exponentially distributed with rate $1/\lambda T$, we find that mixed strategies can be selected for: if on average the environment is a fast-recovery one (low λT), then the mixed strategy invests more in loners than in spores; indeed, for sufficiently low λT , only loners will survive. Conversely, if the environment is a slow-recovery one (high λT), then the mixed strategy invests more in spores (see Fig. 1A), and for sufficiently high λT only spores will persist.

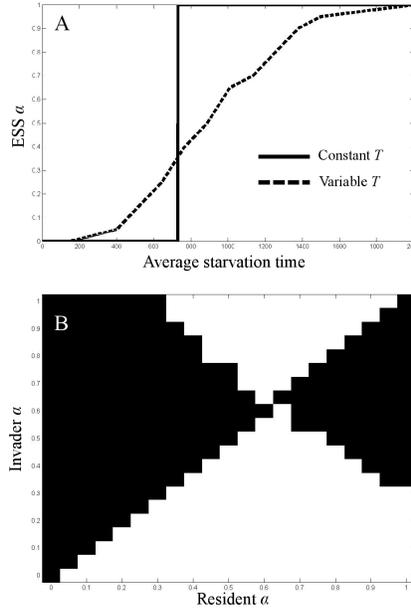


Figure 1: In one environment, deterministic (fixed) starvation times always select for pure strategies while stochastic starvation times can select for mixed strategies. A. The corresponding to the evolutionarily stable strategy (ESS) is shown as a function of average starvation time for both deterministic and stochastic cases. In the deterministic case, if T is under a threshold (calculated in SI) then the all-loners strategy wins; if T is above the threshold, then the all-spores strategy wins. In the stochastic case, extreme average starvation times select for the corresponding pure strategies, but intermediate average starvation times select for mixed strategies (intermediate). B. The pairwise invasibility plot (PIP) for the stochastic case with average starvation time $\lambda T = 1000$. The ESS strategy is $\alpha = 0.63$, which invests 63% in spores and 37% in loners.

The evolutionary stability (ESS) (17) of the mixed strategy is further confirmed by a pairwise invasibility analysis for a fixed stochastic environment with a given intermediate average starvation (see Fig. 1B). Here we assume an exponential distribution of starvation times; assuming uniformly distributed starvation times (13) leads to qualitatively similar results. In the future it will be interesting to explore other distributions (e.g. normal); however, we generally expect similar results to hold. We next study the effects of the model parameters on the evolutionarily stable genotype. The higher the consumption rate c , and implicitly the reproductive rate of solitary amoebae, the more likely it is to select for loners in the intermediate starvation regions; this is intuitive since the more that loners can take advantage of new food, the more advantage they will have over the spores. Similarly, the longer the spore germination delay τ , the bigger the cost incurred by spores, which favors genotypes with more loners. Loners will also be more favored the higher the death rate of spores, d . Conversely, the higher the spore success rate s or the death rate of single cells, d , the better it is for genotypes with more spores (Fig 2A-E).

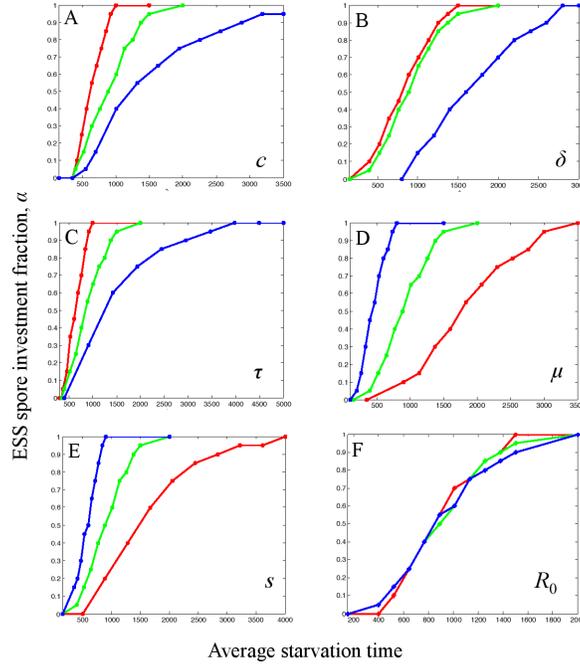


Figure 2: The sensitivity of the model to the parameters. A, B, C, F. Higher consumption rate (c), spore death rate (δ), time needed for spore germination (τ) or resource input favor selection for loners. D, E. Conversely, higher loner decay rate (μ) or higher spore survival and viable germination rate (s) favor selection for spores. In all panels, the lowest value of the parameter is red, the intermediate is green and the highest value is blue. Except for the parameter varied to perform the sensitivity analysis, all other parameters are as in Table S1. Parameter values for the sensitivity analysis: $c = 0.086, 0.173, 0.346$; $\delta = 0.00004, 0.0002, 0.001$; $\tau = 2, 4, 8$; $\mu = 0.001, 0.002, 0.004$; $s = 0.25, 0.5, 0.75$; $R_0 = 104, 108, 1012$.

Finally, we treat the size of the resource pulse, R_0 , separately. If the resource pulse is fixed, then the higher it is, the easier it is to select for loners. However, while varying R_0 has a sigmoidal effect on T_{cr} for constant T (SI Fig B.6), the benefits of varying R_0 in a stochastic environment are marginal and the resulting ESS α varies little with varying resource pulse size (Fig 2F). This is because T is exponentially distributed and the effects of R_0 are sigmoidal, with only a small intermediate region in which the selection for spores or loners is sensitive to varying R_0 . Finally, a more realistic scenario is that the resource pulse is stochastic. Our simulations suggest that introducing stochasticity in the resource pulse size does not select for mixed investment in loners and spores: if the starvation time T is fixed and the only source of stochasticity comes from the resource pulse, it appears that the only evolutionarily stable strategies are the pure strategies (see SI Fig B.7). However, because this is only a simulation result, it is possible that for different parameter combinations resource stochasticity will result in mixed investment. The relationship between the resource

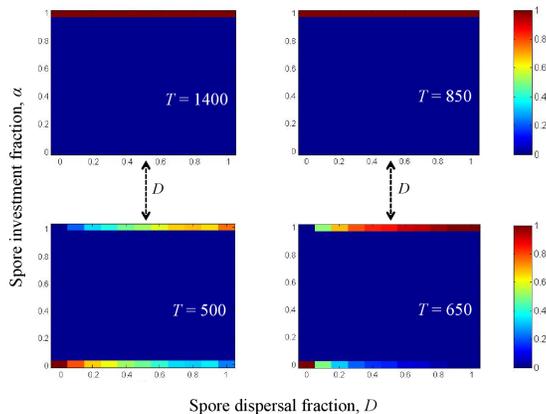


Figure 3: In two environments, if food recovery is deterministic and the environments are sufficiently different, coexistence between the two pure strategies is possible in the faster environment. The more different the environments, the larger the dispersal range that allows coexistence – $T = 1400$, $T = 500$ allow coexistence for the entire dispersal range $0 < D < 1$ while $T = 850$, $T = 650$ allow coexistence only for $D < 0.5$. The color scale, ranging from 0 (blue) to 1 (red), represents the average relative abundances across 60 replicates at the end of 1500 growth/starvation cycles.

pulse and the time to starvation is an interesting one and needs to be further explored, for different distributions of resources and stochastic times; however, since our preliminary analysis suggests that the more interesting and rich behavior seems to be induced by the starvation times, in this paper we choose to perform the entire analysis for a fixed resource input following every starvation period.

Two environments (patches). Including the loners in the analysis already produces interesting behavior (selection of mixed strategies) in one environment. In this second part however we explore whether the extension of our model to a spatially heterogeneous environment connected via spore dispersal can, under certain conditions, favor the coexistence of genotypes, pure or mixed. For simplicity, we consider exactly two environments characterized by different starvation times. Each environment is governed by the same dynamics as above, but they receive and exhaust resources independently of each other (asynchronously); the only element that couples the dynamics in the two environments is the dispersal of spores. When starvation occurs in one environment, a fraction $1-D$ of the spores remain in their home environment while a fraction D are moved to the other environment. Because the dynamics in the two environments are desynchronized, when spores get moved to a new environment they may immediately find food and start the germination process, or they may be lying dormant until food gets introduced into that environment. We consider two scenarios, depending on whether the starvation times in the two environments are deterministic or stochastic. If the two environments are deterministic, in the sense that environment i (for $i = 1,2$) always has a starvation time of length T_i ,

then in the absence of dispersal they each select for a pure strategy. Because we are interested in environments with different starvation times, we study fast recovery environments, which select for all loners ($\alpha = 0$), paired with slow recovery environments, which select for all spores ($\alpha = 1$). When dispersal is allowed between the two environments, coexistence of the two pure strategies can occur for weak dispersal, such that each strategy will dominate in the environment in which it is selected but, because of dispersal, spores will be present in both environments. As the dispersal fraction increases, however, two alternative outcomes can occur depending on how different the starvation times in the two environments are: (i) If the two environments are very different, coexistence in the fast environment can be maintained for a large range of dispersal values. This is because the very fast recovery of food allows the loners to grow quickly and dominate their home environment. (ii) The more similar the environments however, only weaker dispersal can maintain coexistence. For stronger dispersal the spores can easily start to dominate the loners in the fast environment and eventually replace them (Fig 3). If the two environments are stochastic, the behavior is richer. As in the deterministic case, each environment will have its winner in the absence of dispersal; and the more different the two environments, the more different the respective winning genotypes. When dispersal connects the two environments, for low-to-medium values of D , both genotypes coexist in both environments, although, as in the deterministic case, each environment is dominated by the genotype for which it selects. Due to stochasticity, in the presence of new mutations, each of the two coexisting genotypes is in fact likely to be a genotype surrounded by a cloud of its very close relatives (similar to a quasispecies, (18)); in that case, more than two genotypes might coexist due to dispersal and stochasticity in starvation times. Although stochasticity does allow for a richer coexistence than the one found in the deterministic case because mixed strategies coexist in both environments, it does so for a narrower range of dispersal fractions (weak-to-moderate dispersal). As dispersal increases, the two environments become increasingly more connected and new winning strategies evolve that are still dominating their home environments but are better able to deal with the other environment as well. When dispersal becomes high enough there is sufficient transfer between them for a new successful genotype to emerge that is selected for the average of the two environments, and coexistence is lost. As was the case in the deterministic scenario, the dispersal range for which coexistence is maintained depends on how different the two environments are, with more similar environments losing coexistence faster (Fig 4; see also SI Fig B.8 and B.9).

Discussion

We argue that recognizing that loners (non-aggregating cells) contribute to D . discoideum fitness and combining that with a richer ecology can explain the observed coexistence of diverse genotypes, thus resolving the paradox created by the linear competitive dominance in chimeras. The additional fitness com-

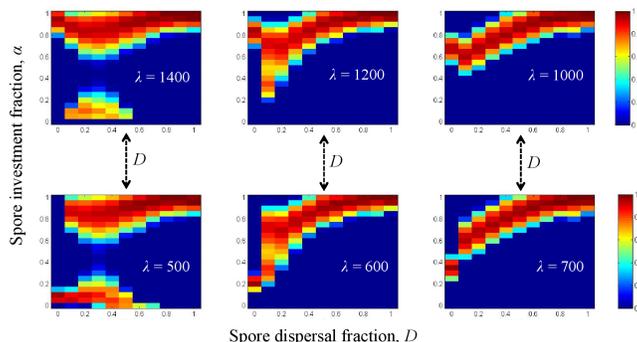


Figure 4: In two environments, if food recovery is stochastic and the environments are sufficiently different, coexistence between two (possibly) mixed strategies is possible in both environments. The more different the environments, the larger the dispersal range that allows coexistence: $T = 1400$, $T = 500$ allow coexistence for approximately $D < 0.6$; $T = 1200$, $T = 600$ allow coexistence only for approximately $D < 0.1$; $T = 1000$, $T = 700$ allow coexistence only for approximately $D < 0.01$. The color scale, ranging from 0 (blue) to 1 (red), represents the average relative abundances across 60 replicates at the end of 1,500 growth/starvation cycles.

ponent represented by the loner cells allows for mixed strategies in variable environments; this, coupled with the existence of at least two different environments connected via weak-to-moderate dispersal ensures the coexistence of at least two genotypes (Fig 5).

These ideas of spatially and/or temporally heterogeneous (variable) environments leading to risk-spreading or long-term optimization strategies including dormancy versus dispersal, persistence versus normal growth versus dormancy, and exploitation versus exploration are well established in ecology from studies of plants (19-25) to those of bacteria (26), planktonic copepods (27), and even social insects (28). Coupled with ideas of colonization-competition trade-offs (29-34), their effects on spatial coexistence of multiple phenotypes or even species has been well studied. However, to the best of our knowledge, they have never been suggested to explain coexistence in *D. discoideum*. Furthermore, in the context of cooperation and altruism, our loners have a strategy similar to the loner strategy described in the game-theoretic literature (35). There, loners are individuals who choose to opt out of the game or social contract. Because they cannot be exploited by defectors, loners persist in the population and allow the recovery of cooperators via rock-paper-scissors type dynamics. Due to their proven viability here and elsewhere (13) and by analogy with the existing theoretical and empirical literature, it therefore makes sense to explore the possibility that the loners contribute to *D. discoideum* fitness, and here we begin to show theoretically the rich behavior that such an expanded strategy space would display.

Much remains to be done empirically before the non-aggregating cells of *D. discoideum* are convincingly proven to be selected for, and before the mecha-

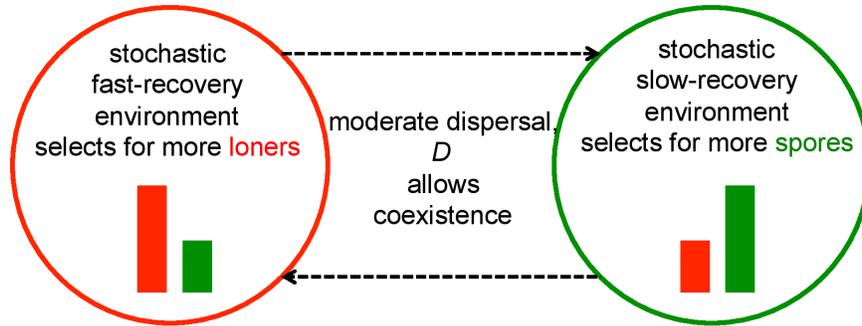


Figure 5: Fast-recovery environments select for investment in loners; slow-recovery environments select for investment in spores. Fast- and slow- recovery environments connected via weak-to-moderate dispersal allow for coexistence of strategies – each strategy dominates its home environment but dispersal allows for it to be present in the other environment as well.

nisms by which a genotype mediates the amount of loner cells left behind are understood. Because quorum sensing is instrumental in the decision to aggregate in *D. discoideum* (reviewed in (5)), one possibility is to search for mechanistic hypotheses there. We hypothesize two related mechanisms by which a *D. discoideum* genotype could lead to mixed investment in loners and spores when grown in a monoculture: (i) direct, signal-mediated quorum activation, and (ii) indirect, resource co-mediated quorum activation. The former posits the genotypes have varying and heritable sensitivities directly to the autoinducer; the latter posits that resource availability mediates a cell’s probability of responding to an autoinducer, and that the genotypes have varying resource starvation tolerances (hence resource co-mediated quorum activation). As resources are depleted, small-scale spatial heterogeneity can lead to some cells of the same genotype sensing abundant resources in their local environment and others sensing sparse resources. For a given level of resources and a given degree of spatial heterogeneity, a fraction of the cells may initiate their developmental program and move towards aggregation, and that fraction can vary across genotypes due to variation in the sensory and transcriptional machinery involved in detecting local resource density. Recent work (13) supports the plausibility of such a mechanism, but as loners still remain even in homogenous food conditions it is unresolved if this is the only mechanism at work in loner formation. We suggest that dose-response experiments examining different genotypes’ responsiveness to the autoinducer at various resource concentrations could test this hypothesis and assess the validity of our model.

It is furthermore important to note that the loners might not be the only additional component of the amoeba fitness. Since the stalk is considered to play an important role in dispersal, it is natural that stalk allocation would also

play an important role; similarly, since cells left behind in the trail of the slug have been shown to remain viable (14), they too possibly need to be included. But because slugs are not always formed and because they travel different distances depending on the environment (therefore shedding different amounts of cells), a more careful analysis and further experiments are necessary to determine exactly how to include these additional components into the *D. discoideum* fitness. Furthermore, our model assumes that different genotypes have identical growth rates both within the same environment and across different environments. Since empirical studies of growth rates are lacking, it would have been arbitrary to posit differential growth. We expect however that, provided more empirical evidence, a further extension of the model to include stalk- and trail-shedding allocations as well as possible frequency-dependent interactions among genotypes could produce an even richer set of predictions.

Here we have made a theoretical case for resolving the paradox of multiple genotype coexistence through considering non-aggregating cells as part of the *D. discoideum* strategy. We further suggest however that, if the loners are shown to be indeed selected for, other existing analyses of *D. discoideum* need to be revisited. For example, because stalk cells undergo apoptosis, *D. discoideum* has been used as a powerful model organism to explore the evolution and maintenance of altruism (reviewed in (5)). But whereas before the fitness of a genotype was well-defined as the number of spores it produced, because of the loners' contribution, fitness is a much more complicated quantity that strongly depends on the environments that genotype will encounter. The same is true for chimeras, where it is possible that a genotype that is better represented in the spores is less represented in the loners and vice versa, which makes the success of a genotype in a chimera dependent on the environments in which the chimeras exist. Including the loners in the analysis could also offer very different interpretations to the stalk/spore measurements performed to evaluate dominance in chimeras (9). Therefore, in this new context, a cheater, whether it be in a clonal or in a chimeric context, is much more challenging to define, motivating a comprehensive consideration of an organism's life history and ecological context when looking for problems of altruism. Finally, we focused our analysis on *D. discoideum* since it is the best studied of all cellular slime molds. However, similar diversity and coexistence has been identified in the majority of cellular slime molds (36) and both intra- and inter-specific chimeric fruiting bodies have been found to occur (12, 37, 38), albeit the interspecific ones are much rarer due to various interspecies barriers to mixing (37-39). As is the case in *D. discoideum*, within chimeras there are asymmetries in the contribution of the different strains to the reproductive tissue (12, 38, 40). Moreover, the social behavior of *D. discoideum* by no means exhausts the enormous range of developmental strategies that are seen in the slime molds (or even in the Dictyostelids). There are species in which a single amoeba can sporulate, species in which the stalk is an extracellular product and species in which cells die and produce a stalk continuously during the course of migration (2, 4, 41). We expect therefore that an extension of this analysis can be performed for other species, which will allow us to explore meaningful theoretical and empirical comparisons between

their different social behaviors as well as their abilities to coexist.

Acknowledgements

We thank J. Bonachela, J. Bonner, E. Cox, T. Gregor and P. Rainey for many useful discussions that helped shape the manuscript and C. Nizak and colleagues for generously sharing their work with us before publication. CET, ADW and SAL acknowledge support from the Templeton Foundation FQEB grant #RFP-12-14. AES acknowledges support from NIH through an NRSA postdoctoral fellowship. ADW acknowledges support from the NSF through a graduate fellowship, NSF DGE-1148900.

References

1. Bonner JT (1967) *The Cellular Slime Molds*, 2nd edn. Princeton Univ. Press, Princeton, NJ.
2. Bonner JT (1982) Evolutionary strategies and developmental constraints in the cellular slime molds. *Am Nat*, 119(4):530-552.
3. Bonner JT (2009) *The Social Amoebae*. Princeton University Press, Princeton, NJ.
4. Raper KB (1984) *The Dictyostelids*. Princeton: Princeton University Press.
5. Strassmann JE, Zhu Y, Queller DC (2011) Evolution of cooperation and control of cheating in a social microbe. *Proc Natl Acad Sci* 108:10855-10862.
6. Fortunato A, Strassmann JE, Santorelli LA, Queller DC (2003) Co-occurrence in nature of different clones of the social amoeba, *Dictyostelium discoideum*. *Mol Ecol* 12:1031-1038.
7. Strassmann JE, Zhu Y, Queller DC (2000). Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. *Nature* 408:965-967.
8. Gilbert OM, Foster KR, Mehdiabadi NJ, Strassmann JE, Queller DC (2007) High relatedness maintains multicellular cooperation in a social amoeba by controlling cheater mutants. *Proc Natl Acad Sci* 104(21):8913-8917.
9. Fortunato A, Queller DC, Strassmann JE (2003) A linear dominance hierarchy among clones in chimeras of the social amoeba *Dictyostelium discoideum*. *J Evol Biol* 16:438-445.
10. Buttery NJ, Rozen DE, Wolf JB, Thompson CRL (2009) Quantification of social behavior in *D. discoideum* reveals complex fixed and facultative strategies. *Curr Biol* 19(16):1373-1377.
11. Buttery NJ, Jack CN, Adu-Oppong B, Snyder KT, Thompson CRL, Queller DC, Strassmann JE (2012) Structured growth and genetic drift raise relatedness in the social amoeba *Dictyostelium discoideum*. *Biol Lett* 8(5):794-797.
12. Sathe S, Khetan N, Nanjundiah V (2013) Interspecies and intraspecies interactions in social amoebae. *J. Evol Biol* 27:349-362.
13. Dubravcic D, van Baalen M, Nizak C (2014) An evolutionarily significant unicellular strategy in response to starvation stress in *Dictyostelium social*

- amoebae [v1; ref status: awaiting peer review, <http://f1000r.es/3hg>] *F1000Research* 2014, 3:133.
14. Kuzdzal-Fick JJ, Foster KR, Queller DC, Strassmann JE (2007) Exploiting new terrain: an advantage to sociality in the slime mold *Dictyostelium discoideum*. *Behav Ecol* 18:433-437.
 15. Michaelis L, Menten ML (1913) Die Kinetik der Invertinwirkung. *Biochem Z* 49:333-369.
 16. Katoh M, Chen G, Roberge E, Shaulsky G, Kupsa A (2007) Developmental Commitment in *Dictyostelium discoideum*. *Eukaryot Cell* 6(11):2038-2045.
 17. Maynard Smith J, Price GR (1973). The logic of animal conflict. *Nature* 246:15-18.
 18. Eigen M, Schuster P (1979) *The Hypercycle: A Principle of Natural Self-Organization*. Berlin: Springer-Verlag.
 19. Cohen D (1966) Optimizing reproduction in a randomly varying environment. *J theor Biol* 12: 119-129.
 20. Hamilton WD, May RM (1977) Dispersal in stable habitats. *Nature* 269:578-581.
 21. Levin SA, Cohen D, Hastings A (1984) Dispersal strategies in patchy environments. *Theor Pop Biol* 26:165-191.
 22. Cohen D, Levin SA (1987) The interaction between dispersal and dormancy strategies in varying and heterogeneous environments. *Mathematical Topics in Population Biology, Morphogenesis and Neurosciences*, eds E. Teramoto and M. Yamaguti (Springer-Verlag, Heidelberg), pp. 110-122.
 23. Campbell BD, Grime JP, Mackey JML (1991) A trade-off between scale and precision in resource foraging. *Oecologia* 87:532-538.
 24. Cohen D, Levin SA (1991) Dispersal in patchy environments: the effects of temporal and spatial structure. *Theor Pop Biol* 39:63-99.
 25. Ludwig D, Levin SA (1991) Evolutionary stability of plant communities and the maintenance of multiple dispersal types. *Theor Pop Biol* 40:285-307.
 26. Kussell E, Kishony R, Balaban NQ, Leibler S (2005) Bacterial persistence: a model of survival in changing environments. *Genetics* 169(4):1807-1814.
 27. Hairston NG, Olds EJ (1984) Population differences in the timing of diapause: adaptation in a spatially heterogeneous environment. *Oecologia* 61(1):42-48.
 28. Hopper KR (1999) Risk-spreading and bet-hedging in insect population biology. *Annual Review of Entomology* 44:535-560.
 29. Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America* 15:237-240.
 30. Levin SA (1974) Dispersion and population interactions. *Am Nat* 108:207-228.
 31. Slatkin M (1974) Competition and regional coexistence. *Ecology* 55:128-134.
 32. Connell JH (1978) Diversity in tropical rain forests and coral reefs. *Science* 199:1302-1310.
 33. Tilman D (1994) Competition and biodiversity in spatially structured habitats. *Ecology* 75(1):2-16.

34. Pacala SW, Rees M (1998) Models suggesting field experiments to test two hypotheses explaining successional diversity. *Am Nat* 152:729–737. 35. Hauert C, De Monte S, Hofbauer J, Sigmund K (2002) Volunteering as a Red Queen mechanism for cooperation. *Science* 296:1129–1132.

36. Sathe S, Kaushik S, Lalremruata A, Aggarwal RK, Cavender JC, Nanjundiah V (2010) Genetic heterogeneity in wild isolates of cellular slime mold social groups. *Microb Ecol* 60:137–148. 37. Bonner JT, Adams MS (1958) Cell mixtures of different species and strains of cellular slime moulds. *J Embryol Exp Morphol* 6:346–356. 38. Kaushik S, Katoch B, Nanjundiah V (2006) Social behaviour in genetically heterogeneous groups of *Dictyostelium giganteum*. *Behav Ecol Sociobiol* 59:521–530. 39. Raper KB, Thom C (1941) Interspecific mixtures in the Dictyosteliaceae. *Am J Bot* 28:69–78.

40. Khare A, Santorelli LA, Strassmann JE, Queller DC, Kuspa A, Shaulsky G (2009) Cheater-resistance is not futile. *Nature* 461:980–982. 41. Atzmony D, Zahavi A, Nanjundiah V (1997) Altruistic behaviour in *Dictyostelium discoideum* explained on the basis of individual selection. *Curr Sci* 72(2):142–145.

SUPPLEMENTAL INFORMATION

1 Experimental confirmation of loner viability

Clonal, natural strains NC34.1, NC105.1, and NC98.1 of *Dictyostelium discoideum*, originally from Little Butt’s Gap, North Carolina, USA (Francis and Eisenberg 1993) were obtained from dictyBase (Fey, et al. 2013) and maintained on *Klebsiella aerogenes* lawns grown on SM agar plates. For growth of amoebae, spores of each strain were inoculated in SorMC buffer (15 mM KH_2PO_4 , 2 mM Na_2HPO_4 , 50 μM MgCl_2 , and 50 μM CaCl_2) supplemented with *Klebsiella* to an OD_{600} of 8 and shaken at 190 r.p.m. For starvation experiments, vegetative cells were harvested from these shaking cultures, washed, and resuspended at $1\text{--}2 \times 10^7$ cells/mL in developmental buffer (DB; 10 nM K/Na₂ phosphate buffer, 2 mM MgSO_4 , 0.2 M CaCl_2 , pH 6.5). 1–2 μL of this cell suspension was placed on a non-nutrient agar plate and allowed to aggregate. To test the viability of cells left behind after aggregation, spores were removed using an inoculation loop and 5 μL of *Klebsiella* at an OD_{600} of 8 in SorMC were added to the remaining cells.

We used cells left behind during the aggregation of starving *D. discoideum* (Fig. 3A) that were allowed to form fruiting bodies on non-nutrient agar (Fig. 3B,C). When these fruiting bodies are removed (Fig. 3D) and fresh bacteria are added as a food source (Fig. 3E), the non-aggregating cells re-grow and deplete the bacteria as expected from Dubravcic et al (2014) and re-aggregate and go on to form normal fruiting bodies, leaving behind a population of non-aggregating

cells themselves (Fig. 3F).

2 Analytical results for one environment

The dynamics of free-living amoebae is described by:

$$\begin{aligned}\frac{dR}{dt} &= -\frac{cR}{R_{1/2} + R} \sum_{\alpha} X_{\alpha} \\ \frac{dX_{\alpha}}{dt} &= \frac{cR}{R_{1/2} + R} X_{\alpha} - \mu X_{\alpha}\end{aligned}\tag{1}$$

for all genotypes α . Here R is the available resource and X_{α} is the abundance of amoebae of genotype α . Amoebae die at rate μ and reproduce and grow following Michaelis-Menten kinetics (Michaelis & Menten 1913), where c is chosen such that $c > \mu$. When resources are no longer able to sustain the growth of amoebae (i.e. when $dX_{\alpha}/dt = 0$ and $R^* = \mu R_{1/2}/(c - \mu)$), these enter a starvation phase. We let $t_{X_0; R_0}^*$ be the time to starvation when the initial total amoeba population is X_0 and the initial resource availability is R_0 . For simplicity of exposition we will henceforth often use t^* (unless otherwise needed for disambiguation) but we will implicitly assume its dependence on initial conditions.

We proceed to calculate the time to starvation t^* , the total abundance of amoebae at time t , $X(t) = \sum_{\alpha} X_{\alpha}(t)$ and the abundance of amoebae of genotype α at time t , $X_{\alpha}(t)$. From equations (1) we obtain the change in amoeba abundance X as a function of resources R :

$$dX/dR = -1 + \frac{\mu}{c} + \frac{\mu R_{1/2}}{cR}\tag{2}$$

which allows us to find X as a function of R , and of the initial conditions:

$$X(R; X_0; R_0) = \left(-1 + \frac{\mu}{c}\right)R + \frac{\mu R_{1/2}}{c} \log R + \text{const.}(X_0, R_0)\tag{3}$$

Here the constant term is determined from the initial conditions. Plugging (3) into the first equation in (1) we find

$$\begin{aligned}\frac{dR}{dt} &= -\frac{cR}{R_{1/2} + R} X \\ &= -\frac{cR}{R_{1/2} + R} \left(\left(-1 + \frac{\mu}{c}\right)R + \frac{\mu R_{1/2}}{c} \log R + \text{const.}(X_0, R_0) \right) =: -\frac{1}{f_1(R; X_0; R_0)}\end{aligned}\tag{4}$$

where the last equality simply indicates notation. Here f_1 is monotonic and positive on the interval of interest. Then we obtain an expression for time as a function of resources:

$$t(R; X_0; R_0) = \int_{R_0}^R (-f_1(y)) dy = \int_R^{R_0} f_1(y) dy =: f_2(R; X_0; R_0)\tag{5}$$

where the last equality is again notation. From this equation, which we solve via numerical integration, we obtain two key quantities. First, we find the time to starvation t^* simply as

$$t^*(X_0; R_0) = \int_{R_{X_0, R_0}^*}^{R_0} f_1(y) dy \quad (6)$$

where R_{X_0, R_0}^* is the equilibrium level of resources in equation (1) with initial conditions X_0, R_0 . Agreement between our formula (6) and simulations can be seen in Figure S3.

Second, by finding the inverse of function f_2 , which exists since f_1 and hence f_2 are strictly monotonic functions of R , we obtain R as a function of t :

$$R(t; X_0; R_0) = f_2^{-1}(t(R; X_0; R_0)) \quad (7)$$

Agreement between our analytical result (7) and simulations is shown in Figure S4(a). From (4) we know that $X = (R_{1/2} + R)/(cRf_1(R; X_0; R_0))$, so substituting $R(t)$ we find the abundance X as a function of time:

$$X(t; X_0; R_0) = \frac{R_{1/2} + f_2^{-1}(t)}{cf_2^{-1}(t)} \frac{1}{f_1(f_2^{-1}(t))} \quad (8)$$

Agreement between our result (8) and simulations is shown in Figure S4(b). Because we assume that amoebae of different genotypes have identical reproductive and death rates, we can write the growth derived from one cell during time t , starting with initial population size X_0 and resources R_0 as:

$$G(t; X_0; R_0) = \frac{X(t; X_0; R_0)}{X_0} \quad (9)$$

Agreement between our result (7) and simulations is shown in Figure S5. Then the abundance of amoebae of genotype α at time t is given by $X_\alpha(t; X_0; R_0) = X_{\alpha,0}G(t; X_0; R_0)$, where $X_{\alpha,0}$ is the initial abundance of genotype α .

We can now proceed to analyze the fate of a genotype when successive periods of food and starvation occur. In what follows, to simplify our analysis, we assume that, after a starvation period, the same amount of initial resources is introduced. Moreover, in order to simplify our notation we will use t^* (but implicitly assume that it depends on the initial resources as well as the initial population size) to denote the time to starvation after all amoebae are active. If the population has a non-zero number of spores, then from the moment resources are introduced it will take time τ for all amoebae to be active. It is at that point (after time τ) that we start to measure the starvation time t^* (see Fig 2). Because spores and loners have different fates, we keep track of each independently; furthermore, we keep track of the number and length of starvation periods that a genotype has experienced. We consider a food period followed by a starvation period as one event and we let $S_{\alpha,k}$ and $L_{\alpha,k}$ denote the abundance of spores, respectively of loners of genotype α after the k th food phase (i.e. at the beginning of the starvation phase of event k , see Fig 6).

We distinguish two cases:

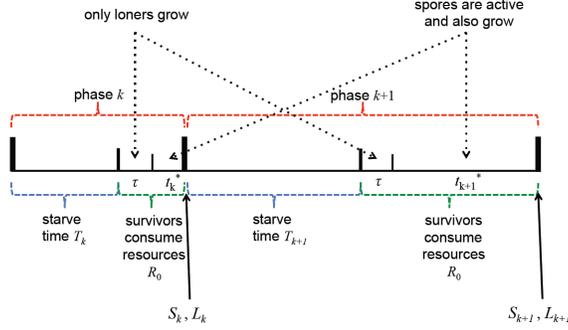


Figure 6: Events k and $k + 1$ with their food (growth) and starvation phases, showing where the spores and loners are measured in our analytical description above.

- (i) the loners finish the available food before the spores have had a chance to germinate; this can only be the case when there are at least some genotypes in the population that invest in loners. In this case we can write

$$\begin{aligned} S_{\alpha,k+1} &= e^{-\delta(T_{k+1}+t_{k+1}^*)} S_{\alpha,k} + \alpha e^{-\mu T_{k+1}} L_{\alpha,k} G(t_{k+1}^*) \\ L_{\alpha,k+1} &= (1 - \alpha) e^{-\mu T_{k+1}} L_{\alpha,k} G(t_{k+1}^*) \end{aligned} \quad (10)$$

where for simplicity of notation we used $G(t_{k+1}^*) = G(t_{k+1}^*; e^{-\mu T_{k+1}} L_k; R_0)$, which is the growth of a cell before the resources are depleted, given that the initial number, $e^{-\mu T_{k+1}} L_k$, of cells is given by the number of surviving loners after starvation time T_k , and that the initial resource input is R_0 . Notice that for the growth function we do not use an index α . This is because the initial conditions after starvation event k might contain loners of many different genotypes. Thus, when we say L_k we mean all loners, of all possible genotypes, after phase k . Note that, as mentioned above, t^* also depends on the initial conditions – however, to simplify the notation, since the initial conditions are the same as those in the argument of G , we simply use t^* .

- (ii) the spores can complete their germination, in which case we can write:

$$\begin{aligned} S_{\alpha,k+1} &= \alpha \left(e^{-\delta(T_{k+1}+\tau)} s S_{\alpha,k} + e^{-\mu T_{k+1}} L_{\alpha,k} G(\tau) \right) G(t_{k+1}^*) \\ L_{\alpha,k+1} &= (1 - \alpha) \left(e^{-\delta(T_{k+1}+\tau)} s S_{\alpha,k} + e^{-\mu T_{k+1}} L_{\alpha,k} G(\tau) \right) G(t_{k+1}^*) \end{aligned} \quad (11)$$

where as before, for simplicity, we denote $G(\tau) = G(\tau; e^{-\mu T_{k+1}} L_k; R_0)$ to be the growth of a cell during time τ , given that the initial population is made of the surviving loners and the initial resources are R_0 . Similarly,

$G(t_{k+1}^*) = G(t_{k+1}^*; e^{-\delta(T_{k+1}+\tau)} s S_k + e^{-\mu T_{k+1}} L_k G(\tau); R(\tau; e^{-\mu T_{k+1}} L_k; R_0))$ is the growth of a cell in the time before resources are depleted, given that the new initial population size is given by the active spores that have survived starvation and successfully completed germination and the loners which have grown for time τ ; the amount of resources available is that left from the initial R_0 , after the loners have consumed food during time τ . Here, as well, the growth is determined by all genotypes in the population; hence the growth term does not depend on the genotype α . Note that, as mentioned above, t^* also depends on the initial conditions – however, to simplify the notation, since the initial conditions are the same as those in the argument of G , we simply use t^* .

For different T_k this problem is hard (if not impossible) to solve analytically. However, when the environment is entirely deterministic (i.e. all starvation times have the same length) analytical results are possible. When all starvation periods have the same length T , our simulations show that the system converges to equilibrium values S_α^* , L_α^* and t^* . At this equilibrium, the system will either be such that we are in case (i) (loners finish the food before spores can germinate) or in case (ii) (spores can complete the germination process). If the former, then at equilibrium spores never get to germinate, but simply decay at rate δ ; so it is trivial that the winning genotype will be one that does not invest in spores at all, i.e. $\alpha = 0$. The more interesting scenario occurs when at equilibrium we are in case (ii), so that genotypes that invest in spores can potentially have a benefit. In this case, we need to explore what strategies can be present at equilibrium. For this, we perform an analysis to assess what strategies are evolutionarily stable. At equilibrium, a resident monoculture population of type α will satisfy the following:

$$\begin{aligned} S_\alpha^* &= \alpha \left(\overbrace{e^{-\delta(T+\tau)} s S_\alpha^*}^A + \overbrace{e^{-\mu T} G_\alpha(\tau) L_\alpha^*}^{B_\alpha} \right) G_\alpha(t^*) \\ L_\alpha^* &= (1 - \alpha) \left(e^{-\delta(T+\tau)} s S_\alpha^* + e^{-\mu T} G_\alpha(\tau) L_\alpha^* \right) G_\alpha(t^*) \end{aligned} \quad (12)$$

where the subscript α in $G_\alpha(\tau)$ and $G_\alpha(t^*)$ simply means that there is only one type in the population and where we denote $A = e^{-\delta(T+\tau)} s$ and $B_\alpha = e^{-\mu T} G_\alpha(\tau)$. The above equations imply that $S_\alpha^*/L_\alpha^* = \alpha/(1 - \alpha)$ for $\alpha \neq 1$. When $\alpha = 1$ then $L_\alpha^* = 0$, as expected. Then from (12) we obtain that the growth of a cell before starvation in an environment where only type α is present is given by

$$G_\alpha(t^*) = \frac{1}{\alpha A + (1 - \alpha) B_\alpha} \quad (13)$$

This holds for $\alpha \neq 1$; however, for $\alpha = 1$ we conclude directly from (12) that $S_1^* = A G_1(t^*) S_1^*$, which implies that $G_1(t^*) = 1/A$. Thus, (13) holds for all $\alpha \in [0, 1]$.

If we introduce a very small population of an invader β into a resident

population $\alpha \neq 1$, then the growth of β is described by the equations

$$\begin{aligned} S'_{\beta,\alpha} &= \beta \left(AS_{\beta,\alpha} + B_\alpha L_{\beta,\alpha} \right) G_\alpha(t^*) \\ L'_{\beta,\alpha} &= (1 - \beta) \left(AS_{\beta,\alpha} + B_\alpha L_{\beta,\alpha} \right) G_\alpha(t^*) \end{aligned} \quad (14)$$

The subscript α signifies, as above, that the resident population is of type α . Since the invader genotype is introduced at very low levels, its immediate growth occurs in the environment where growth is still determined by the resident genotype, such that $G_\alpha(t^*)$ is in fact given by (13). The only exception is when $\alpha = 1$, i.e. the resident is all-spores. Then any genotype $\beta \neq 1$ will have loners that will be able to grow during time τ as if they were alone in the environment; thus, in this case, the growth in time τ is in fact $G_\beta(\tau)$ and not $G_1(\tau)$ and the matrix becomes

$$\begin{aligned} S'_{\beta,1} &= \beta \left(AS_{\beta,1} + B_\beta L_{\beta,1} \right) G_1(t^*) \\ L'_{\beta,1} &= (1 - \beta) \left(AS_{\beta,1} + B_\beta L_{\beta,1} \right) G_1(t^*) \end{aligned} \quad (15)$$

Next we calculate in general the growth of genotype β in an α -monoculture; for $\alpha \neq 1$ this can be found from (14) to be $\lambda_{\beta,\alpha} = \beta A G_\alpha(t^*) + (1 - \beta) B_\alpha G_\alpha(t^*)$, which can be further written as:

$$\lambda_{\beta,\alpha} = \frac{\beta A + (1 - \beta) B_\alpha}{\alpha A + (1 - \alpha) B_\alpha} \quad (16)$$

for $\alpha \neq 1$, while for $\alpha = 1$ we obtain from (15)

$$\lambda_{\beta,1} = \frac{\beta A + (1 - \beta) B_\beta}{A} \quad (17)$$

We first explore when the pure strategies can invade or be invaded by other strategies.

[noitemsep]

- (a) 0 is not invadable by strategy $\alpha \neq 0$ if and only if $\lambda_{\alpha,0} < 1$, which is equivalent to $A < B_0$. Since this latter condition is independent of α , we conclude that if $A < B_0$, then 0 is not invadable by any strategy $\alpha \neq 0$. Conversely, if $A > B_0$, then 0 is invadable by all strategies.
- (b) 0 can invade strategy α if and only if $\lambda_{0,\alpha} > 1$, which is equivalent to $A < B_\alpha$.
- (c) 1 is not invadable by strategy $\alpha \neq 1$ if and only if $\lambda_{\alpha,1} < 1$, which is equivalent to $A > B_\alpha$.
- (d) 1 can invade $\alpha \neq 1$ if and only if $\lambda_{1,\alpha} > 1$, which is equivalent to $A > B_\alpha$. If $A = B_\alpha$ then 1 and α are neutral with respect to each other.

From these conditions we conclude that a strategy α can be invaded by either strategy 0 (if $A < B_\alpha$) or by strategy 1 (if $A > B_\alpha$). Therefore, *an intermediate strategy α can never be an ESS and the only possible ESSes are the pure strategies*. (In the threshold case $A = B_\alpha$, 1 and α are neutral, so α is again not an ESS). Next we will show that there can be at most one ESS.

From (a) above, we know that 0 is ESS if $A < B_0$; using (c) this also means that 0 invades 1, which means that 1 cannot be an ESS. Similarly, if 1 is an ESS, then from (c) we know that $A > B_\alpha$ for all $\alpha \neq 0$; this implies that $A > B_0$ as well, which means that 0 cannot be an ESS. Thus, *there can be at most one ESS for a given set of parameters*. Finally, we explore whether there can be no ESSes (i.e. whether neither of the pure strategies is an ESS). One possibility occurs when 0 and 1 are neutral with each other, i.e. $A = B_0$. In that case we find the critical starvation time threshold

$$T_{\text{cr}} = \frac{\log(G(\tau)) - \log s + \delta\tau}{\mu - \delta} \quad (18)$$

such that if $T < T_{\text{cr}}$ then 0 is an ESS and if $T > T_{\text{cr}}$ then 1 is an ESS (Fig. ??A).

Another possibility for there not to be any ESS is if $A \geq B_0$ (i.e. 0 is not an ESS) and $A < B_\alpha$ for some $\alpha \in (0, 1)$ (i.e. 1 is not an ESS). However, we conjecture and confirm via simulations for the parameters of interest in this paper that for biologically relevant parameter regimes either 0 or 1 will be an ESS, except when these are neutral to each other. This last case is given by (18).

From equation (18) it is also easy to see how the parameters of our model affect the threshold T_{cr} : the right hand side of the above equation is decreasing in s and μ and increasing in τ , δ and $G(\tau)$, the latter of which is an increasing function of the resource input R_0 , an increasing function of the consumption rate c and a decreasing function of $R_{1/2}$. Thus, we conclude that loners are favored for decreasing spore success rate s , decreasing loner mortality rate μ , and decreasing $R_{1/2}$ and for increasing time to germination τ and increasing spore mortality rate δ . Finally, we treat the size of the resource pulse, R_0 , separately. If the resource pulse is fixed, then the higher it is, the easier it is to select for loners. Varying R_0 has a sigmoidal effect on T_{cr} for deterministic T (Fig 2), suggesting that for low and high values of the resource pulse the benefit of increasing R_0 is only marginal. Thus, since T is exponentially distributed, the benefits of varying R_0 in a stochastic environment are marginal and the resulting ESS α varies little with varying resource pulse size.

Finally, a more realistic scenario is that the resource pulse is stochastic. Our simulations suggest that introducing stochasticity in the resource pulse size does not select for mixed investment in loners and spores: if the starvation time T is fixed and the only source of stochasticity comes from the resource pulse, our simulations find that the only evolutionarily stable strategies are the pure strategies. However, because this is only a simulation result, it is possible that for different parameter combinations resource stochasticity will result in

mixed investment. The relationship between the resource pulse and the time to starvation is an interesting one and needs to be further explored, for different distributions of resources and stochastic times; however, since our preliminary analysis suggests that the more interesting and rich behavior seems to be induced by the starvation times, in this paper we choose to perform the entire analysis for a fixed resource input following every starvation period. A full sensitivity analysis for stochastic T is shown in Fig. 2.

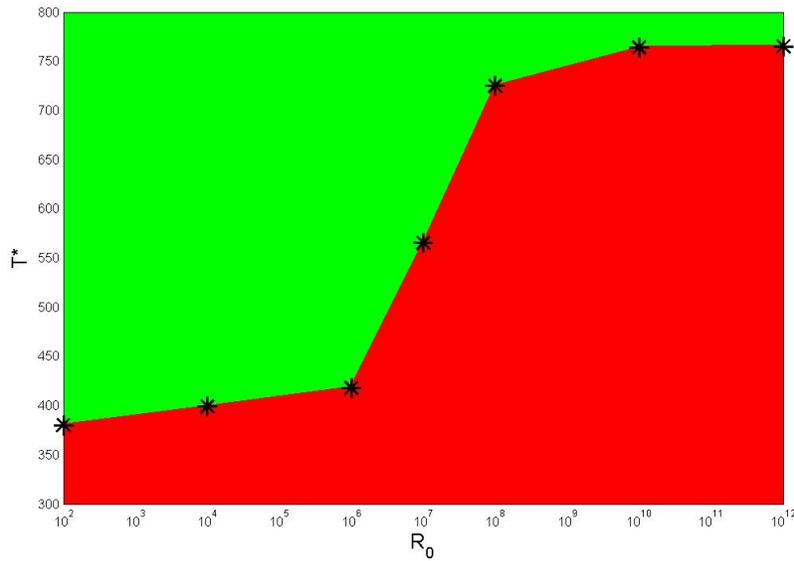


Figure 7: Varying R_0 has a sigmoidal effect on the threshold T_{cr} : low values of R_0 lead to low threshold values and thus favor spores; high values of R_0 lead to high threshold values and thus favor loners; in between, for a narrow range of intermediate values, there is a sudden jump. Green = all-spores; red = all-loners. All parameters are as in Table C.1.

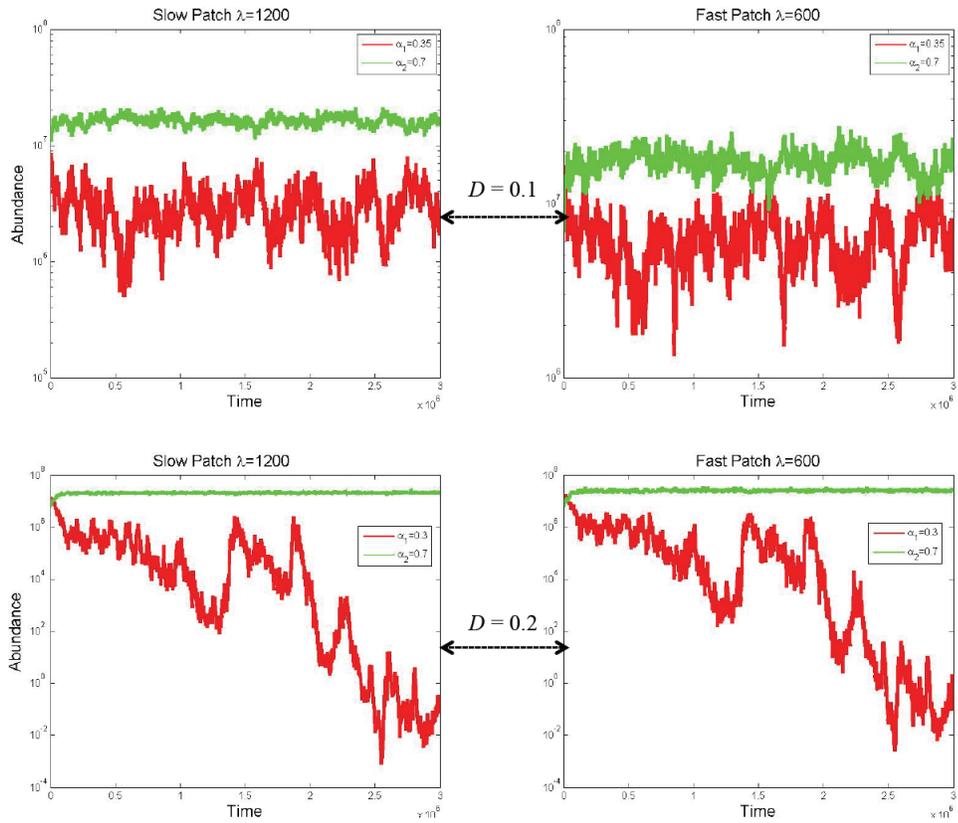


Figure 8: Long-term simulations show that the coexistence of strategies for more similar environments with stochastic starvation times is only possible for low values of dispersal. For $D = 0.1$ there is coexistence between strategies $\alpha = 0.35$ and $\alpha = 0.7$ while for $D = 0.2$ coexistence is lost. Green = strategies with higher spore investment; red = strategies with higher loner investment.

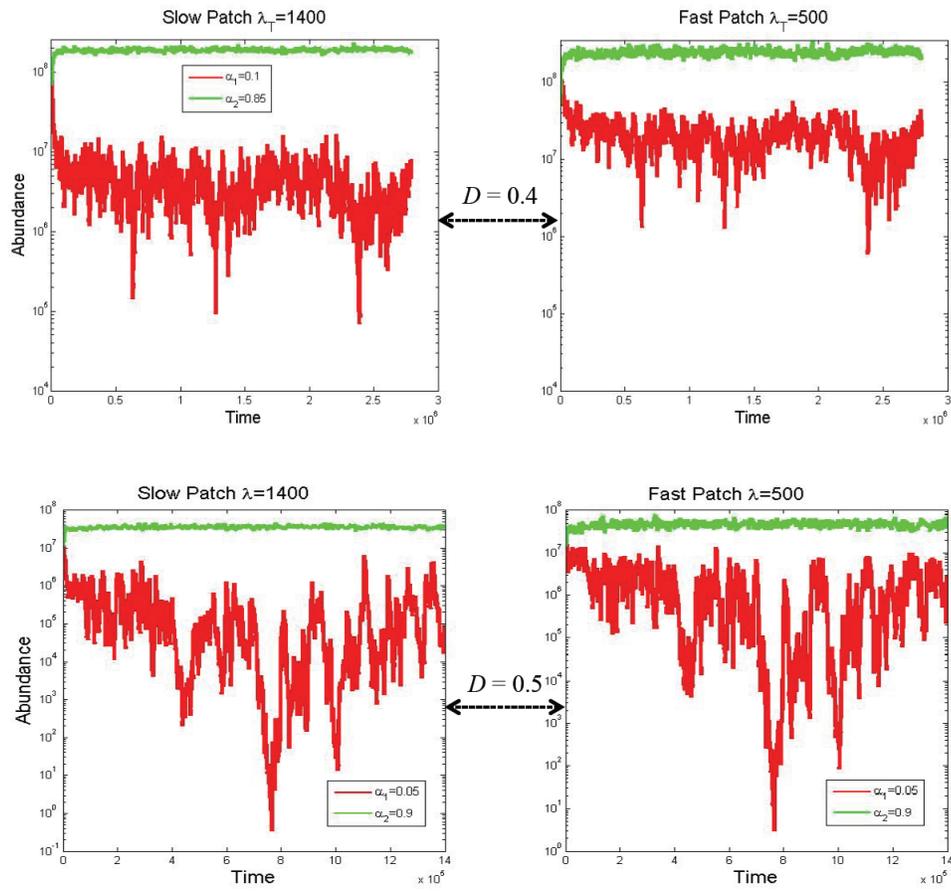


Figure 9: Long-term simulations show that the coexistence of strategies for less similar environments with stochastic starvation times is possible for low-to-medium values of dispersal. For both $D = 0.4$ and $D = 0.5$ there is coexistence between strategies ($\alpha = 0.1$ and $\alpha = 0.85$ for the former; $\alpha = 0.5$ and $\alpha = 0.9$ for the latter). For higher values of D coexistence is lost. Green = strategies with higher spore investment; red = strategies with higher loner investment.

3 Details of the simulations

All simulations were run using MATLAB R2013b. `ode15s` was used for numerical integration of B1, and the parallel computing toolbox was used to more rapidly sample replicates for trajectories, invasion analysis, and sensitivity analysis.

One environment (patch). Populations were initialized with 21 genotypes $\alpha_i = 0.05i$, $i = 0, \dots, 21$ whose initial abundances were independently drawn from a standard log-normal distribution and subsequently normalized so that the entire population contained 10^8 cells. The cells of the different genotypes were then split into spores with probability α_i and loners with probability $1 - \alpha_i$.

An initial resource pulse of magnitude 10^8 was added and the trajectories governed by equation (1) were numerically integrated, with spores remaining inactive until τ hours after the resource pulse. The simulation was stopped when the resources could no longer sustain a growing population, that is, when the resources crossed R^* . If the resources were consumed by the active loners before the spores were able to activate, the spores remained dormant and unaffected (i.e. there is no cost to spores aborting germination). After the resources ran out, each subpopulation of genotype i split into spores with probability α_i and loners with probability $1 - \alpha_i$. Only a fraction s of spores were viable upon germination. Starvation times, T_k , were drawn from an exponential distribution with mean λ_T . T is the time between the end of growth (t^*) and the arrival of the following resource pulse. During the starvation time, spore populations decayed exponentially at rate δ while the loners decayed exponentially at rate $\mu > \delta$. At the end of the starvation time, a new resource pulse of size 10^8 arrived.

Invasion Analysis. Residents were established at an initial abundance of 10^8 and invaders with initial abundance equal to a half that of the resident. Single patch growth-starvation cycles were run for 1000 cycles and replicated 4 times. If the invader had, on average across replicates, a higher abundance than the resident at the end of the 1000th growth phase, the corresponding square in the pairwise invasibility plot (PIP) was colored black, otherwise the corresponding square was colored white. The diagonal entries, where residents and invaders were neutral, were not simulated but instead set to black as a default.

Sensitivity analysis. Parameters c , μ , δ , τ , s and R_0 were varied to study their effects. Our estimates of α_{ESS} and α^* , were obtained by running 20 replicate single-patch simulations for 1000 growth/starvation cycles each, and finding the α_i with the highest abundance, on average across replicates.

Multiple environments (patches). We ran simulations of two patches undergoing desynchronized growth-starvation cycles, with dispersal from one patch to the other happening during spore formation at the end of the dispersing patch's growth period. At the end of growth, a fraction D of the successfully formed spores moved to the neighboring patch while $1 - D$ stayed at the home patch after which newly formed spores were governed by the same delay assumptions as in the single patch. One patch was chosen as a \hat{O} slow-recovery \hat{O} patch in which starvation times were drawn from an exponential distribution with

Table of notation and parameter values

Parameter	Use	Value
c	consumption rate	0.173 (4-hour doubling time)
$R_{1/2}$	resource concentration at which the reaction rate is at half-maximum	10^7
R_0	resource input after starvation	10^8
s	fraction of aggregating cells that become viable spores	0.504
τ	time needed for spore germination	4 hours
μ	death rate of amoebae	0.002
δ	death rate of spores	0.0002
α, β	fraction of aggregating amoebae	variable
T	length of starvation	variable
λ_T	average length of starvation	variable

Table 1: Table of notations used in the text and the baseline parameter values used for simulations, unless otherwise specified in the figure legends.

high average time λ_T , and the other patch was chosen to be a fast-recovery patch with low λ_T . To explore the effect of the similarity between patches, we varied the difference between the high and low λ_T 's.

Sampling Trajectories. In both single patch and 2-patch simulations, trajectories were sub-sampled to reduce the memory costs of simulations and interpolated at the same timepoints to allow averaging across replicates. Sub-sampling of trajectories was performed by noting the time and population vector for a patch at the end of that patch's growth phase. For the two desynchronized patches, this means that the sampling times of each patch differ, as patch i was only sampled at the end of patch i 's growth period. To later obtain average trajectories such as those displayed in Fig C.1 and C.2, the sampling times and population vectors of each patch were interpolated at 1000 equally-spaced time points on the shortest time-interval of all the replicates being averaged.

1cm

References

Dubravic D, van Baalen M, Nizak C (2014). An evolutionarily significant unicellular strategy in response to starvation stress in *Dictyostelium* social amoebae [v1; ref status: awaiting peer review, <http://f1000r.es/3hg>] F1000Research 2014, 3:133 (doi: 10.12688/f1000research.4218.1)

Fey, P., Dodson, R., Basu, S., Chisholm, R. L. (2013). One Stop Shop for Everything *Dictyostelium*: dictyBase and the Dicty Stock Center. *Dictyostelium discoideum* Protocols. *Methods Mol. Biol.* 983:59-92, edited by Ludwig Eichinger

and Francisco Rivero.

Francis, D. and Eisenberg, R. (1993). Genetic structure of a natural population of *Dictyostelium discoideum*, a cellular slime mould. *Mol Ecol.* 2(6):385-391.

Michaelis, L, Menten, ML (1913). Die Kinetik der Invertinwirkung. *Biochem Z* 49: 333-369.

Chapter 2: Prey carrying capacity modulates the effect of predation on prey diversity

Jacob Socolar¹, Alex Washburne²

Author affiliations:

¹Department of Ecology and Evolutionary Biology, Princeton University.

²Program for Quantitative and Computational Biology, Princeton University.

Author emails: ¹jacob.socolar@gmail.com; ²awfour@princeton.edu

Both authors contributed equally to this work.

Keywords: Predator; Coexistence; Prey-switching; island biogeography; stochastic dynamics

An expanded version of this paper has been accepted at The American Naturalist.

Abstract

Understanding the role of predation in regulating prey diversity is a major goal in ecology, with profound consequences for community dynamics, ecosystem structure, and conservation practice. Deterministic differential equation models predict that some predation regimes, such as prey-switching predation, should promote prey coexistence and increase prey diversity. However, such models do not capture stochastic population fluctuations that are ubiquitous in empirical study sites and nature reserves. In this paper, we examine the effects of prey-switching predation on the species richness of prey communities with demographic noise. We show that in finite, discrete prey populations, the ability of prey-switching predation to promote diversity depends on the carrying capacity of the prey community and the richness of the source pool for prey. Identical predation regimes may have opposite effects on prey diversity depending on the size and productivity of the habitat or the metacommunity richness. Statistical properties of the fluctuations of prey populations determine the effect of stabilizing mechanisms on species richness. We discuss the implications of this result for empirical studies of predation in small study areas and for the management of small nature reserves.

Introduction

Predators influence numerous features of biological communities, including the diversity of their prey (Terborgh and Estes, 2010). Intuitively, predators may reduce prey diversity by eating their prey into oblivion (Schoener and Spiller, 1996; Blackburn et al., 2004), but they also can promote prey diversity by reducing interspecific competition (Paine, 1969; Chase et al., 2002). This diversity-promoting effect occurs when predators consume a potentially dominant prey species (Paine, 1966) or when specialist predators prevent any prey species from becoming abundant (Terborgh et al., 2001; Janzen, 1970; Connell, 1971).

Much of the literature on predator-mediated coexistence has focused on differential equation models (Chesson, 2000; Chase et al., 2002). In finite populations, however, demographic stochasticity may play a significant role in observed community dynamics, especially when species are rare (Turelli, 1980; Nisbet and Gurney, 1982; Lande et al., 2003). Prey populations in a single study area or nature reserve may be sufficiently small that stochastic processes influence important system properties, including species richness. A nuanced understanding of predator-mediated coexistence in finite populations requires a trophic theory of island biogeography that explores how predators influence inherently stochastic colonization/extinction dynamics and statistical patterns of richness in small prey communities (Holt, 2009).

Nascent trophic theories of island biogeography assume that finite predator and prey populations interact stochastically, but thus far have treated a limited subset of predator-prey interactions, namely where predators do not affect prey populations, an assumption termed *donor control* (Holt, 2009; Gravel et al., 2011). Contrary to the assumption of donor control, predators generally reduce the mean population sizes of their prey, thereby increasing the probability of stochastic extinction, all else being equal (Holt et al., 2008). Thus, when analyses of Lotka-Volterra style models predict that predation should promote diversity by increasing the relative fitness of an invader, the actual impact of the predator on a finite prey population is unclear. In very large prey populations, stabilizing predation regimes should reliably increase prey diversity. In sufficiently small prey populations, the elevated risk of stochastic extinction may outweigh the stabilization afforded by a particular predation regime.

Elucidating the effects of predation on the diversity of finite communities has ramifications for empirical studies of predation and for protected area management. Empirical studies of small populations may fail to detect a diversity-promoting effect of predators even when that role indeed exists in larger natural communities. On the other hand, a predator’s diversity-promoting effect may vanish in small populations, leading theory and large-scale empirical studies to make perverse management recommendations for small nature reserves.

In this paper, we describe and implement a family of stochastic simulations to elucidate the effects of stabilizing predation on the species richness of finite prey communities. We focus on prey-switching predators that change their behavior through time to consistently target the most abundant prey (Murdoch, 1969; Holt, 1984; Abrams and Matsuda, 1996, 2003, 2004). In differential equa-

tion models, such predators stabilize prey communities via negative frequency dependent selection, provided that the predator does not deterministically drive its prey extinct. This fact, coupled with the simplicity of a single prey-switching predator (compared to myriad specialist predators), makes prey-switching predation an attractive regime for investigating the tension between stabilizing ecological interactions and intensified stochasticity resulting from predation. Furthermore, empirical evidence documents prey-switching predators in nature: various predators form search images (Martin, 1988) or shift the timing (e.g. Elliott, 2004) and/or location (e.g. Murdoch, 1977) of foraging to better target abundant prey resources (Smout et al., 2010).

In finite prey populations with demographic variation in lifetime and fecundity, we find that for a given prey-switching predation regime, there is a critical prey carrying capacity above which predation increases mean prey richness relative to a zero-predation community, and below which predation depresses prey richness. This carrying capacity is determined by tension between the stabilizing effect of prey-switching predation and the destabilizing effect of reduced prey abundance. Increasing metacommunity species richness shrinks the region of parameter space where prey-switching predation increases diversity. These results can be understood in terms of the statistical properties of colonization/extinction dynamics in the presence and absence of predators.

Methods

We wish to investigate the dynamics of a community that is subject to both stabilizing predator-prey interactions and demographic stochasticity. To do so, we use a simulation model that we describe and motivate in three steps. First, we present a system of ordinary differential equations that captures the stabilizing predator-prey interaction in our model. Second, we build a stochastic process that corresponds to the deterministic system with added demographic noise. By *corresponds to*, we mean that over infinitesimal time windows the expected change in the stochastic community (conditioned on its current state) equals the deterministic change of the system of differential equations. Third, we supply the computational details of how we simulate the stochastic process (see supplemental material for computer code). We summarize the simulations that we present in this paper, including sensitivity analysis and modifications to the model that we use to evaluate the robustness of our results to relaxations of various model assumptions.

The deterministic system

Predator-Prey Model

Considerable work has been done in the analysis of prey-switching predators in deterministic, continuous, well-mixed systems (e.g. see Holt, 1984; Abrams and Matsuda, 1996, 2003, 2004). To marry our study with their work, we construct a differential system similar to theirs to govern the stabilizing dynamics in our

simulations. We conceptualize a predator-prey community on a patch or island, where population dynamics are the outcome of birth and death processes and a small immigration rate. We begin by writing general equations for a well-mixed, continuous-state predator-prey system with one predator and one prey.

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K} \right) + m - \rho P \frac{N}{K} \quad (1)$$

$$\frac{dP}{dt} = \alpha \rho P \frac{N}{K} - \delta P.$$

The prey population N grows logistically with an intrinsic growth rate r , a carrying capacity K , a small immigration rate m , and removal via a bilinear (Holling type I) predation term that depends on the predator population P and the predation rate ρ . Eventually, we explore other functional forms for predation to assess the robustness of our results.

The predator population increases as it harvests prey, but decreases via a constant death rate δ . The parameter α^{-1} represents the average number of prey that a predator must consume in order to replace itself. The model has a single non-trivial equilibrium at

$$\begin{aligned} \bar{N} &= \frac{\delta K}{\alpha \rho} \\ \bar{P} &= \frac{rK}{\rho} \left(1 - \frac{\delta}{\alpha \rho} \right) + m. \end{aligned} \quad (2)$$

The non-trivial equilibrium is stable when $\bar{P} > 0$ or, for $m \approx 0$, $\rho > \frac{\delta}{\alpha}$. Equilibrium prey populations decrease monotonically with ρ whereas \bar{P} is maximized when $\rho \approx 2\frac{\delta}{\alpha}$.

We obtain a simple multi-species version of the system by assuming neutral competition among prey:

$$\begin{aligned} \frac{dN_i}{dt} &= rN_i \left(1 - \frac{\sum N_j}{K} \right) + m - \rho P \frac{N_i}{K} \\ \frac{dP}{dt} &= \alpha \rho P \frac{\sum N_j}{K} - \delta P. \end{aligned} \quad (3)$$

These equations represent a mean-field neutral prey community subject to non-selective predation.

We assume neutrality in the prey community not because we believe prey communities to be neutral in their dynamics, but as a parsimonious starting point. A neutral prey community allows us to explore the role of prey-switching predators in maintaining prey diversity without being overwhelmed by the large space of particular assumptions about competitive interactions between prey species. We break the neutral assumption when we incorporate prey-switching, but the neutral prey community remains a benchmark against which we compare communities containing prey-switching predators.

Prey Switching

We now modify the predation term to incorporate prey-switching. We do so with a term that allows us to vary the strength of prey-switching independent of the total amount of prey consumed. Thus, we retain the overall predation rate $\rho P \frac{\sum N_j}{K}$, but we select which prey species is killed based on its relative abundance using a tunable parameter z that determines the strength of prey-switching. By definition, prey-switching predators preferentially target more abundant prey, so that the diet ratio of species i and j , $f_i(\mathbf{N})/f_j(\mathbf{N})$, increases more quickly than the ratio of those species' abundances, N_i/N_j . Many functional forms of the species-specific predation rate have been implemented for prey-switching predators (reviewed by van Leeuwen et al., 2013; Morozov and Petrovskii, 2013). For its algebraic simplicity, we use the form

$$f_i(\mathbf{N}) = \rho P \frac{\sum N_i}{K} \frac{(N_i)^z}{\sum_j (N_j)^z} \quad (4)$$

with $z \geq 1$. This yields identical diet ratios to the form used by Elton and Greenwood (1970), but the community-wide predation rate remains $\rho P \frac{\sum N_j}{K}$ regardless of the value of z . van Leeuwen et al. (2013) critique Elton & Greenwood's form and derive a more principled—but more complex—alternative based on an individual-based encounter process. For our purposes, a critical advantage of our phenomenological functional form is the ability to vary prey-switching independent of the overall predation rate. This property allows us to develop intuition about our stochastic system by decomposing a prey species' variance in abundance into the whole-community variance affected by ρ and the between-species covariance affected by z . We later demonstrate that our results are robust to incorporating the functional form of prey-switching predation covered by van Leeuwen et al. (see Robustness Checks, below).

The full community dynamics are thus given by

$$\begin{aligned} \frac{dN_j}{dt} &= rN_j \left(1 - \frac{\sum N_j}{K} \right) + m - \rho P \frac{\sum N_j}{K} \frac{(N_j)^z}{\sum (N_i)^z} \\ \frac{dP}{dt} &= \alpha \rho P \frac{\sum N_j}{K} - \delta P. \end{aligned} \quad (5)$$

This system of n species of prey and one predator has one predator-free fixed point where $\bar{P} = 0$ and $\bar{N}_{tot} = K$, and one predator-containing fixed point where $\bar{N}_{tot} = \frac{\delta K}{\alpha \rho}$, where by symmetry $\bar{N}_j = \frac{\delta K}{\alpha \rho n}$ for all prey species j and $\bar{P} = rK \left(1 - \frac{\bar{N}_{tot}}{K} \right)$. As for the 1-prey 1-predator system in equation 1, the predators can persist when $\rho > \frac{\delta}{\alpha}$.

The stochastic system

We construct a stochastic analog to equation 5 in a manner similar to Haegeman and Loreau (2011). Populations are composed of discrete individuals, and

populations increase or decrease by one individual as a result of four fundamental events: prey birth/migration, natural prey death, predation, and predator death. We assume that these events occur at exponentially distributed time intervals. We then construct a stochastic process by defining propensities of these four events to ensure that over short (infinitesimal) timescales the expected change in the stochastic system per unit time follows the trajectory defined by equation 5. We assume that prey competition results in a density-dependent birth process, while death is density-independent. Respectively, the propensities are:

$$\begin{aligned}
n_j^+ (\mathbf{N}, P) &= \max \left\{ 0, rN_j \left(1 - \frac{\sum N_j}{K} \right) + m + \gamma_j N_j \right\} \\
n_j^- (\mathbf{N}, P) &= \gamma_j N_j \\
\tilde{p}^+ (\mathbf{N}, P) &= \rho P \frac{\sum N_j}{K} \\
p^- (\mathbf{N}, P) &= \delta P
\end{aligned} \tag{6}$$

where n_j^+ is the propensity of birth/migration events of species j , n_j^- is the propensity of natural death events of prey species j , \tilde{p}^+ is the propensity of predation events community-wide, and p^- is the propensity of predator deaths. Each event has a specific impact on the community. Birth/migration in species j causes a unit increase in the population of species j (and consequently in $\sum N_j$), while a natural death causes a unit decrease. Predation causes a unit decrease in the prey community size $\sum N_j$, and the decrease occurs in species j with probability $\frac{(N_j)^z}{\sum (N_i)^z}$. To account for the imperfect trophic efficiency in our differential system (equation 5), each predation event has a probability $\alpha < 1$ of causing a unit increase in the predator population. Predator deaths cause a unit decrease in the predator population. To ensure that predators don't go extinct, we set a reflecting wall at $P = 1$, so the predators can never have fewer than one individual in the population.

The parameters γ_j allow the possibility that different prey species have different intrinsic per-capita death rates, and consequently different average lifetimes, $1/\gamma_j$. Allowing γ_j to vary by species enables us to change the generation times of prey independently without disrupting the mean-field dynamics.

To confirm that this stochastic model produces trajectories over short time intervals whose mean is given by equation 5, note that as long as $\sum N_j$ is not much larger than K (more precisely, as long as $rN_j \left(1 - \frac{\sum N_j}{K} \right) + m + \gamma_j N_j > 0$), the expected change in a prey species' abundance ΔN_j over a short time interval Δt is

$$\begin{aligned}
\mathbb{E} [\Delta N_j | \Delta t] &= n_j^+ - n_j^- \\
&= rN_j \left(1 - \frac{\sum N_j}{K} \right) + m - \rho P \frac{\sum N_j}{K} \frac{(N_j)^z}{\sum (N_i)^z},
\end{aligned} \tag{7}$$

and likewise the expected change in predator abundance in some small unit time is

$$E[\Delta P|\Delta t] = \alpha \rho P \frac{\sum N_j}{K} - \delta P. \quad (8)$$

Note further that the model rarely attains value of $\sum N_j$ large enough to cause the mean trajectories to deviate from the deterministic expectation. Birth rates drop to zero as the community approaches this value, and deaths are much more frequent than migrations.

One crucial difference between the stochastic model and its deterministic analog is the tendency for the stochastic model to display oscillatory dynamics. In our model, such oscillations result from the resonant amplification of demographic stochasticity (McKane and Newman, 2005); similar effects can also result from exogenous environmental stochasticity Nisbet and Gurney (1982). Below (see Robustness checks), we also consider a non-oscillatory system to confirm that our main results are not peculiarities related to quasicycles in stochastic Lotka-Volterra systems.

Robustness checks

We wish to ensure that the species richness patterns predicted by our model are not artifacts peculiar to the Lotka-Volterra system that we implement. Therefore, we also consider several related systems which serve as robustness checks for our general results.

Predator functional responses

To begin, we consider deviations from the Holling Type I functional response and the simple form of prey switching that we have discussed thus far. First, we replace the unbounded, linear predation term $f(\mathbf{N}) = \rho \frac{\sum N_j}{K}$ with a saturating Holling type II term $f(\mathbf{N}) = \frac{\rho \frac{\sum N_j}{K}}{1 + \rho h \frac{\sum N_j}{K}}$ where ρ is the attack rate and h is the handling time. Second, we consider a mechanistically justified functional form for prey-switching derived by van Leeuwen et al. (2013). In their model, the per-predator rate of consumption of prey species i in an n -species prey community is

$$f_i(\mathbf{N}) = \frac{\rho_i \tilde{N}_i \sum_{k=1}^n s_{ik} \rho_k \tilde{N}_k}{\sum_{k=1}^n \rho_k \tilde{N}_k \left(1 + \sum_{j=1}^n s_{kj} T_{kj} \rho_j \tilde{N}_j \right)} \quad (9)$$

where ρ_i is the attack rate of predators on species i , s_{ij} is the similarity of species i to species j , T_{ij} is the handling time (the average time required between catching species i to attacking species j) and $\tilde{N}_i = \frac{N_i}{K}$ is the density of species i (Figure S6). For our purposes, we set $T_{ij} = 0.5$; $s_{ij} = 0.5$ for $i \neq j$ and $s_{ij} = 1$ for $i = j$; and $\rho_i = \rho$ constant for all species i .

Invariant and independently fluctuating predators

In the presence of demographic stochasticity, Lotka-Volterra systems can exhibit strong oscillatory behavior as a result of resonant amplification of demographic stochasticity, and indeed our simulations exhibit strong oscillations (Figure S7). The resulting community-wide prey bottlenecks increase extinction rates, thereby reducing the likelihood that a prey-switching predator will increase diversity. Because oscillations are not ubiquitous in real-world predator-prey systems, we also examine a non-oscillatory system. To do so, we replace the predator from the Holling type II differential system with a constant predator population \bar{P} , equal to the equilibrium population from that system. This invariant predator population consumes prey with a Holling type II functional response.

These systems produce predator-prey dynamics with considerably reduced variance compared to the oscillatory cases. To increase the predator-induced variance without producing quasicycles, we construct analogous systems where the predator population is not constant, but instead varies according to a discrete Ornstein-Uhlenbeck process with propensities

$$\begin{aligned} p^+ &= \lambda\bar{P} + \gamma P \\ p^- &= (\gamma + \lambda)P \end{aligned} \tag{10}$$

where \bar{P} is the equilibrial predator population in the corresponding Holling type II system. As before, we set a reflecting wall for the predator population at reflecting wall at $P = 1$. The variation in predator abundance is independent of the prey abundance, and we refer to this system as an independently fluctuating predator system.

Species-specific generation times

By using a single intrinsic prey death rate across species, we might create a system that is unreasonably prone to instability and oscillation. We address this issue by examining prey species with different generation times (see Appendix I; Table S2). For these systems, we implement a Holling type I functional response in the predator. Note that the community dynamics are no longer consistently neutral, even without prey switching (see Appendix II).

Density-dependent death

In our model, density dependence exists in the prey because birth rates decline as prey increase. A more general model of density-dependence is

$$\begin{aligned} n_j^+ (\mathbf{N}, P) &= \max \left\{ 0, rN_j - \beta \frac{r}{K} N_j \sum_i N_i + m + \gamma_j N_j \right\} \\ n_j^- (\mathbf{N}, P) &= \gamma_j N_j + (1 - \beta) \frac{r}{K} N_j \sum_i N_i \end{aligned} \tag{11}$$

where $\beta = 1$ is our density-dependent birth model. As a robustness check, we consider a model of density-dependent death, $\beta = 0$. Populations with density-dependence entirely in death are more volatile and thus have higher long-term variances in population size and qualitatively different extinction-colonization dynamics following predator introduction (see Appendix II).

Metacommunity richness

To investigate the influence of the metacommunity richness M on the ability of predation to promote diversity, we simulate communities with Holling type I predation across the full range of values of ρ and K used in this paper for $M \in \{10, 20, 40\}$. Additionally, for $K = 1000$ and M ranging from 10 to 360, we determine the attack rate ρ^* above which predators suppress diversity below the neutral expectation (see Appendix II).

Simulations

We simulate predator-prey communities using MATLAB R2013b. MATLAB code for simulation is available in the supplementary materials, described in Appendix VI. We implement simulations using a Gillespie algorithm (Gillespie, 1977) in which timesteps are drawn from an exponential distribution whose rate constant, η , is the inverse of the total propensity for events of any kind

$$\eta = \left[(\tilde{p}^+ + p^-) + \sum_j (n_j^+ + n_j^-) \right]^{-1}. \quad (12)$$

At the end of each timestep, a single event is selected with probability proportional to the event's propensity. The time and state of the system are updated, and the process repeated for T time units.

Parameter choices and initial conditions

To ensure biological realism in our simulations, we turn to the empirical literature to inform our parameterizations. Throughout this paper, we fix the prey death rate $\gamma = 0.1$. This is a free choice that specifies the timescale, and the remainder of the parameters are chosen relative to γ . Table 1 (Holling type I model) and Table S1 (all models) give the numerical values and biological interpretations for all simulation parameters. We encourage the reader to visit Appendix I, where we outline parameter estimation for our model, drawing on empirical investigations of demographic parameters, trophic efficiencies, predator-prey biomass ratios, trophic cascades, and so forth. Similar measurements, if available, can be used to parameterize our model for any particular natural system. In each simulation, we initialize the community near its equilibrium in the corresponding parameterization of the mean-field model, thus ensuring that transient dynamics alone cannot account for differences between the stochastic system and the mean-field equilibrium.

Table 1: Parameters for Holling type I model

Parameter	units	Value(s)	Interpretation
δ	time^{-1}	0.1	Intrinsic prey deathrate; inverse prey lifespan
γ	time^{-1}	0.05	Intrinsic predator deathrate; inverse predator lifespan
r	time^{-1}	5	Maximum prey population growth rate
K	prey	200 – 2000	Carrying capacity; system size; inverse competition intensity
M	species	30; 10 – 360 (fig. 3)	Metacommunity richness
m	$\frac{\text{prey}}{\text{time}}$	0.3	Migration rate
α	$\frac{\text{predators}}{\text{prey}}$	10^{-3}	Ratio of predator births to predation events
ρ	$\frac{\text{prey}}{\text{predators}\cdot\text{time}}$	0 – 500*	Predator attack rate (prey deaths per predator)
z	–	1 – 2	Prey-switching exponent

Stationarity check

We are interested in the asymptotic or stationary species-richness of communities at island-biogeographic equilibrium. To simulate this quantity, we must determine the relaxation time required to ensure that the imprint of initial conditions fades and the system reaches dynamic equilibrium. For every simulated result in this paper, we present a stationarity check in the supplement. Each figure in the paper corresponds to an equivalently numbered supplementary figure in Appendix III that presents the final result alongside results obtained by stopping the simulation after shorter time intervals. These stationarity checks confirm that even as we vary functional forms and parameters, simulations approximate the stationary species richness by $T = 2000$.

Transient dynamics

Experimental manipulations of predation (Bangs and Fritts, 1996; Schoener and Spiller, 1996; Heinlein et al., 2010) result in transient dynamics. Experimenters might wish to know whether the dynamics observed following a predator introduction provide reliable information about the eventual stationary state. To explore this question, we simulate predator introductions into a predator-free community. We consider the introduction of one individual predator with ordinary population dynamics and a Holling type I functional response, and also of an invariant, equilibrium-sized predator population with a Holling type II functional response (mimicking studies that monitor an experimental predator introduction for less time than is typically required for predators to reproduce; e.g. Heinlein et al., 2010; Stier et al., 2014). We monitor the dynamics by examining the community state every 0.25 time units for 50 time units (5 prey generations). Across the full range of values examined for ρ and K , we simulate predator introductions into twenty independent replicate prey communities near stationarity (see Appendix II for details). We ask whether the initial change in species richness is in the same direction as the final change, and we examine patterns in the time evolution of richness towards stationarity.

Results

Species richness landscapes

As expected, predators that do not engage in prey-switching ($z = 1$) never increase prey diversity for any value of the prey carrying capacity K . For such predators, any increase in predation intensity leads to a decrease in the average prey species richness. Prey-switching predators, on the other hand, increase diversity at low attack rates and high values of K , and this effect becomes more pronounced as z increases (Figure 1). These results are all in line with the predictions of the mean-field model.

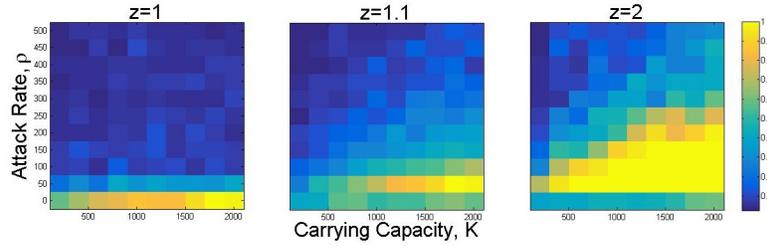


Figure 1: Local prey species richness relative to metacommunity richness as a function of the carrying capacity K , predator attack rate ρ , and prey-switching exponent z in a Holling type I predator-prey community. Prey-switching predators increase prey diversity only when attack rates are sufficiently low, and the range of diversity-promoting attack rates decreases as K gets small. All data points represent the averaged richness of 4 communities simulated for $T = 20000$ time units. See Figure S1 for stationarity checks.

However, for every value of z examined, sufficiently high predation rates lead to decreased diversity (Figure 1). For a particular value of K , the neutral community at $\rho = 0$ serves as a baseline for species-richness in absence of predators. There always exists an attack rate ρ^* above which even a prey-switching predator depresses diversity below that of the predator-free community. The value of ρ^* increases monotonically with K . Thus, the identical predation regime can promote diversity for large K and inhibit diversity for small K . The qualitative pattern in prey species-richness as a function of ρ and K is robust to several variations in our model assumptions, including a saturating (Holling II) predation rate, invariant and independently fluctuating predator populations, species-specific generation times for predators and prey, a different functional form for prey-switching introduced by van Leeuwen et al. (2013), and prey density dependence via death rates instead of birth rates (Figure 2).

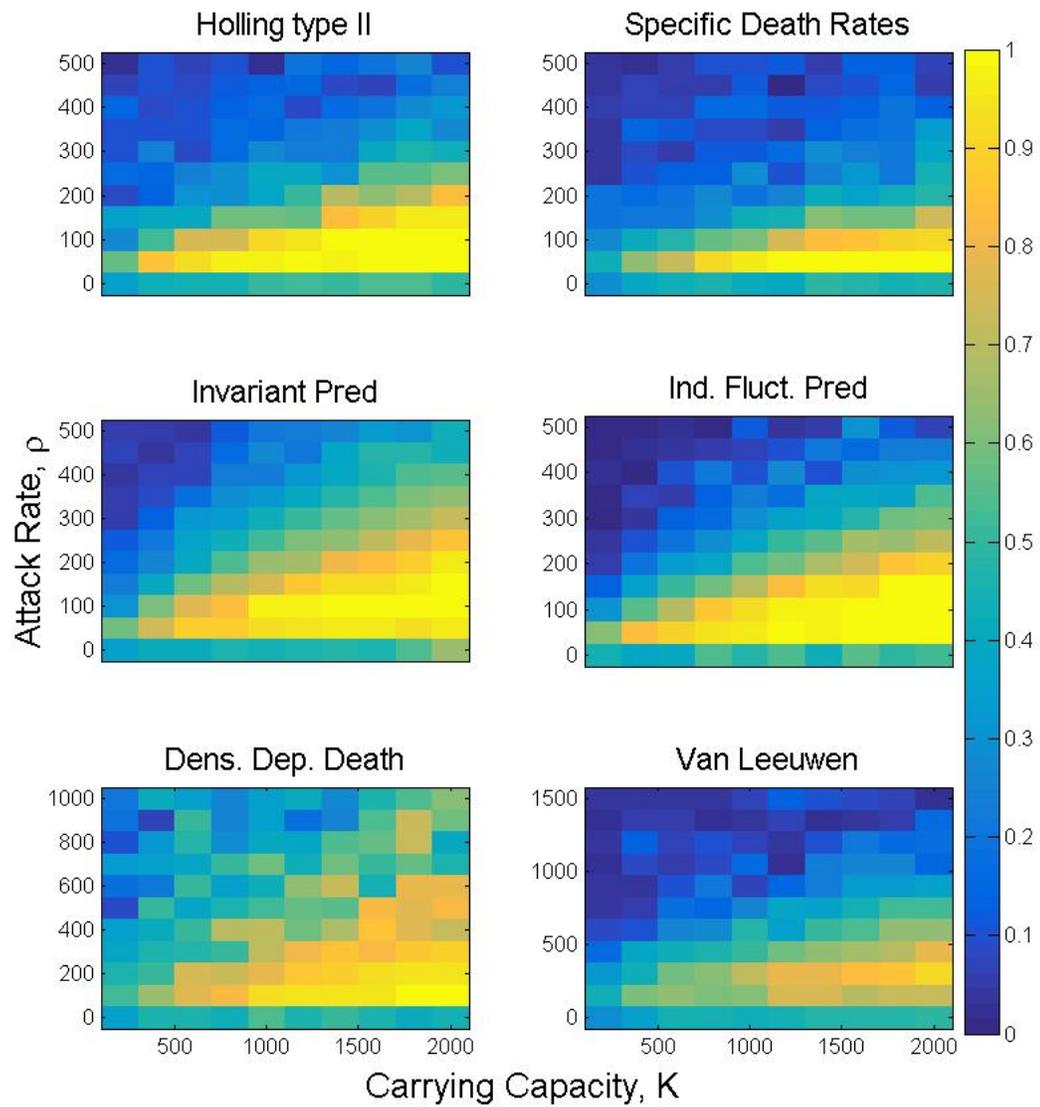


Figure 2: The qualitative patterns evident in Figure 1 are preserved across a variety of alternative models for predator-prey dynamics. All data points represent the relative richness of the local community to the metacommunity richness averaged over 4 communities simulated for $T = 2000$ time units. Note the different scale for predator attack rates ρ for the van Leeuwen and Density dependent death simulations (see appendices I & II for explanation). See Figure S2 for stationarity checks.

In all cases where $z > 1$, the prey species-richness isoclines are wedge-shaped contours that can be identified either by their species richness or by the carrying capacity K of the predator-free neutral community that yields the same average species richness. The lower boundaries of the species-richness isoclines change only slightly with K and very rapidly with ρ , especially around $\rho = 50$, which is the minimum value of ρ required for the persistence of the predator in the mean-field model; these lower boundaries define the region where increasing predation intensity will increase prey community diversity. The slope of the upper boundaries of the species-richness isoclines, and therefore the width of the wedge, increases as the amount of prey switching increases. As expected, more strongly prey-switching predators promote diversity across a broader region of parameter space.

The metacommunity richness M also plays a role in determining the attack rate ρ^* above which prey richness decreases below the predator-free average. Increases in M lead to decreases in ρ^* , meaning that predation is less likely to promote prey diversity in speciose metacommunities (Figure 3a). However, even in highly speciose metacommunities, prey-switching predators promote diversity across a broad range of parameter space as long as K and z are sufficiently high (Figure 3b).

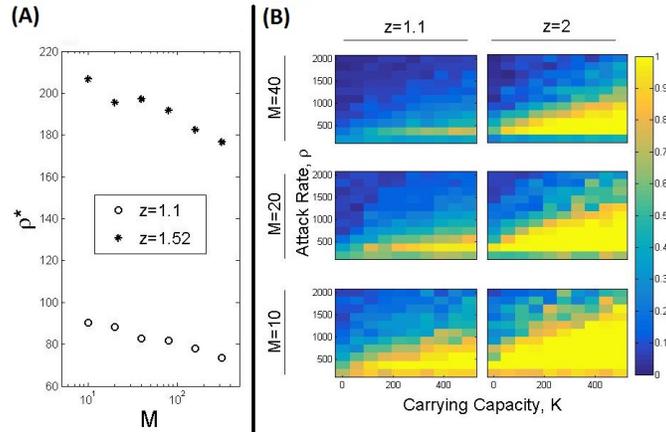


Figure 3: For constant $K = 1000$, increasing metacommunity richness decreases the range of predation intensities over which prey-switching predation improves species richness. (A) ρ^* decreases with increasing metacommunity richness. (B) Prey switching increases species richness over a narrower range of ρ , K , and z when prey metacommunities contain more species. See Figure S3 for stationarity checks.

Variance and quasicycles

Invariant or independently fluctuating predators tend to promote diversity over wider ranges of attack rates than predators with ordinary population dynamics (Figures 1 & 2). These differences can be understood in terms of the effect of

predators on prey population variance, which shapes the diversity landscapes of Figures 1 & 2 because the likelihood of extinctions increases when the population size decreases to below-average values. For all functional forms explored in this paper, predators increase the variance relative to the mean of the total prey population. Some functional forms, including Holling types I and II with ordinary predator population dynamics, cause oscillatory behavior and especially large increases in relative variance.

Because the expected community size grows linearly with K , a system-size expansion in K (van Kampen, 1981) suggests that the variance in community size should also grow linearly with K . The variances calculated from the replicate simulations behind Figure 1 indeed show approximately linear increases (Figure S7a). This trend becomes apparent only when communities with and without predators are disaggregated. Controlling for the expected community size, communities with Holling type I predators have variances roughly one hundred times that of communities without predators (Figure S7a). This massive increase in variance manifests as sustained periodic oscillations in the total community size (Figure S7b-d). These noise-induced oscillations, referred to as quasicycles, result from the interaction of predator-prey feedback and demographic stochasticity, and they persist even when the predator and every prey species has a species-specific death rate. Quasicycles substantially increase the variance in prey population beyond what could be expected without predator-prey feedback (Nisbet and Gurney, 1982; McKane and Newman, 2005). The increase in variance is associated with bottlenecks in total prey abundance (Figure S7b).

Comparing the species richness plots (Figures 1 & 2) for Holling type I & II predators and invariant predators (which do not produce quasicycles), we see that the increased variance associated with the quasicycle shrinks the region of parameter space where predators promote diversity. By comparing the invariant and independently fluctuating predators, we see that variation in predator populations reduces the ability of a prey-switching predator to maintain diversity. However, we also see that prey population size still modulates the effect of predation on prey diversity, even in the absence of quasicycles.

Transience

Following the introduction of a predator to a predator-free community, the expected species richness of a community enters a transient period before converging to the stationary expectation. For most parameter combinations, the initial change in species richness is in the same direction as the long-term change. However, when predators weakly reduce prey richness in the long term (ρ is slightly greater than ρ^*), predator introduction may cause a transitory increase in prey richness (Figures 4 & S14). This effect is most pronounced when the introduction involves a single individual predator. In this case, the transient effect on species richness probably results from the transient growth of the predator population. Because the per-prey attack rate is given by the product ρP , a predator population that has not yet reached its equilibrium abundance has exactly the same effect on prey dynamics as a larger predator population with

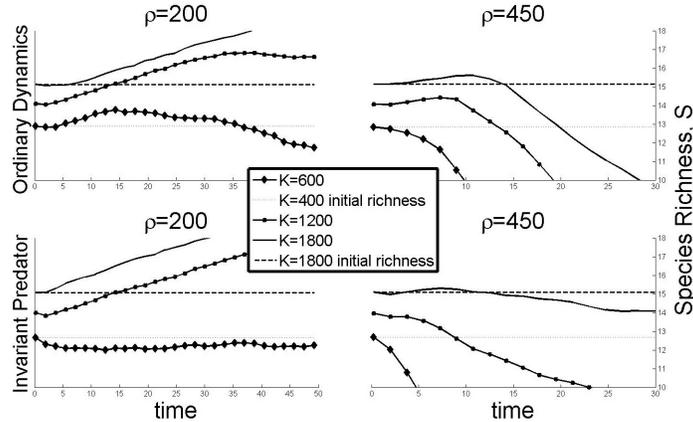


Figure 4: Predator introduction causes transient changes in prey richness generally in the same direction as the long-term change. However, introducing predators that weakly suppress richness in the long term can increase richness in the short term. The effect is most pronounced for introductions of small (sub-equilibrial) predator populations that exhibit transient growth. Trajectories show the average richness of twenty replicate communities (see Appendix II).

a lower attack rate. Even in the case of invariant predation, substantial time may pass before diversity begins to decline, presumably because the predator initially targets abundant species and is unlikely to kill rare species until the relative abundance distribution evens out (Figure 4).

Extinctions and species richness

A formal treatment of species richness in our model requires an examination of prey extinction-colonization dynamics. In this section, we combine simulations and analytic results to develop intuition for how prey-switching predators increase or decrease prey diversity in an island biogeographic context, where species richness is determined by a colonization-extinction equilibrium.

Predators do not directly affect colonization rates; they affect species richness entirely via extinctions. The propensity of extinctions at time t is

$$\psi^-(t) = n_1^-(t)\Psi_1(t) \quad (13)$$

where $n_1^-(t)$ is the propensity for death of a singleton and $\Psi_1(t)$ the number of singletons in the community. The average rate of extinctions is the expectation of equation 13 over the stationary distribution of community composition.

$$\mathbb{E}[\psi^-] = \mathbb{E}[n_1^-] \mathbb{E}[\Psi_1] + \text{Cov}[n_1^-, \Psi_1]. \quad (14)$$

To reduce extinction rates, a stabilizing mechanism must sufficiently decrease the expectation of n_1^- , the expectation of Ψ_1 , and/or the covariance between n_1^- and Ψ_1 .

Therefore, to understand how the parameters ρ , z , and K affect species richness, we need to understand how they affect the means and covariance on the right side of equation 14. We estimate analytically (Appendix V) and show by simulation (Figure 5a) that n_1^- increases with ρ and decreases with K and z . We show by simulation that for $z > 1$, the covariance $\text{Cov}[n_1^-, \Psi_1]$ increases with ρ , and that the effect is larger for small values of K and large values of z (Figures 5a & S8). Analysis of the effect of model parameters on Ψ_1 and the covariance $\text{Cov}[n_1^-, \Psi_1]$ is difficult (Appendix V) but simulations suggest that a U-shaped relationship between Ψ_1 and ρ accounts for much of the humped relationship between ρ and species richness. The reduction in Ψ_1 for intermediate ρ is magnified by increasing K and/or z over the range of parameters we consider in this paper (Figures 5a & S8).

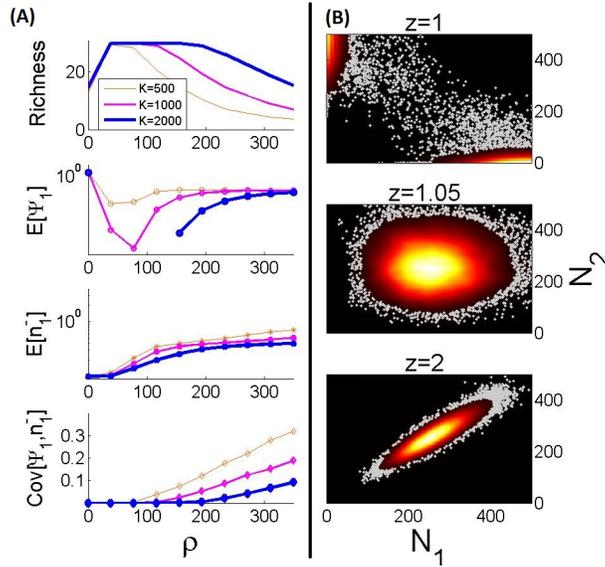


Figure 5: (A) Species richness in our model is controlled by the extinction rate. Based on equation 14, the scaling of species richness with ρ and K can be understood in terms of the scaling of the expectation of Ψ_1 , the expectation of n_1^- , and the covariance between Ψ_1 and n_1^- . (B) Prey switching can increase the covariance between prey populations, pulling the distribution of the population of a given species away from the boundaries where singletons occur and extinctions take place.

To understand the effect of z on Ψ_1 , we note that z does not affect the total community size or its variance. Instead, z increases the covariance between species by suppressing fluctuations in relative abundance (Figure 5b). We further note that the variance of population size of a particular species, given that it is present in the community, relates to the variance in the total community size:

$$\sigma_C^2 = S\sigma_N^2 + S(S-1)\sigma_{i,j} \quad (15)$$

where σ_C^2 is the variance in community size, σ_N^2 the variance in population size, and $\sigma_{i,j}$ the covariance between populations of different species. Symmetries in our model ensure that $\sigma_{i,j}$ is constant for all pairs of prey species. The increased covariance $\sigma_{i,j}$ necessarily entails a decrease in σ_N^2 . Thus, increasing z decreases Ψ_1 by moving the multivariate prey densities away from the boundaries where singletons occur (Figures 5b, S9). Prey-switching predation of the van Leeuwen functional form similarly increases the covariance between prey populations (Figure S9). Since frequency-dependent stabilizing mechanisms like prey-switching predation dampen fluctuations in relative abundances, other statistics beyond σ_N^2 and $\sigma_{i,j}$, such as the variance in relative abundance, σ_R^2 , may yield useful scale-free proxies for extinction rates (Figure S13, Appendix V).

Discussion

Whether in oceanic mesocosms, tropical islets, or remnant continental wildlands, a field biologist always encounters finite, discrete populations with stochastic dynamics. Species appear and vanish from islands or study plots, and these stochastic colonizations and extinctions are fundamental to metacommunity theory and island biogeography. When differential equations approximate our study systems well, they provide profound mechanistic insight and predictive power (Paine and Levin, 1981; Perelson et al., 1996). But some systems display important stochastic dynamics that cannot be captured by differential equations alone. We have shown that when prey species richness is the variable of interest, some predator-prey systems belong to this category across a range of parameters that the field biologist might encounter in nature, particularly where the intensity of competition between prey is high (K is small). Differential equations continue to provide important insight, and indeed they form the backbone of our simulation model, but they fail to predict key features of our system.

For all of our simulations, simple analysis of the corresponding differential equations suggests that the predator should stabilize the system and promote prey diversity (Holt, 1984; Abrams and Matsuda, 1996, 2003, 2004). By incorporating demographic stochasticity in finite populations, we show that the effect of prey-switching predation depends predictably not only on the predator’s attack rate and degree of prey switching, but also on the prey carrying capacity and the prey metacommunity richness. These variables together determine whether a given predation regime will increase or decrease the richness of an otherwise neutral prey community.

For a given prey carrying capacity K and prey-switching exponent z , there is a critical predation intensity ρ^* below which equilibrium prey richness is increased relative to the predator-free community, and above which equilibrium prey richness is reduced. Similarly, for a given attack rate ρ and prey-switching exponent z , there is a critical community size K^* below which equilibrium prey richness is reduced, and above which prey richness is increased. Increasing the metacommunity richness, M , decreases the relative abundance of each prey

species and consequently decreases ρ^* . This fundamental result is robust to several important variations in model assumptions, but it is invisible both to stability analysis of ordinary differential equations and to analysis of island-biogeographic communities which lack the niche detail of predator-prey interactions.

We suspect that the phenomena described here are not restricted to prey communities regulated by a single prey-switching predator, but might also extend to communities regulated by a suite of specialist predators (Janzen, 1970; Connell, 1971) or generalist predators regulating hierarchically competitive prey (Paine, 1966, 1969; Caswell, 1978). Any stabilizing predation regime that increases prey death rates will result in a tension between stabilization and stochastic intensification to determine prey diversity. We chose to consider prey-switching predators because simulations of a single prey-switching predator are simple to implement, easy to interpret, and contain fewer parameters than simulations of an entire suite of specialist predators.

Empirical Consequences

We estimate our model parameters from the empirical literature (see Appendix I). Doing so helps to ensure that the behavior we observe might exist in nature, because it exists at reasonable values and combinations of parameters. Moreover, our procedures for determining model parameters rely on empirically available measurements from real-world systems. Our model can therefore be fit to a real world system using empirical data that might reasonably be available.

Scale-dependence of predator effects

Predators are known to regulate prey diversity in many systems, sometimes increasing prey diversity (Paine, 1966; Stier et al., 2014) and sometimes decreasing prey diversity (Schoener and Spiller, 1996; Crooks and Soulé, 1999). Understanding the influence of predation on prey diversity is crucial for ecologists attempting to resolve the relative importance on top-down and bottom-up processes as controls on ecosystems (Dobson, 2014) because in some systems, predation is essential for the maintenance of species diversity (Terborgh, 2012). Empirical observations of predation affirm the theoretical notion that predators create a rare-species advantage by targeting abundant prey (Murdoch, 1969; Martin, 1988). Even when individual predators do not alter their preferences, the numerical response of specialist predators can produce a similar effect.

It is unsurprising that this rare species advantage is not always sufficient to ensure that predators increase prey diversity. With sufficiently large attack rates (or numerical responses), predators increase the probability of prey extinction regardless of switching. This result is confirmed by many recorded instances of predators decimating prey diversity to such an extent that prey switching is unlikely to have ameliorated the consequences in any meaningful way (Savidge, 1987; Blackburn et al., 2004).

The important insight from our models is that parameters other than attack rates, such as the metacommunity richness M or the carrying capacity K , can reverse the effect of predation on species richness. The carrying capacity K is of considerable interest, because in nature it is controlled by factors such as productivity and area. In systems where prey communities remain well-mixed and productivity-per-area remains constant across a range of spatial extents, horizontal slices of the plots in Figures 1 and 2 can be read as species-area relationships (SARs; Figure S10). Many predation regimes that increase diversity at large spatial scales reduce prey diversity at small scales. This possibility complicates the interpretation of empirical studies of the effects of predators on prey diversity. Landmark studies of predation and prey diversity have occurred at very small spatial scales, in mesocosms or on small islands (Schoener and Spiller, 1996; Blackburn et al., 2004; Stier et al., 2014). For example, Spiller and Schoener’s (1998) seminal work on Bahaman spider communities and their *Anolis* predators examined replicate communities with fewer than 300 spiders each in the absence of predators, a value at the low-end of the range of the parameter space that we consider. These studies have done an admirable job elucidating the effects of predation at small scales, but caution is warranted in applying the results to large blocks of continuous habitat.

Predators and prey frequently operate on different spatial scales, with predators ranging across larger areas than their prey, and coupling spatially distinct food webs (McCann et al., 2005). By confining both predators and prey to a single well-mixed patch, our models fail to account for this reality. However, our model results hold for both invariant predator populations and independently fluctuating predator populations where predator dynamics are independent of prey dynamics in the patch. Whereas the Holling type I model represents a limiting case where predator populations respond strongly to prey populations at the patch scale, these represent the opposite case, mimicking predators that range so widely that the dynamics at the patch scale do not appreciably affect the amount of predation experienced at the patch scale. Furthermore, experimental manipulations of predation often confine prey and predators to a single patch (Spiller and Schoener, 1998), or even restrict predator movement below that which is available to prey (Stier et al., 2014).

Conservation in habitat fragments

Just as caution is warranted in extrapolating empirical results from small scales to large, any empirical or theoretical understanding of the diversity-promoting effects of a predator in a large, intact system might not apply in small isolates. Habitat fragmentation is a ubiquitous global process that creates such isolates (Wilcove et al., 1986; Skole and Tucker, 1993; Fahrig, 2003; Kareiva, 1987). Increasingly, managing for predator persistence or reintroduction is viewed as a tool to enhance conservation outcomes for species at lower trophic positions through top-down effects (Nimmo et al., 2014). However, managing for predator persistence (or reintroduction) in fragmented habitats may negatively impact prey diversity even when the predator has unambiguous positive impacts in in-

tact systems. Note, however, that our results are relevant for cases when the focal species of conservation interest are potential prey. In some highly fragmented systems (e.g. Brazil's Atlantic forest), predators and prey alike are in serious danger of extinction (Pardini, 2004). Our model shows that managers cannot assume, based on the ability of prey to coexist with each other and with their predators in large landscapes, that such coexistence will be maintained in small forest reserves; in the face of evidence that predators enhance prey coexistence in large blocks of habitat, our model raises concern that fragmentation could reverse the predator's beneficial effect.

Coexistence in hyperdiverse communities

Our model predicts that predators are more likely to increase species richness on habitat patches subject to immigration from a depauperate metacommunity than on patches subject to immigration from a speciose metacommunity. On its face, this result is contradicted by the mounting evidence that predation plays a crucial role in the maintenance of diversity in hyperdiverse communities, such as in humid tropical forests (Janzen, 1970; Connell, 1971; Bagchi et al., 2014). However, at sufficient spatial scales switching predation continues to increase diversity across a reasonably broad region of parameter space, even in hyperdiverse communities. Furthermore, our results apply exclusively to systems where predators reduce the total density of their entire prey guild to substantially lower values than could be achieved in the absence of predation. Because tropical trees compete ferociously for canopy space (Hubbell, 2001), our model assumptions are probably invalid for tropical trees.

More individuals hypothesis

The tendency of predators to reduce diversity at small scales can be viewed as a natural consequence of the more individuals hypothesis (Srivastava and Lawton, 1998), which suggests that sites containing more individuals tend to contain more species and has received substantial empirical support (Hurlbert, 2004). Where prey-switching predation decreases species richness in our models, it does so partly by reducing the density of individuals, so a focal patch contains fewer individuals and fewer total species. It is tempting to imagine that this sampling effect causes a prey-switching predator to increase beta diversity, especially given a prey-switching predator's tendency to increase diversity at large scales, but to reduce diversity at small scales (Whittaker, 1960). However, this presumption rests on a shaky premise, because our models of richness at small spatial scales represent the richness of single small isolates, not of small patches integrated into a larger landscape and interconnected by migration. Such a model would be a useful extension of our work.

Extensions

Our investigations leave multiple open questions for future work. One interesting result from our simulations is that prey-switching predation increases the covariance between time-series datasets of prey populations. The increased covariance in communities subjected to prey-switching predation is due to prey-switching predation dampening stochastic fluctuations in relative abundances. Although the mean prey population sizes remain constant, the increased correlation with increased prey-switching predation is a stochastic analog the observation that prey-switching predation can change apparent competition to apparent mutualism between prey species (Holt, 1977; Abrams and Matsuda, 1996): stochastic decreases in the relative abundance of one species will lead to transient increases in the death rates of other, more abundant species and so transient stochastic increases in harvesting of one of two equally abundant species should lead to a transient decrease in the abundance of the other species. The prey populations we've considered still satisfy Holt's notion of apparent competition in the behavior of their means; this stochastic apparent mutualism exists only in species that stochastically fluctuate about their mean. This increase in temporal covariance between prey populations under prey-switching predation is empirically measurable - as is the hypothesized concomitant reduction in variance of prey population size - and thus fruitful grounds for testing our hypotheses about predators' effects on statistical properties of prey populations.

We hypothesize that the existence of ρ^* and K^* extend to more general systems with frequency-dependent death in finite, stochastic communities. This points the way to future studies of predator-prey communities with multiple predators, specialist predators/pathogens, non-neutral niche structure among the prey, and extrinsic environmental variability. The degree of prey-switching could be influenced by population size (Abrams and Matsuda, 2003), as well as patterns in spatial variation of prey relative to the home range of predators. Multi-patch extensions of our model could shed new light on species persistence after habitat fragmentation (Terborgh et al., 2001), and could elucidate the potential stabilizing and destabilizing impacts of predators that spatially couple local food webs (McCann et al., 2005).

If ρ^* and K^* generalize to a large class of realistic top-down mechanisms of biodiversity maintenance, then their scaling properties have major implications for our understanding of diversity in natural systems, the management of predators, and reserve design. For example, if top predators such as wolves (Becker, 2008) or lions (Bissett et al., 2012) exhibit very weak prey-switching, they may have beneficial effects on species richness that are only apparent in large swaths of connected habitat containing large prey populations. Contemporary management decisions about reserve size such as the range expansion of the Mexican Gray Wolf (*Canis lupus baileyi*) are often motivated by arguments of population viability (USFWS, 2014), but the diversity of the predators' diet (Reed et al., 2006; Merkle et al., 2009) and the potential scaling of the predators' impact on the robustness of prey populations with reserve size would motivate range expansions of prey-switching top predators on the basis of the viability of

the prey. The large historical ranges of these top predators compared to their smaller contemporary ranges mean that despite the predators' persistence their full effects on ecosystem structure & function on the landscape scale are not currently realized (Terborgh and Estes, 2010).

Acknowledgements

We thank our colleagues S. Levin, H. Horn, A. Dobson, J. Dushoff, A. Hein, C. Torney, J.E.S. Socolar, and D. Wilcove for many useful discussions and feedback. Two anonymous reviewers provided comments that greatly improved this manuscript. ADW and JBS both acknowledge support from NSF graduate research fellowships.

References

- Abrams, P., and H. Matsuda. 1996. Positive indirect effects between prey species that share predators. *Ecology* 77:610–616.
- . 2003. Population dynamical consequences of reduced predator switching at low total prey densities. *Population Ecology* 45:175–185.
- . 2004. Consequences of behavioral dynamics for the population dynamics of predator-prey systems with switching. *Population Ecology* 46:13–25.
- Bagchi, R., R. E. Gallery, S. Gripenberg, S. J. Gurr, L. Narayan, C. E. Addis, R. P. Freckleton, and O. T. Lewis. 2014. Pathogens and insect herbivores drive rainforest plant diversity and composition. *Nature* 506:85–88.
- Bangs, E. E., and S. H. Fritts. 1996. Reintroducing the gray wolf to central Idaho and Yellowstone National Park. *Wildlife Society Bulletin* pages 402–413.
- Becker, M. S. 2008. Applying predator-prey theory to evaluate large mammal dynamics: wolf predation in a newly-established multiple-prey system. ProQuest.
- Bissett, C., R. T. Bernard, and D. M. Parker. 2012. The response of lions (*Panthera leo*) to changes in prey abundance on an enclosed reserve in South Africa. *Acta Theriologica* 57:225–231.
- Blackburn, T. M., P. Cassey, R. P. Duncan, K. L. Evans, and K. J. Gaston. 2004. Avian extinction and mammalian introductions on oceanic islands. *Science* 305:1955–1958.
- Caswell, H. 1978. Predator-mediated coexistence: a nonequilibrium model. *The American Naturalist* 112:127–154.

- Chase, J. M., P. A. Abrams, J. P. Grover, S. Diehl, P. Chesson, R. D. Holt, S. A. Richards, R. M. Nisbet, and T. J. Case. 2002. The interaction between predation and competition: a review and synthesis. *Ecology Letters* 5:302–315.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics* 31:343.
- Connell, J. H. 1971. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. *In* P. J. den Boer and G. Gradwell, eds., *Dynamics of Populations*. Pudoc, Washington.
- Crooks, K. R., and M. E. Soulé. 1999. Mesopredator release and avifaunal extinctions in a fragmented system. *Nature* 400:563–566.
- Dobson, A. P. 2014. Yellowstone wolves and the forces that structure natural systems. *PLoS Biology* 12:e1002025.
- Elliott, J. M. 2004. Prey switching in four species of carnivorous stoneflies. *Freshwater Biology* 49:709–720.
- Elton, R. A., and J. J. D. Greenwood. 1970. Exploring apostatic selection. *Heredity* 25:629–633.
- Fahrig, L. 2003. Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 34:487–515.
- Gillespie, D. 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* 81:2340–2361.
- Gravel, D., F. Massol, E. Canard, D. Mouillot, and N. Mouquet. 2011. Trophic theory of island biogeography. *Ecology Letters* 14:1010–1016.
- Haegeman, B., and M. Loreau. 2011. A mathematical synthesis of niche and neutral theories in community ecology. *J Theor Biol* 269:150–65.
- Heinlein, J., A. Stier, and M. Steele. 2010. Predators reduce abundance and species richness of coral reef fish recruits via non-selective predation. *Coral Reefs* 29:527–532.
- Holt, A. R., Z. G. Davies, C. Tyler, and S. Staddon. 2008. Meta-analysis of the effects of predation on animal prey abundance: evidence from UK vertebrates. *PLoS One* 3:e2400.
- Holt, R. D. 1977. Predation, apparent competition, and the structure of prey communities. *Theoretical Population Biology* 12:197–229.
- . 1984. Spatial heterogeneity, indirect interactions, and the coexistence of prey species. *American Naturalist* 124:377–406.

- . 2009. Toward a trophic island biogeography. *In* J. B. Losos and R. E. Ricklefs, eds., *The Theory of Island Biogeography Revisited*. Princeton University Press, Princeton.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography, vol. 32. Princeton University Press, Princeton, New Jersey USA.
- Hurlbert, A. H. 2004. Species–energy relationships and habitat complexity in bird communities. *Ecology Letters* 7:714–720.
- Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. *The American Naturalist* 104:501–528.
- Kareiva, P. 1987. Habitat fragmentation and the stability of predator–prey interactions. *Nature* 326:388–390.
- Lande, R., S. Engen, and B.-E. Saether. 2003. *Stochastic Population Dynamics in Ecology and Conservation* (Oxford Series in Ecology and Evolution). illustrated ed. Oxford University Press, USA.
- Martin, T. E. 1988. On the advantage of being different: nest predation and the coexistence of bird species. *Proceedings of the National Academy of Sciences* 85:2196–2199.
- McCann, K., J. Rasmussen, and J. Umbanhowar. 2005. The dynamics of spatially coupled food webs. *Ecology Letters* 8:513–523.
- McKane, A. J., and T. J. Newman. 2005. Predator-prey cycles from resonant amplification of demographic stochasticity. *Physical Review Letters* 94:218102.
- Merkle, J. A., P. R. Krausman, D. W. Stark, J. K. Oakleaf, and W. B. Ballard. 2009. Summer diet of the Mexican gray wolf (*Canis lupus baileyi*). *The Southwestern Naturalist* 54:480–485.
- Morozov, A., and S. Petrovskii. 2013. Feeding on multiple sources: towards a universal parameterization of the functional response of a generalist predator allowing for switching. *PloS One* 8:e74586.
- Murdoch, W. W. 1969. Switching in general predators: Experiments on predator specificity and stability of prey populations. *Ecological Monographs* 39:335–354.
- . 1977. Stabilizing effects of spatial heterogeneity in predator-prey systems. *Theoretical Population Biology* 11:252–273.
- Nimmo, D. G., S. J. Watson, D. M. Forsyth, and C. J. A. Bradshaw. 2014. Dingoes can help conserve wildlife and our methods can tell. *Journal of Applied Ecology* pages 281–285.

- Nisbet, R. M., and W. S. C. Gurney. 1982. *Modelling fluctuating populations*. Wiley Chichester ; New York.
- Paine, R. T. 1966. Food web complexity and species diversity. *American Naturalist* 100:65–75.
- . 1969. A note on trophic complexity and community stability. *The American Naturalist* 103:91–93.
- Paine, R. T., and S. A. Levin. 1981. Intertidal landscapes: disturbance and the dynamics of pattern. *Ecological Monographs* 51:145–178.
- Pardini, R. 2004. Effects of forest fragmentation on small mammals in an Atlantic Forest landscape. *Biodiversity & Conservation* 13:2567–2586.
- Perelson, A. S., A. U. Neumann, M. Markowitz, J. M. Leonard, and D. D. Ho. 1996. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* 271:1582–1586.
- Reed, J. E., W. B. Ballard, P. S. Gipson, B. T. Kelly, P. R. Krausman, M. C. Wallace, and D. B. Wester. 2006. Diets of free-ranging Mexican gray wolves in Arizona and New Mexico. *Wildlife Society Bulletin* 34:1127–1133.
- Savidge, J. A. 1987. Extinction of an island forest avifauna by an introduced snake. *Ecology* 68:660–668.
- Schoener, T. W., and D. A. Spiller. 1996. Devastation of prey diversity by experimentally introduced predators in the field. *Nature* 381:691–694.
- Skole, D., and C. Tucker. 1993. Tropical deforestation and habitat fragmentation in the amazon. satellite data from 1978 to 1988. *Science* 260:1905–1910.
- Smout, S., C. Asseburg, J. Matthiopoulos, C. Fernández, S. Redpath, S. Thirgood, and J. Harwood. 2010. The functional response of a generalist predator. *PLoS One* 5:e10761.
- Spiller, D. A., and T. W. Schoener. 1998. Lizards reduce spider species richness by excluding rare species. *Ecology* 79:503–516.
- Srivastava, D. S., and J. H. Lawton. 1998. Why more productive sites have more species: an experimental test of theory using tree-hole communities. *The American Naturalist* 152:510–529.
- Stier, A. C., K. M. Hanson, S. J. Holbrook, R. J. Schmitt, and A. J. Brooks. 2014. Predation and landscape characteristics independently affect reef fish community organization. *Ecology* 95:1294–1307.
- Terborgh, J. 2012. Enemies maintain hyperdiverse tropical forests. *The American Naturalist* 179:303–314.

- Terborgh, J., and J. Estes. 2010. Trophic cascades : predators, prey, and the changing dynamics of nature. Island Press, Washington, DC.
- Terborgh, J., L. Lopez, P. Nunez, M. Rao, G. Shahabuddin, G. Orihuela, M. Riveros, R. Ascanio, G. H. Adler, T. D. Lambert, and L. Balbas. 2001. Ecological meltdown in predator-free forest fragments. *Science* 294:1923–1926.
- Turelli, M. 1980. Niche overlap and invasion of competitors in random environments ii. the effects of demographic stochasticity. Pages 119–129 *in* W. Jager and H. Rost, eds. *Biological Growth and Spread*. Springer, Heidelberg.
- USFWS. 2014. Draft - Environmental impact statement for the proposed revision to the nonessential experimental population of the Mexican Wolf *Canis lupus baileyi* .
- van Kampen, N. G. 1981. *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam ; New York : New York.
- van Leeuwen, E., A. Brannstrom, V. A. A. Jansen, U. Dieckmann, and A. G. Rossberg. 2013. A generalized functional response for predators that switch between multiple prey species. *Journal of Theoretical Biology* 328:89–98.
- Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. *Ecological monographs* 30:279–338.
- Wilcove, D. S., C. H. McLellan, and A. P. Dobson. 1986. Habitat fragmentation in the temperate zone. *Conservation Biology* 6:237–256.

Supplemental information: Appendices I-VI

1 Appendix I: Parameter Estimation

We turn to the empirical literature where possible (and the theoretical literature where necessary) to obtain realistic estimates for the parameters in our models. Our goal is to ensure that our results of interest occur parameter values that correspond to realistic natural communities. Table S1, an expanded version of table 1, includes every parameter used in our models. Below, we explain our choices of values for each parameter in turn.

Lifespans: δ and γ

The choice of δ is an arbitrary free choice that specifies the timescale in our model. We use $\delta = 0.1$, and thus the expected lifetime of a prey (given that it does not succumb to predation) is $T_{prey} = 10$. Because predators are generally longer-lived than their prey, we set the per-capita death rate of predators $\gamma = 0.05$. Thus, predators have twice the average lifespan of their prey.

This choice reflects the average body mass ratio of predators to prey (roughly 100 : 1; Barnes et al., 2010; Brose et al., 2006) and the allometric relationship between body size and lifespan in mammals ($\log(\text{lifespan}) \approx 0.2 \log(\text{body mass})$; Sacher, 1959). Note, however, that the variance about both of these quantities is large, and a wide range of ratios $\frac{\gamma}{\delta}$ are reasonable. Note also that in our models, predators live on average more than twice as long as prey, because prey lives are routinely cut short by predation.

Population growth: r and K

When a (predator-free) community is nearly empty, the prey population growth rate r is given by $r = \delta(r_0 - 1)$, where r_0 is the average lifetime reproductive output of an individual under conditions of abundant resources, and in the absence of predation. Thus we obtain an estimate of r by obtaining values of r_0 from the literature. r_0 is variable across real-world systems, and note that in sexually reproducing species, r_0 is obtained by dividing by two the average number of offspring per individual. Below, we provide a sampling of empirical values: In absence of predation, a North American songbird might have a lifespan of four years (Calder, 1990) and a reproductive output of perhaps 4 individuals per pair per year, given abundant resources (Martin, 1995). In this case, $r_0 \approx 8$, and $r \approx 0.7$ (in this example, one time unit is ≈ 0.4 years). A black-billed whistling-duck (*Dendrocygna autumnalis*) might have a lifespan approaching eight years (Clapp and Klimkiewicz, 1982), and a reproductive output (under abundant resources and no predation) as high as 14 individuals per pair per year (Rohwer, 1988), yielding $r \approx 5.6$ (in this example, one time

Table S1: Parameters used in MATLAB simulations

Parameter	MATLAB Name	Units	Value(s)	Interpretation
δ	D	time^{-1}	0.1	Intrinsic prey deathrate; inverse prey lifespan
γ	G	time^{-1}	0.05	Intrinsic predator deathrate; inverse predator lifespan
r	r	time^{-1}	5	Maximum prey population growth rate
K	k	prey	200 – 2000	Carrying capacity; system size; inverse competition intensity
M	M	species	30; 10 – 360 (fig. 3)	Metacommunity richness
m	m	$\frac{\text{prey}}{\text{time}}$	0.3	Migration rate
α	alpha	$\frac{\text{predators}}{\text{prey}}$	10^{-3}	Ratio of predator births to predation events
ρ	P	$\frac{\text{prey}}{\text{predators} \cdot \text{time}}$	0 – 500*	Predator attack rate (prey deaths per predator)
z	z	–	1 – 2	Prey-switching exponent
Additional Parameters used in Robustness Checks				
h	h	$\frac{\text{predators} \cdot \text{time}}{\text{prey}}$	10^{-3}	Handling time (Holling Type II)
T_{ij}	h	$\frac{\text{predators} \cdot \text{time}}{\text{prey}}$	10^{-3}	Handling time (van Leeuwen)
ρ_i	p	$\frac{\text{prey}}{\text{predators} \cdot \text{time}}$	0 – 1500	Attack rate (van Leeuwen)
s_{ij}	s	–	1 if $i = j$ 0.15 if $i \neq j$	Prey similarity (van Leeuwen)
\bar{P}	PO	predators	$\frac{rK}{\aleph} (1 - \frac{\gamma}{\alpha \aleph})^\dagger$	Mean predator population; (Invariant and Indep. Fluctuating predators)
λ	L	time^{-1}	10	Mean reversion (Independently Fluctuating predator pop.)
δ_i and γ'	D(i)	time^{-1}	See table S2	Species-specific death rates for prey and predator

* Note that in the density dependent death simulation⁵⁸, $0 \leq \rho \leq 1000$. (Appendix II).† Where $\aleph = \rho(1 - \frac{h\gamma}{\alpha})$.

unit is ≈ 0.8 years). A female harbor seal (*Phoca vitulina*) might have a lifetime reproductive output of roughly 25 pups, yielding $r \approx 1.2$ (Härkönen and Heide-Jørgensen, 1990). Female adders (*Vipera berus*) in field conditions might have a lifetime reproductive output of 10.5 offspring on average, so a lower bound on r in this system would be $r = 0.43$ (Madsen and Shine, 1994), while a female leatherback turtle (*Dermochelys coriacea*) might lay 60 eggs every 5 years for 30 years after maturity, yielding $r \approx 17.9$ (Avens et al., 2009; Price et al., 2004). In insects, documented lifetime fecundities range from 5.2 to 2469, yielding a range of r from 0.16 to 123 (Honěk, 1993). Among marine invertebrates, fecundity under optimal conditions may be difficult to estimate. For example, one sea hare (*Tethys californicus*) can lay half a billion eggs in half a year, but an unknown fraction (presumably very small) are fertilized and recruit into the population, even under optimal conditions (MacGinitie, 1934).

We pick $r \approx 5$ as a reasonable ballpark value; similar values (6.0 ± 0.4) have been found for meadow voles (*Microtus pennsylvanicus*) (Turchin and Ostfeld, 1997).

We examine several values of K , all between $K = 100$ and $K = 2000$. The upper bound corresponds roughly to, for example: the number of individual birds breeding in one square kilometer of temperate or tropical forest (Terborgh et al., 1990); the number of lizards in two square kilometers of dry Sonoran desert (González-Romero et al., 1989); one tenth the number of fish in a 3.8 hectare lake (average depth 8 m) in Michigan (Brown and Ball, 1943); the number of trees > 10 cm diameter-at-breast-height in four hectares of tropical forest (Condit et al., 1999); or the number of bacteria in 10^{-4} grams of dry soil under maize cultivation in China (Li et al., 2002). Note that in our model, K represents the population density that would be realized in absence of predation, so the densities mentioned above tend to be somewhat lower than the appropriate densities for comparison. Counteracting this effect, population densities in our model should be thought of as referring to one guild of competing species, and so the densities mentioned above tend to be somewhat higher than the appropriate densities for comparison.

As an aside, the web-spider communities studied by Spiller and Schoener (1998) contained fewer than 300 individuals per replicate in the absence of predators.

Colonizations: m and M

In our model, the colonization rate is determined by the migration rate m and the number of species $M - S$ present in the metacommunity but absent in the local community. Realistic values for the metacommunity richness M range from $M = 1$ (e.g. the richness of amphibians on Antigua and Barbuda) to $M > 600$ (the number of breeding birds near Iquitos, Peru); $M > 2000$ (the richness of reef fish in the coral triangle—though some of these are narrowly endemic to smaller sub-regions); or even $M > 5000$ (the richness of freshwater fishes in the Amazon Basin—though again, many are sub-regional endemics).

In most of our simulations (all figures except Figure 3), we use $M = 30$,

which approximates the richness of snakes in the coastal plain of North Carolina, the richness of ungulates in the arid savanna of central Kenya, the richness of ladybirds in England, or the richness of birds breeding in a New England hardwood forest (Holmes et al., 1986). We also test the effect of varying M , using values ranging from 10 to 360.

In our model, m represents the rate at which immigrant individuals of all species arrive in the focal patch. Realistic values for m in nature vary tremendously, depending on the isolation of the focal patch and the biology of the focal guild. On remote oceanic islands, immigrations of terrestrial mammals and reptiles are rare indeed. In flow-through mesocosms, immigrations could account for a majority of the individuals present. We use $m = 0.3$. Note that at equilibrium, prey have on average one offspring per lifetime, and in absence of predation, prey therefore have one offspring per $\delta^{-1} = 10$ timesteps, and the community produces $\frac{K}{10}$ offspring per timestep. Thus for $K = 2000$, we expect that migrants account for roughly 0.1% of individuals in the population, while for $K = 200$, that figure rises to over 1%.

Our choices of m and M should be viewed as somewhat arbitrary, except that we chose them intentionally to produce partly-filled communities for the range of values of K that we consider. This allows both diversity-promoting and diversity-suppressing effects of predators to express themselves in simulations.

Predator production: α

In our model, α represents the probability that a predation event will produce an additional predator, and α^{-1} is the average number of prey that a predator must consume in order to reproduce once. Our functional form assumes that the numerical response of the predator population is a linear function of the number of prey consumed, an assumption that is valid in at least some systems (Lawton et al., 1975). Alpha depends on the efficiency ϵ of biomass conversion between trophic levels and on the biomass ratio τ of predators to prey: $\alpha = \frac{\epsilon}{\tau}$. In nature, the efficiency of biomass conversion ϵ is variable because it depends not only on the efficiency of nutrient uptake in the digestive system and the theoretical efficiency of conversion by cells in culture, but also on the metabolic costs associated with homeostasis and foraging (Calow, 1977). In a review of invertebrates, Calow (1977) encountered minimum biomass conversion efficiencies of 11% in leaf-eating lepidopteran larvae, and maximum efficiencies of $\approx 70\%$ in the mite *Steganacarus magnus* (with an outlying efficiency of 95% in the sea star *Asterias rubens*). Conversion efficiencies in vertebrates are lower (5-25% in juvenile salmon; Calow 1977), and as expected are lowest in endotherms (4.6% in mice; Timon and Eisen, 1970). An absolute lower-bound to the efficiency of biomass conversion is the trophic efficiency (the ratio of predator production to prey production). In the North Sea, the estimated average trophic efficiency is 3.7%, while the estimated trophic efficiency among organisms with mass less than 256 g is 27% (Jennings et al., 2002).

The ratio τ of individual predator mass to prey mass is similarly variable in nature. Extremes exist in marine systems, where the black swallower (*Chias-*

modon niger) takes prey more massive than itself ($\tau < 1$; Jordan, 1905), while a blue whale (*Balaenoptera musculus*) subsists on Antarctic krill (*Euphasia superba*) 380 million times less massive than itself ($\tau = 3.8 \cdot 10^8$; Tomilin, 1959; Clarke, 1976). Average biomass ratios of 100:1 appear to be typical in marine food webs ($\tau = 100$; Jennings et al., 2002; Barnes et al., 2010). This ratio is reasonable across a variety of systems (terrestrial and aquatic) and taxa (endotherms and ectotherms), though endotherms tend to have higher ratios while invertebrates tend to have lower ratios (Brose et al., 2006).

For our purposes, we fix $\alpha = 10^{-3}$, meaning that predators consume on average 1000 prey to self-replace. This choice might reflect, for example, $\epsilon = 0.1$ and $\tau = 100$ (reflecting the average predator-prey interaction in the North Sea food web; Jennings et al., 2002); or $\epsilon = .05$ and $\tau = 50$; et cetera.

Predation: ρ and z

Determining an appropriate predation rate ρ is crucial to our results, because the behavior we report is uninteresting if it occurs only for unrealistically high values of ρ . Direct estimation is precluded by the dearth of empirical studies that have succeeded in simultaneously measuring the total number of individual prey, the total number of individual predators, and the total rate of predation events in a natural system. However, in the context of our model, we can derive an empirical estimate for ρ by noting that the expected community size of prey without predators is $\bar{x}_0 = K$, whereas the expected community size with predators is $\bar{x}_\rho = \frac{\gamma K}{\alpha \rho}$ in the Holling Type I model. Thus, $\rho = \frac{\gamma \bar{x}_0}{\alpha \bar{x}_\rho}$. Following local predator extinctions, prey populations have been observed to increase by as much as 10-fold or 100-fold ($\frac{\bar{x}_0}{\bar{x}_p} \in [10, 100]$; Terborgh et al., 2001). Conservatively, we consider attack rates no higher than $\rho = 10 \frac{\gamma}{\alpha} = 500$. There is also a lower bound for ρ , namely the minimum value required for persistence of the predator population in the mean field model $\rho > \frac{\gamma}{\alpha} = 50$. We consider no nonzero predation rates lower than this value.

Empirical estimates of the prey-switching parameter z exist for both Minke Whales (*Balaenoptera acutorostrata*) and the Hen Harrier (*Circus cyaneus*) (Smout and Lindstrom, 2007; Smout et al., 2010). In both of these studies, z is allowed to vary by prey species and is parameterized in the context of a Holling Type II functional response for three prey species. The studies find species-specific values ranging from 1 to 4, with values between 1 and 2 being typical across species. In our investigation, we consider no values for z less than one, as these would imply that the predator preferentially targets rare species. When $z = 2$, our functional form for prey switching is a special case of a general form derived from recent theoretical work on prey switching, suggesting that it is reasonable to consider values as high as $z = 2$ (Morozov and Petrovskii, 2013). We therefore consider the range $z \in [1, 2]$.

Additional parameters

Handling times: h and T_{ij} Because these parameters serve as robustness checks on our general results, we sought high-end estimates to demonstrate that our qualitative results persist across a broad range of model assumptions. We set the handling times at 10^{-3} . Since a predator must consume on average $\alpha^{-1} = 10^3$ prey to self-replace, this means that a predator must spend on average 1 time unit, or five percent its expected lifetime, handling prey just to self-replace.

van Leeuwen predation: ρ_i and s_{ij} In van Leeuwen’s functional form, attack rates, prey similarity, and prey switching are intertwined. For $i = j$, we set the similarity $s_{ij} = 1$, meaning that prey of the same species are entirely self-similar. Otherwise, we set $s_{ij} = 0.15$, a low value that ensures that significant prey switching occurs in the functional response. For these choices, it was necessary to examine higher values of ρ_i than the values of ρ considered in other simulations in order to obtain a similar range of overall attack rates.

Invariant and independently fluctuating predator populations: \bar{P} and λ In simulations with invariant predator populations, predator populations are fixed at the equilibrium predator population in the corresponding Holling type II differential equation $\bar{P} = \frac{rK}{\aleph} (1 - \frac{\gamma}{\alpha\aleph})$, where $\aleph = \rho(1 - \frac{h\gamma}{\alpha})$.

In simulations with independently fluctuating predator populations, predators fluctuate according to a discrete Ornstein-Uhlenbeck process with propensities $p^+ = \lambda\bar{P} + \gamma P$ and $p^- = (\gamma + \lambda)P$. \bar{P} defines the mean, and λ determines the strength of mean reversion. We construct this model primarily to motivate mathematical understanding of the system, and we do not attempt to estimate λ from data. Instead, we examine $\lambda = 0$ (invariant predators) and $\lambda = 10$, a value chosen because it produces variance in predator populations similar to what we observe from the Holling type II predators with ordinary population dynamics.

Note that because we set a reflecting wall at $P = 1$, the process is not truly an Ornstein-Uhlenbeck process, and the true mean is slightly greater than \bar{P} .

Species-specific death rates δ_i and γ' In some simulations, we give every prey species, plus the predator, a unique species-specific death rate. We selected the death rates δ_i by drawing from a log-normal distribution with $E[\log(\delta_i)] = -2$ and variance $\text{Var}[\log(\delta_i)] = 1/4$. Of the $M + 1$ death rates drawn in this manner, we assign the minimum value to the predator γ' to ensure that it retains a longer average lifetime than its prey. The generation times are drawn once and then are fixed through time and across replicate communities. The randomly drawn generation times used in our analyses are available in Table S2.

Table S2: Species-specific death rates

	Death rate (time ⁻¹)
Predator	0.0311
Prey	0.0575
	0.0738
	0.0763
	0.0776
	0.0785
	0.0793
	0.0878
	0.0879
	0.0903
	0.0921
	0.0928
	0.1199
	0.1209
	0.1246
	0.1286
	0.1333
	0.1349
	0.1376
	0.1407
	0.1583
	0.1588
	0.1592
	0.1630
	0.1852
	0.2338
	0.2357
	0.2366
	0.2685
	0.2778
	0.2912

2 Appendix II: Supplementary methods

Non-neutrality with species-specific death rates

Although every species has the same expected per-capita growth rate at all points in time, the model with species-specific death rates is not neutral in the sense that two species with the same abundance but different values of δ will have different propensities for birth & death despite having the same average change in population size per unit time. This implies that individuals belonging to these different species will have different birth and death propensities, and that longer-lived species will tend to have higher lifetime fecundity than shorter-lived species in growing communities and lower lifetime fecundity in shrinking communities. Furthermore, long-lived species will have lifetime fecundities that integrate over longer timescales, and consequently their fecundities may better approach the average fecundities of organisms in the mean-field limit.

One surprising consequence of this non-neutrality is that species with longer generation times outperform species with shorter generation times. This hierarchical competition vanishes in the mean-field limit, but persists in finite communities, presumably due to fluctuations about the equilibrium community size that favor species with longer-generation times and less volatile populations. A full analysis of this hierarchical competition would distract from the focus of this paper; for now we present evidence for the hierarchical competition in Figure S11.

Behavior of density-dependent death simulations

Prey communities with density-dependent death exhibit two major differences in their island-biogeographic dynamics compared to prey-communities with density-dependent birth: they are more volatile, and the death propensity of a rare species does not necessarily increase with predator attack rate ρ , owing to the reduction in the propensity of death due to intra-guild competition. For the sake of contrast, we illustrate these effects by comparing a single prey population N regulated by density-dependence that occurs either entirely in the birth process, to a population N_D regulated by density-dependence that occurs entirely in the death process. In the context of equation 11 of our main paper, these correspond to $\beta = 1$ and $\beta = 0$, respectively.

The increased volatility of communities with density-dependent death is associated to higher rates of extinction due to higher stationary variance in population size. Recall the propensities of a predator-free prey population with density-dependence in birth and populations not much larger than K are

$$\begin{aligned} n^+ &= rN \left(1 - \frac{N}{K}\right) + \delta N \\ n^- &= \delta N \end{aligned} \tag{S1}$$

and the propensities for a population with density dependence in death are

$$\begin{aligned}\hat{n}^+ &= rN_D + \delta N_D \\ \hat{n}^- &= \frac{r}{K}N_D^2 + \delta N_D.\end{aligned}\tag{S2}$$

Over a short time-interval, the variance in the change of a population size is

$$\text{Var}[\Delta N] = n^+ + n^-\tag{S3}$$

which, for a predator-free prey community near its carrying capacity K , is

$$\begin{aligned}\text{Var}[\Delta N] &= 2\delta K \\ \text{Var}[\Delta N_D] &= 2(r + \delta)K.\end{aligned}\tag{S4}$$

It is clear that the populations with density-dependence in birth are less volatile; $\text{Var}[\Delta N] < \text{Var}[\Delta N_D]$. This short-term volatility in ΔN is not equivalent to the stationary variance in population size (and the associated extinction rates), but since these processes have the same drift near their mean, the near-equilibrium volatility of ΔN gives us the relative variance of isolated populations with density-dependence entirely in birth or entirely in death. Confirmation of the stationary variance requires an appropriate linear-noise approximation from either a martingale problem formalized by Stroock and Varadhan (1979), a Kramer-Moyal expansion, or a van Kampen system-size expansion. Because of its rigor and direct relation to our problem - the dependence of predator-mediated coexistence on system size - we choose to illustrate van Kampen's method below.

The master equation specifying the time-evolution of the probability mass function $f(N, t)$ is

$$\dot{f} = [(\mathbb{E}^1 - 1)n^- + (\mathbb{E}^{-1} - 1)n^+] f(N, t)\tag{S5}$$

where $\mathbb{E}^m f = f(N + m, t)$ is the step-operator defined by van Kampen (1981).

Decompose N into two variables at different scales, K and $K^{1/2}$:

$$\begin{aligned}N &= Kx + K^{1/2}\xi \\ f(N, t) &= \Pi(\xi, t)\end{aligned}\tag{S6}$$

and substitute these into equation S5 to obtain

$$\begin{aligned}\frac{\partial \Pi}{\partial t} - K^{1/2} \frac{\partial \Pi}{\partial \xi} \frac{\partial x}{\partial t} &= \left(-K^{-1/2} \frac{\partial}{\partial \xi} + \frac{1}{2} K^{-1} \frac{\partial^2}{\partial \xi^2} + \dots \right) \\ &\cdot (r + \delta) (Kx + K^{1/2}\xi) \Pi \\ &+ \left(K^{-1/2} \frac{\partial}{\partial \xi} + \frac{1}{2} K^{-1} \frac{\partial^2}{\partial \xi^2} + \dots \right) \\ &\cdot (Kx + K^{1/2}\xi) (\delta + rx + K^{-1/2}r\xi) \Pi.\end{aligned}\tag{S7}$$

Collecting terms of order $O(K^{1/2})$, we get

$$\frac{\partial x}{\partial t} = rx(1-x) \quad (\text{S8})$$

which defines the mean-field dynamics of the prey population relative to K . Note the stationary solution to the normalized mean-field system is $x^s = 1$.

Next, collect terms of order $O(1)$, we get

$$\begin{aligned} \frac{\partial \Pi}{\partial t} &= -(r+\delta) \frac{\partial}{\partial \xi} \xi \Pi + \frac{1}{2} x (r+\delta) \frac{\partial^2 \Pi}{\partial \xi^2} \\ &\quad + (\delta + 2rx) \frac{\partial}{\partial \xi} \xi \Pi + \frac{1}{2} x (\delta + rx) \frac{\partial^2 \Pi}{\partial \xi^2} \\ &= -r(1-2x) \frac{\partial}{\partial \xi} \xi \Pi + \frac{1}{2} x (2\delta + r + rx) \frac{\partial^2 \Pi}{\partial \xi^2}. \end{aligned} \quad (\text{S9})$$

We note that the stationary solution to this Fokker-Plank equation will be a Gaussian. We substitute in the stationary value of $x = x^s$

$$\frac{\partial \Pi}{\partial t} = r \frac{\partial}{\partial \xi} \xi \Pi + \frac{1}{2} (2\delta + 2r) \frac{\partial^2 \Pi}{\partial \xi^2} \quad (\text{S10})$$

and see that the stationary solution for fluctuations in density-dependent death ξ_D^s will be a Gaussian with mean 0 and variance

$$\text{Var}[\xi_D^s] = 1 + \frac{\delta}{r}. \quad (\text{S11})$$

In the model with density-dependent birth, we can follow the same steps and obtain the Fokker-Plank equation

$$\frac{\partial \Pi}{\partial t} = -r(1-2x) \frac{\partial}{\partial \xi} \xi \Pi + \frac{1}{2} x (2\delta + r - rx) \frac{\partial^2 \Pi}{\partial \xi^2} \quad (\text{S12})$$

resulting in stationary variance

$$\text{Var}[\xi^s] = \frac{\delta}{r}. \quad (\text{S13})$$

Thus, a Gaussian approximation for the stationary distributions of N and N_D yield

$$\begin{aligned} \text{Var}[N] &\approx \frac{\delta}{r} K \\ \text{Var}[N_D] &\approx \left(1 + \frac{\delta}{r}\right) K. \end{aligned} \quad (\text{S14})$$

Our instantaneous volatilities at $N = N_D = K$ calculated above give us the same relative values of the stationary variances of N and N_D due to their

having the same drift. Regardless, the result is clear: communities with density-dependent death have higher variance in community size than communities with density-dependent birth. For our choice of parameters, $\delta = 0.1$ and $r = 5$, the stationary variance of communities with density-dependent death is 51 times that of communities with density-dependent birth. Consequently, the timescale of neutral drift to extinction will be much faster for N than N_D , and migration rates must be increased to obtain the same average species richness and realistic-looking species-abundance distributions in absence of predators. To that end, all our simulations of density-dependent death had migration rates that were 51 times that of the normal, density-dependent birth migration rate, i.e. $m_D = 51m = 15.3$.

Communities with density-dependent death have an additional feature that causes qualitative difference in their extinction-colonization dynamics. In our ‘‘Analytic Results’’ section, we assume that the death propensity for a singleton invader, n_1^- , should always increase with ρ . This is a reasonable assumption for density dependent birth. Revisiting the singleton death propensities for density-dependent birth,

$$n_1^- = \delta + \rho \frac{\sum_i N_i}{K} \frac{P}{1 + \sum_i N_i^z} \quad (\text{S15})$$

we can make a conservative estimate that n_1^- will increase under prey-switching predation where $\frac{\partial P}{\partial \rho} > 0$ since $\sum_i N_i^z$ will shrink faster than $\sum_i N_i$ as ρ is increased for $z > 1$. In equation 2 of the main paper we see that \bar{P} is not monotonic with ρ , but if fluctuations in n_1^- are small we can approximate n_1^- by substituting in \bar{P} to get

$$n_1^- \approx \delta + \sum_i N_i \frac{r \left(1 - \frac{\gamma}{\alpha \rho}\right)}{1 + \sum_i N_i^z} \quad (\text{S16})$$

which increases with ρ for $z > 1$, since $\sum_i N_i^z$ will shrink faster than $\sum_i N_i$ as ρ increases. From this, we postulate that singleton death rate always increases with increasing predation intensity for communities with density-dependence entirely in birth.

However, for density-dependent death, this is not the case. The singleton death propensity is

$$n_1^- = \delta + \frac{\sum_i N_i}{K} \left(r + \frac{\rho P}{1 + \sum_i N_i^z} \right) \quad (\text{S17})$$

which can decrease with ρ if deaths due to intraguild competition $\frac{r}{K} \sum_i N_i$ decrease faster than the increase in the death due to predation on singletons. This leads to higher values of ρ necessary to obtain the same singleton death propensity, and may yield a slower rate of increase in the extinction rate with respect to ρ , and consequently a higher ρ^* as observed in Figure 2. The death rate of singletons in communities with density-dependence entirely in death

may decrease with increasing prey-switching predation intensity if intraguild competition is a major cause of mortality.

When we simulated communities with density dependent death, we found that predators increased prey diversity over a wide region of attack rates ρ , and the breadth of this region increased rapidly with K . To present an adequate picture of the changes in species richness with ρ and K , we found it necessary to examine a wider range of values for ρ than in other simulations.

Calculation of ρ^* for different metacommunity richnesses (Figure 3)

To calculate ρ^* for different metacommunity richness M , 4 replicate communities were simulated from the standard, Holling I functional form of predation for $K = 1,000$ and each value of $\rho = [0, 50, 100, 150, 200, 250, 300]$, $M = [10, 20, 40, 80, 160, 320]$ and $z = [1.1, 1.5]$. Simulations ran for $T = 3,000$ time units (300 prey generations) and species richnesses, S_ρ , over the last 2,000 time units were averaged for more robust estimates. For each value of M , the predator-free richness, S_0 , was obtained and then MATLAB's `interp1` function used to interpolate the species-richness curves, S_ρ , and obtain the value of $\rho^* > 0$ such that $S_{\rho^*} = S_0$.

Transient dynamics (Figure 4)

We wish to understand the transient dynamics of species richness following predator introductions to neutral communities. To initialize the neutral communities near the stationary distribution of the predator-free community for a given value of K , we initialize four replicate communities near the mean-field equilibrium and simulate the dynamics for $T=20000$ timesteps. We then average the species-abundance distributions from these four replicates, and we use this averaged distribution to initialize twenty new communities, each of whose dynamics we simulate for $T=1000$ timesteps. We then simulate predator introductions into each of these twenty replicate communities, and we monitor the time-evolution of each replicate. We monitor the dynamics by examining the community state every 0.25 timesteps for 50 timesteps (5 prey generations). Finally, for each time point, we average the species-abundance distributions across replicates to provide a general picture of the transient dynamics.

Extinction & Species Richness Analysis

Prey communities were simulated for $T = 40,000$ timesteps and sampled every time unit for the last 39,000 time units. The following parameters varied factorially: $K \in \{500, 1000, 2000\}$, $\rho \in \{\frac{7n\gamma}{9\alpha}\}_{n=0}^9$, and $z \in \{1, 1.05, 1.1, 2\}$. Average species richness, $E[n_1^-]$, $E[\Psi_1]$, and $\text{Cov}[n_1^-, \Psi_1]$ were obtained by taking the time-average of the respective quantity for each combination of ρ , K and z , where

$$n_1^- = \delta + \rho \frac{\sum_j N_j}{K} \frac{P}{1 + \sum_j N_j^z} \quad (\text{S18})$$

was the propensity of death for a singleton in the community calculated at each time point.

3 Appendix III: Stationarity checks

We are interested in the asymptotic or stationary species-richness of communities at island-biogeographic equilibrium. To simulate this quantity, we must determine the relaxation time required to ensure that the imprint of initial conditions fades and dynamic equilibrium is reached.

We first sought a conservative estimate of the relaxation time in the Holling type I model as follows. We surmised that the community would take longest to reach equilibrium at high carrying capacities and low predation rates. Since we explore parameter values where $0 \leq \rho \leq 500$ and $100 \leq K \leq 2000$, we used ten simulations with $\rho = 0$ and $K = 2000$, initialized near the deterministic equilibrium, to obtain estimates of the relaxation time for the species-abundance distribution. We estimate this time to be $T=20000$ (approximately 2,000 prey generations, Figure S4). In performing this stationarity check, we noted that species richness equilibrates much faster than the entire species-abundance distribution, approximating the long-term expectation by $T=2000$ (Figure S5). The qualitative response of species richness to changes in model parameters is apparent even earlier, by $T=500$. These observations allow us to perform many more simulations than would be possible otherwise. We establish our principal results in Figure 1 by simulating communities to $T=20000$ timesteps, and we examine species-abundance distributions only for communities simulated to $T=20000$. Most other results are based on simulations to $T=2000$ timesteps.

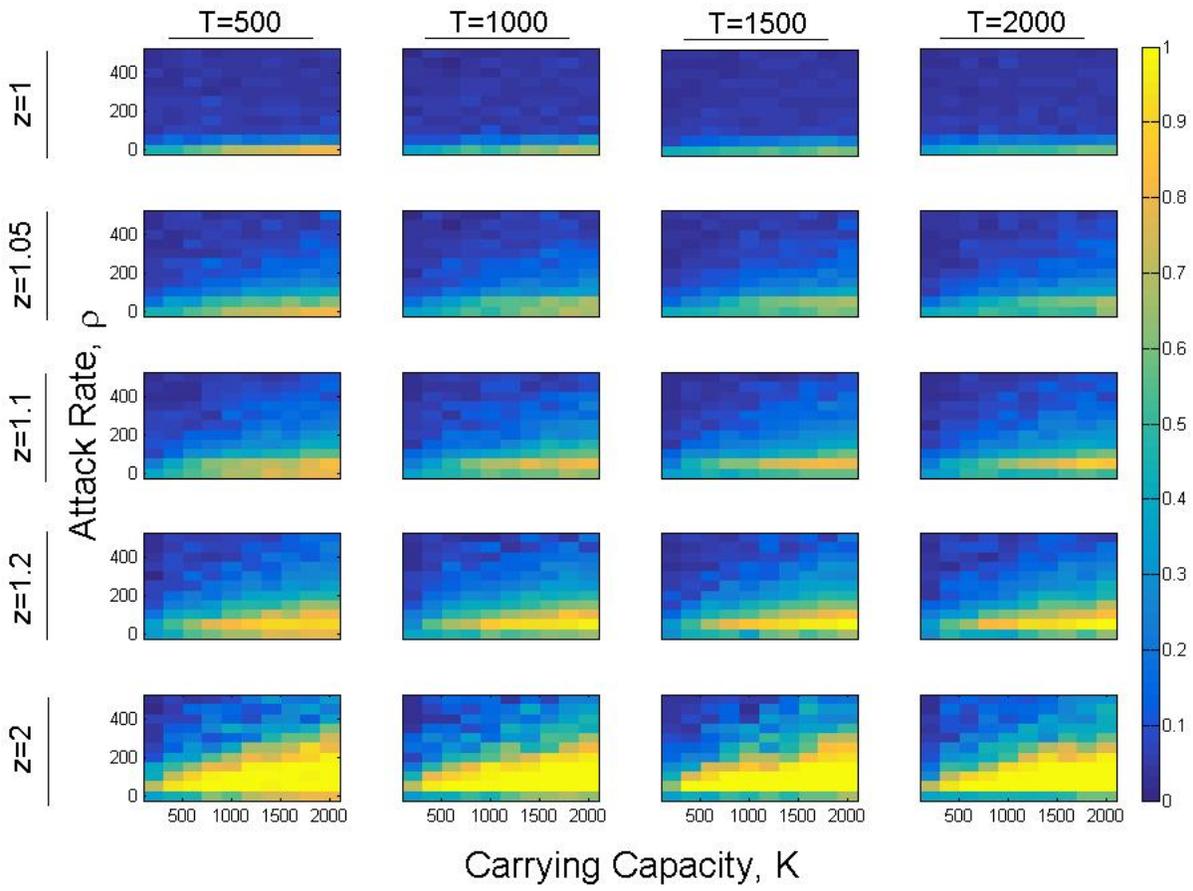


Figure S1: Stationarity check for Figure 1. Note: all simulations began as communities filled with all M species. The simulations with the longest relaxation time are the bottom right pixels of these figures where $\rho = 0$ and $K = 2,000$.

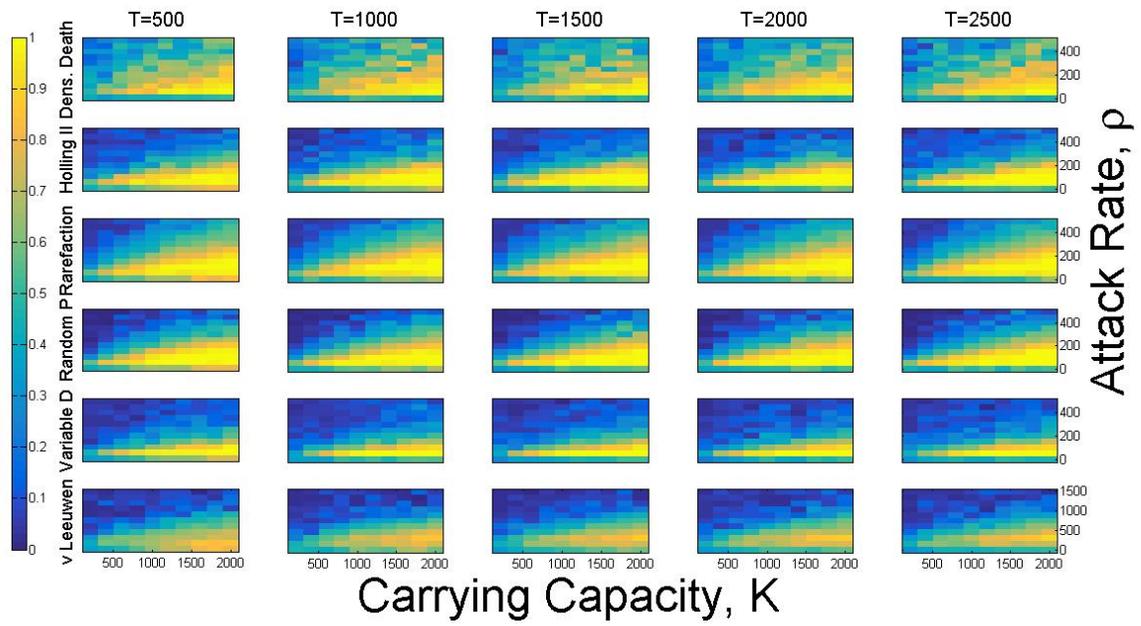


Figure S2: Stationarity check for Figure 2

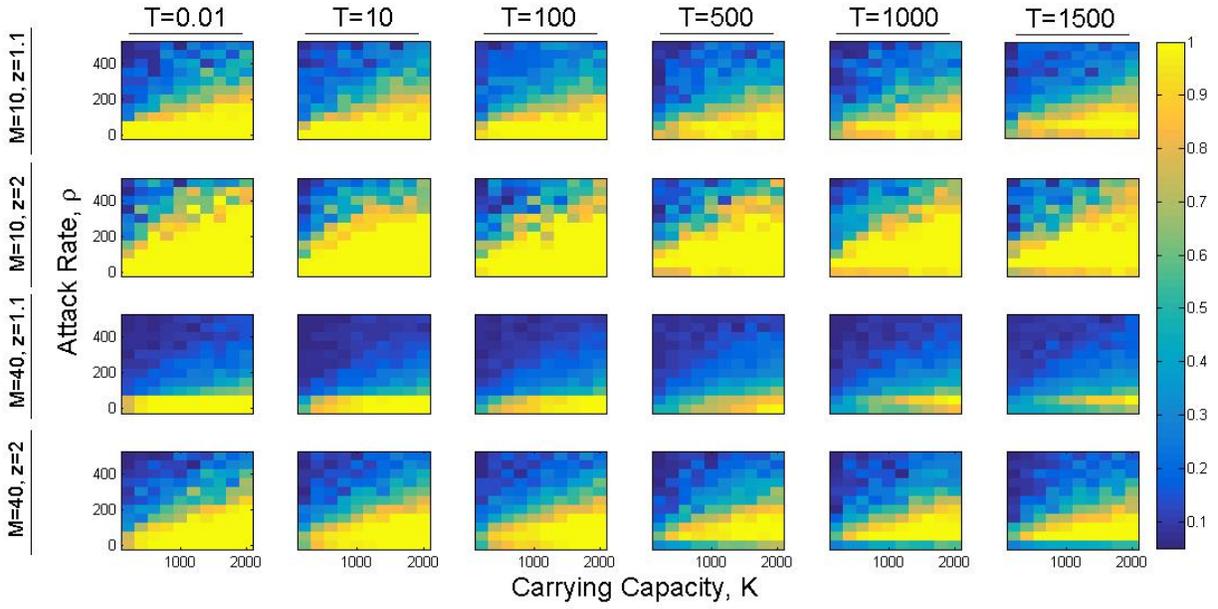


Figure S3: Stationarity check for Figure 3. For simulations varying M , we were unsure whether high or low values of M would have the longest relaxation times. We assumed that the relaxation time would vary monotonically with M , if at all, and so chose values of M outside the range considered in the corresponding figure of the paper ($M \in [10, 20, 40]$). Here we have $M \in [5, 45]$ and $z \in [1.1, 2]$.

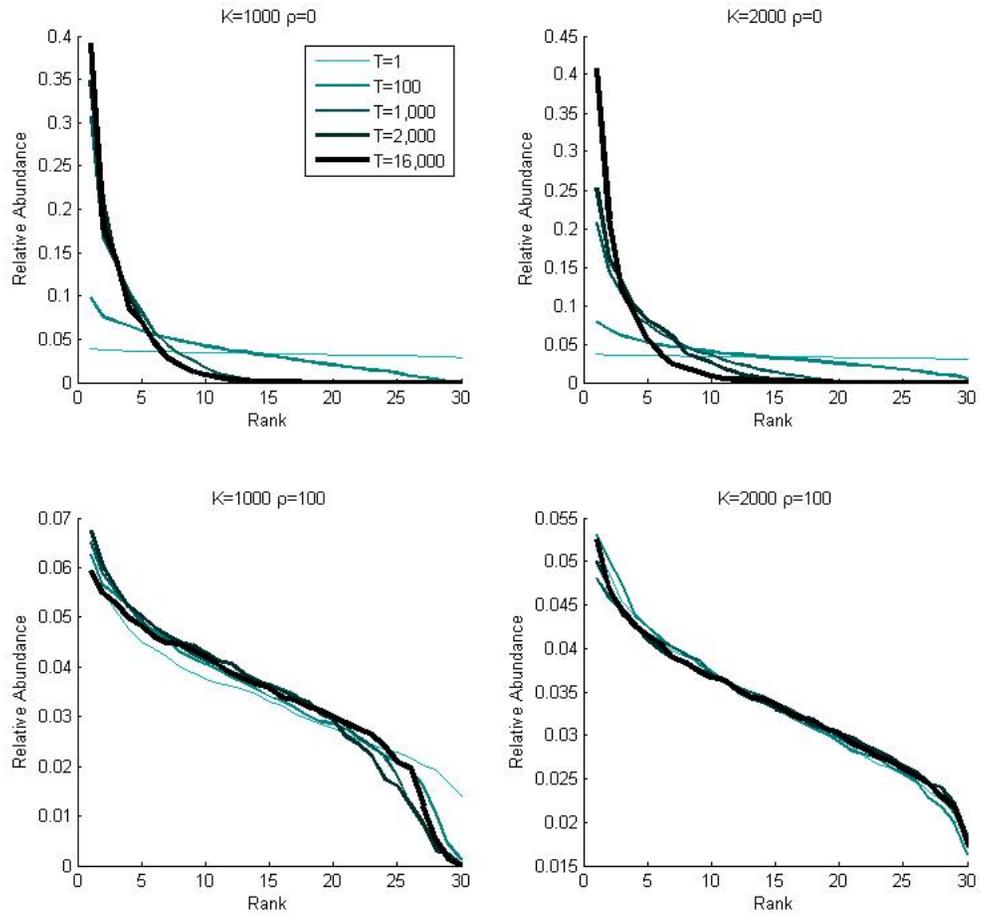


Figure S4: Stationarity of SAD over 10 replicates for $K \in [1000, 2000]$ and $\rho \in [0, 100]$, $z = 1.5$. Prey-switching predation increases the evenness of the community by decreasing the individual fitness of species with high relative abundances. The final SAD is determined by a balance of neutral drift yielding log-series-like distributions in the top panel and deterministic forcing for evenness.

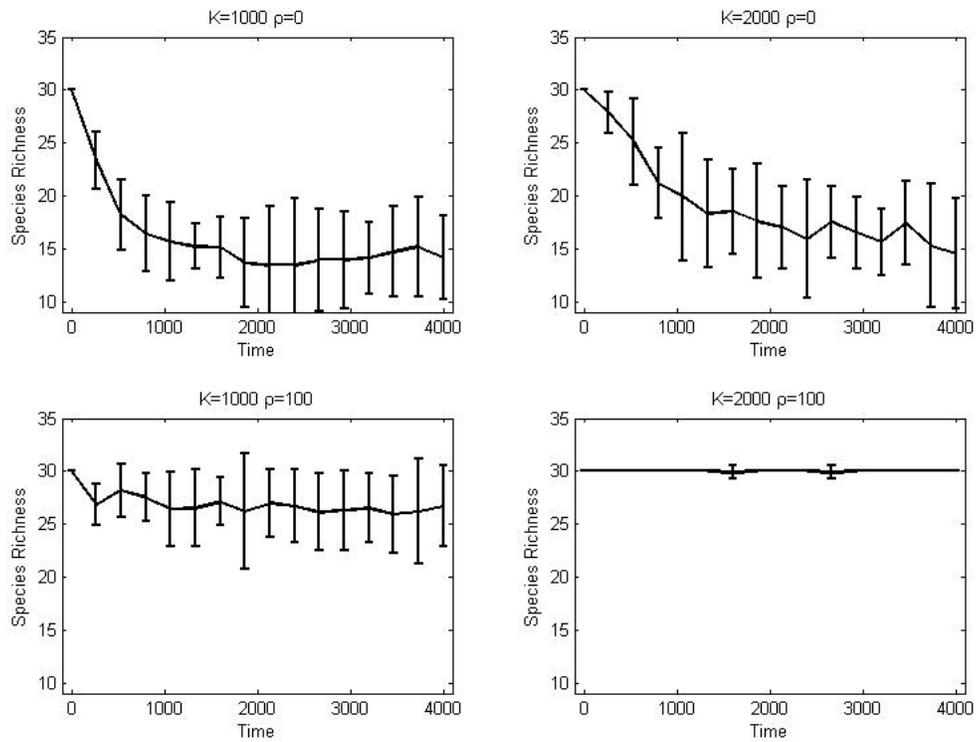


Figure S5: Relaxation of species richness over 10 replicates for $K \in [1000, 2000]$ and $\rho \in [0, 100]$, $z = 1.5$. Error bars show $\pm 2\sigma$. The predator-free system has the longest relaxation time and thus sets the lower bound on time needed for stationarity in species richness; for all simulations of species richness in this paper, $K \leq 2,000$ and $T = 2,000$.

4 Appendix IV: Additional supplementary figures

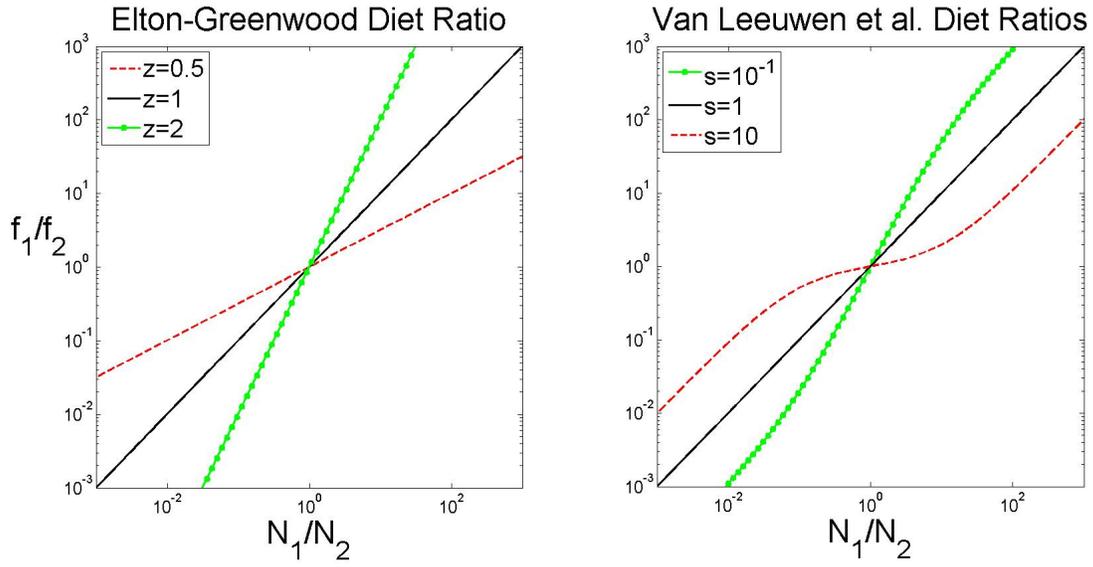


Figure S6: Prey-switching predators disproportionately consume the more abundant prey. (A) We define the per-capita rate of predation of species i as $\rho \frac{\sum_j N_j}{K} \frac{(N_i)^z}{\sum_j (N_j)^z}$, giving a diet ratio $f_i/f_j = \left(\frac{N_i}{N_j}\right)^z$. (B) Van Leeuwen et al. (2013), derived a per-capita rate of predation based on individual-level assumptions of attack rates, handling times, and the probability of switching from one prey species to another (see Robustness Checks)

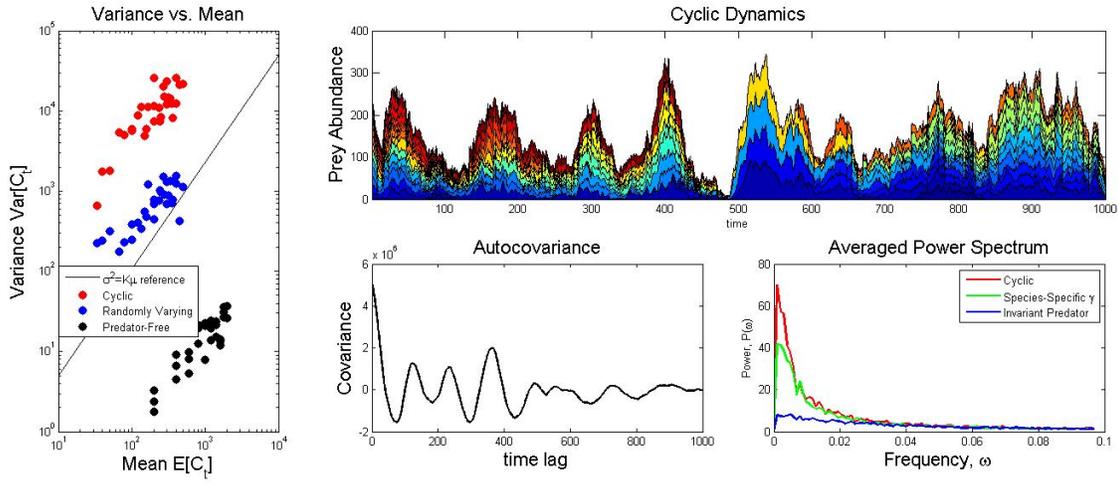


Figure S7: Demographic stochasticity causes sustained oscillations in prey community size which increase the variance of prey populations over time. (A) Variance in community size vs. expected community size across replicate communities used to generate species richness plots in Figure 2. Black dots represent communities where $\rho = 0$; red dots are communities with ordinary Holling I predators, and blue dots are communities with independently fluctuating predators. All predator-containing communities plotted here fall within $\rho \in [250, 350]$. The variance increases approximately linearly with the mean: $\text{Var}[C_t] = \text{const} + E[C_t]$. (B) A stacked time-series plot of prey abundances for a species-specific death rates simulation. The prey and predator have generation times that are non-integer multiples of one another, but oscillations nevertheless dominate the community dynamics. (C) The auto-covariance indicates the existence of quasicycles. (D) An averaged power spectrum of prey community size over 42 replicate simulations for $K = 1000$, $T = 2000$, and $\rho = 250$.

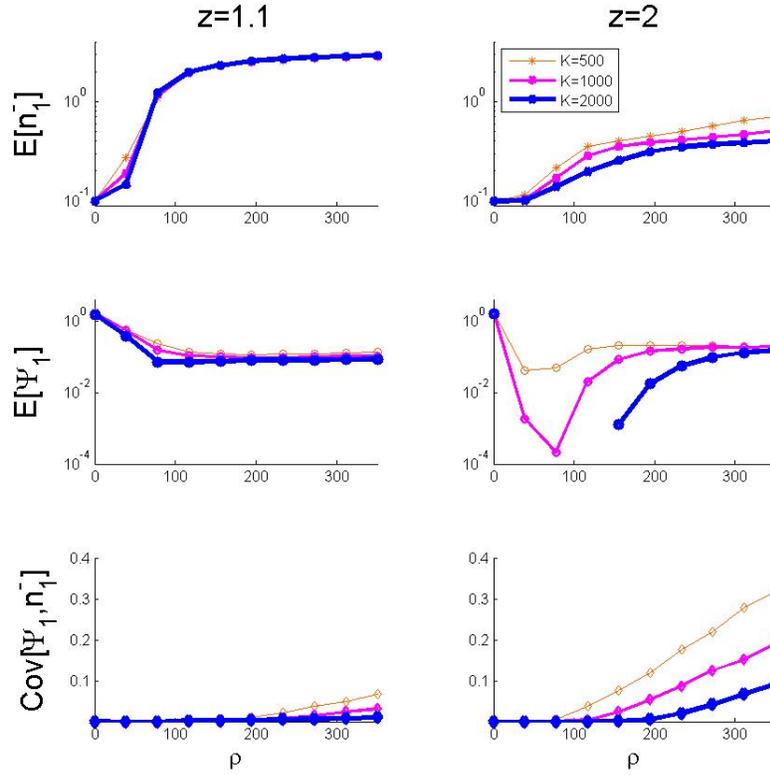


Figure S8: Extinction rates as a function of ρ , K and z . Breaking down the average rate of extinctions into $E[n_1^-]$, $E[\Psi_1]$, and $\text{Cov}[n_1^-, \Psi_1]$ reveals that much of the gains in species richness for low values of ρ are attributable to decreases in $E[\Psi_1]$, which we believe is related to decreases in the variance of relative abundance that shrink the fluctuations of prey community composition away from the boundaries where singletons occur. A major cause of increased extinction rates (and decreased species richness) for high values of ρ is the increased covariance $\text{Cov}[n_1^-, \Psi_1]$, an effect that is pronounced for higher values of z and lower values of K .

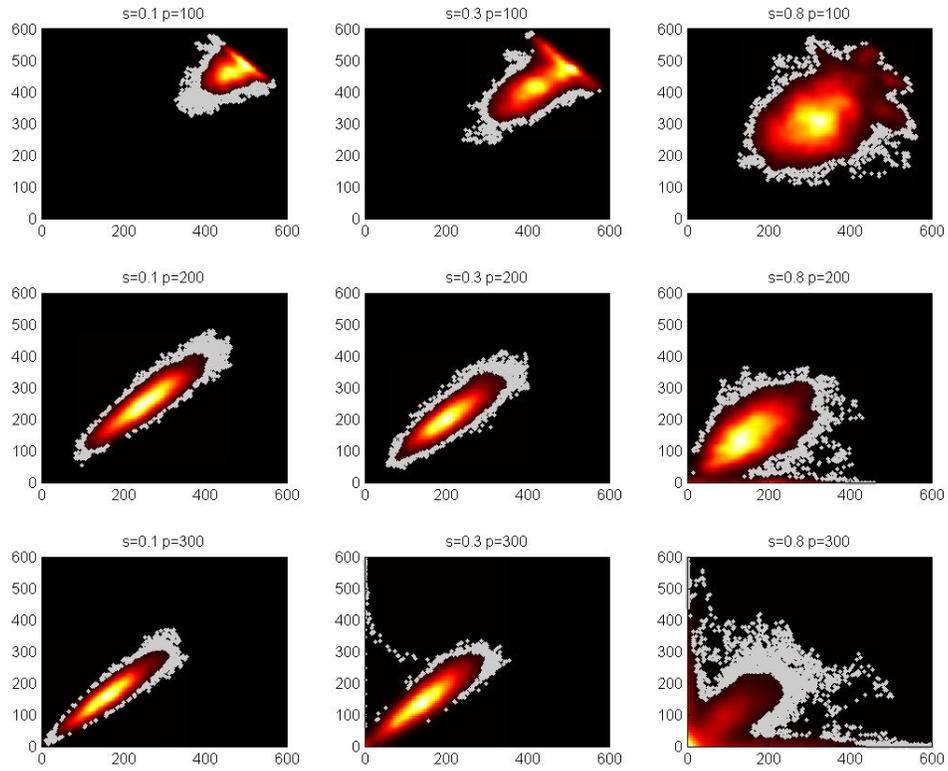


Figure S9: Van Leeuwen prey-switching predation does not separate the overall predation rate and the degree of prey-switching due to the costs of switching between prey. Consequently, increasing s , the similarity between prey (decreasing prey switching), leads to increasing rates of predation seen as the bulk of the densities moving closer to the origin from the left-hand column to the right-hand column. Increasing predation intensity has a similar effect of increasing the variance in community size while also increasing the covariance between prey populations and decreasing the mean abundance of prey.

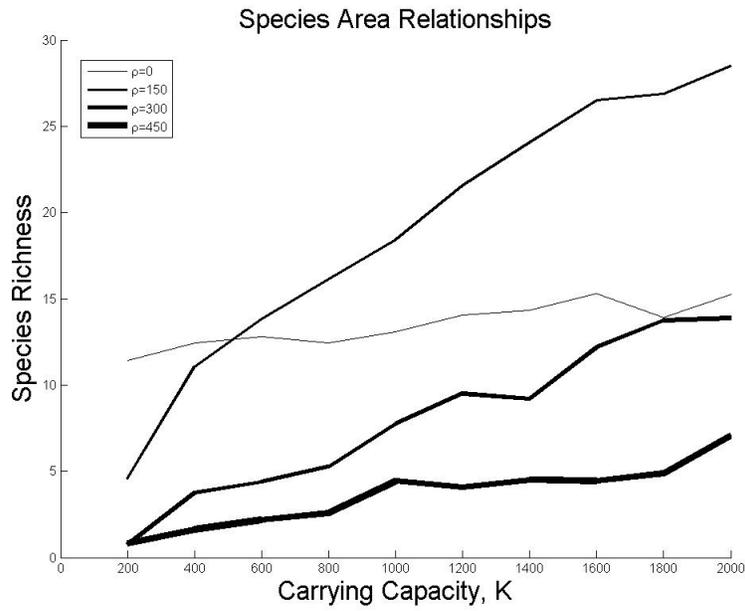


Figure S10: Cross-sections along lines of constant ρ in Figures 1-3 capture species-area relationships; prey-switching predation increases the slope of SARs as plotted here for the average of 4 replicates sampled at 4 time points, $T \in [5000, 10,000, 15,000, 20,000]$ and $z = 1.5$. K^* is the value of K where SARs cross, which requires that the species richness increase faster with K at $\rho = \rho^*$ than at $\rho = 0$.

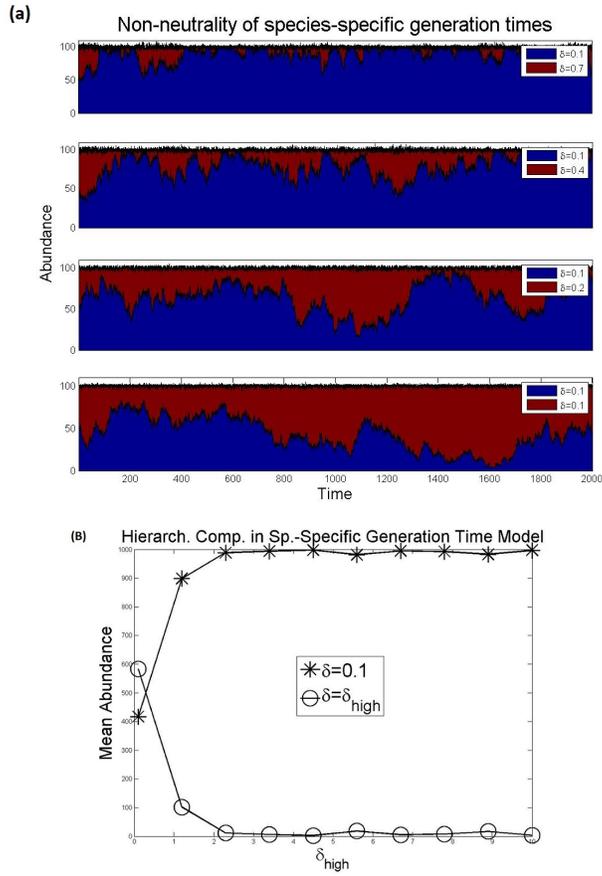


Figure S11: Species-specific generation times yield non-neutral prey communities. (a) Stacked plots of 2-species predator-free communities run for $T=2000$ shows that the species with longer generation times have higher mean abundances. When $\delta_1 = \delta_2$, the communities are neutral (bottom panel of trajectory plots) but as $|\delta_1 - \delta_2|$ increases the competitive superiority of the longer-lived species becomes apparent in a decreased mean abundance of the shorter-lived species (b).

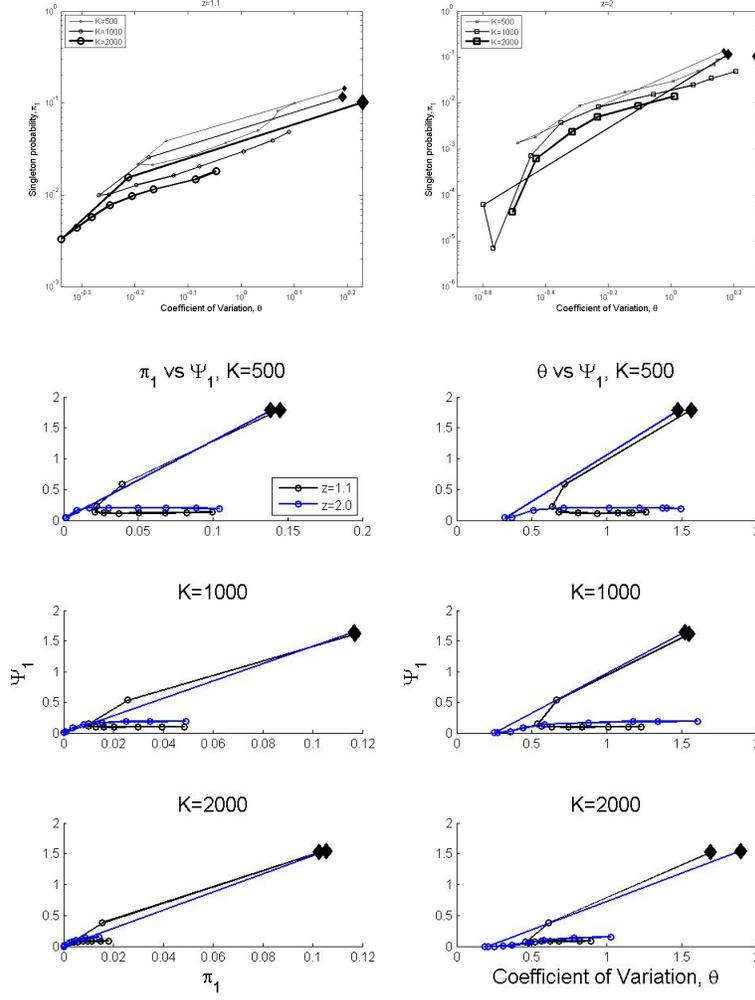


Figure S12: θ is a good predictor of π_1 (top panel), but both are poor predictors of Ψ_1 (bottom plots). Lines follow paths of increasing ρ for $\rho \in [0, 350]$, with diamonds indicating the end where $\rho = 0$.

5 Appendix V: Supplementary analysis

Changes of singleton death propensity n_1^- with K

A decrease in n_1^- with K comes from the decrease in the apostatic death function. Recall

$$n_1^- = \delta + \rho \frac{\sum_j N_j}{K} \frac{P}{1 + \sum_j N_j^z}. \quad (\text{S19})$$

A formal proof that $E[n_1^-]$ decreases with parameters will require integrating over the currently unknown distribution of community composition, but informal intuition about changes in n_1^- can be obtained by looking at its functional form near the expected community composition. Substituting in the mean-values of N_j and P for the predator with a bilinear functional form of predation, we get \tilde{n}_1^- , an approximation for n_1^-

$$\begin{aligned}\tilde{n}_1^- &= \delta + r \frac{\gamma}{\alpha\rho} \left(1 - \frac{\gamma}{\alpha\rho}\right) \frac{K}{1 + M \left(\frac{K\gamma}{\alpha\rho M}\right)^z} \\ &\approx \delta + h(\rho)M^{z-1}K^{1-z}\end{aligned}\tag{S20}$$

where $h(\rho) = r \left(\frac{\gamma}{\alpha\rho}\right)^{1-z} \left(1 - \frac{\gamma}{\alpha\rho}\right)$. Under this approximation, we would expect that n_1^- decreases with K , decreases with z , increases with M , and increases with ρ so long as $z > 1$. All of these changes resonate with intuition: singleton invaders should be less likely to be eaten by a prey-switching predator if they enter into large communities with fewer species (thereby decreasing the abundance of the invader relative to other species) and a predator with low-attack rate that highly favors the more abundant species.

This calculation points to a useful quantity in the analysis of this system: $\frac{\alpha\rho}{\gamma}$, which is the expected number of offspring, P_0 , produced by a single predator invading a community saturated with K prey. This quantity is also the fold-decrease in prey populations observed following predator introduction, an empirically tractable quantity that we used to estimate the parameter ρ . We define $P_0 = \frac{\alpha\rho}{\gamma}$ and rewrite equation S20

$$\tilde{n}_1^- \approx \delta + rK^{1-z} (MP_0)^{z-1} (1 - P_0^{-1}).\tag{S21}$$

We see that for $z > 1$, \tilde{n}_1^- is monotonically increasing with P_0 , and not just ρ .

Investigating the number of singletons Ψ_1

One may be able to compute Ψ_1 through an iterative scheme but often such schemes provide little intuition about Ψ_1 and therefore leave us without rules-of-thumb about how/why a prey-switching predator reduces the number of singletons and how that scales with ρ and z (e.g. Nisbet and Gurney, 1982). We attempted to find such rules of thumb by considering the probability that a species is a singleton given that it is present in the community. We denote this probability π_1 . If migrations are zero and if the propensity of singleton deaths were constant, π_1 is the quasistationary probability that an extant species is a singleton. This quantity would be inversely related to the mean time to extinction, τ_E (Nisbet and Gurney, 1982):

$$\tau_E = \frac{1}{n_1^- \pi_1}.\tag{S22}$$

In seeking an approximation for π_1 , we hypothesized that migration has a negligible effect in our model, that n_1^- approximately constant, and that we

could search for a reasonable approximation to π_1 based on some function of descriptive statistics of the distribution of prey population size such as the mean, μ_N and variance, σ_N^2 , conditioned on the prey species being present in the community. In particular, decreases in the mean and increases in the variance might be tied to increases in π_1 . We developed these hypotheses from the change in the probability density at $x = 1$ with changes in mean and variance of a unimodal distribution with mean greater than 1, small relative variance, and negligible skewness.

We explored various combinations of the mean and variances (e.g. the coefficient of variation $\theta = \frac{\sigma_N}{\mu_N}$) in search of a useful and empirically measurable quantity that predicts π_1 , and consequently extinction rates, thus yielding rules of thumb for the statistical signatures of prey communities that would benefit from prey-switching predation. Such a relationship between θ and extinction rates would open the door to incorporating known scaling laws between the mean & variance (van Kampen, 1981; Taylor and Woiwod, 1980; Kilpatrick and Ives, 2003) and might provide fruitful discussion about the potential scaling of a prey-switching predator's effect on prey diversity in natural systems.

However, although the coefficient of variation was a good approximation for π_1 and both changed non-monotonically with ρ in much the same way as the species richness, they provided a poor approximation of Ψ_1 due to the confounding and logically circular effect of the number of species in the community at stationarity. Conditioned on there being fewer species, μ_N and σ_N^2 should rise, and in fact we see for high ρ an increase in the coefficient of variation of species conditioned on their being present in the community (see Figure S12). It remains an open challenge to find heuristics and rules of thumb based on measurable properties of prey populations that can predict the effect of a predator on prey diversity given its attack rate and degree of prey switching.

Investigating $\text{Cov}[n_1^-, \Psi_1]$

Decreasing the variance in relative abundance of prey populations will change the distribution of the number of singletons at any given point of time. For example, consider a 2-species community. When the variance in relative abundance is non-zero and variance in community size zero, the multivariate prey density falls on the line $N_i + N_j = C$ and there can be only one singleton at any point in time. When a singleton is present in the community, $\Psi_1 = 0$ or $\Psi_1 = 1$ and thus n_1^- takes on its second-lowest value

$$n_1^- = \delta + \rho \frac{C}{K} \frac{P}{1 + (C-1)z}. \quad (\text{S23})$$

Therefore, when a singleton is present, $\text{Cov}[n_1^-, \Psi_1] < 0$ when σ_R^2 is large and σ_C^2 small. When the variance in relative abundance is zero, then $N_1 = N_2$ and the distribution of the number of singletons will fall entirely on either $\Psi_1 = 0$ or $\Psi_1 = 2$, causing n_1^- to take its highest value

$$n_1^- = \delta + \rho \frac{P}{K} \quad (\text{S24})$$

when both species are singletons.

This analysis is complicated in multi-species communities with intermediate degrees of variance along and orthogonal to the simplex, but we suspect that intuition from the 2-species case extends to multispecies communities. For a given community size, C , n_1^- is maximized when the community is perfectly even ($N_i = N_j = N$ for all i, j), and for communities that are perfectly even n_1^- increases as the community gets small. We hypothesize that in general n_1^- increases monotonically along some direction of increasing evenness and decreasing C . We hypothesize $\text{Cov}[n_1^-, \Psi_1]$ increases faster with ρ for higher z because the predator decreases the variance in relative abundance thereby increasing the variance of Ψ_1 , and this decrease in variance is greater for a more effective prey-switcher (higher z). When there are above-average numbers of singletons in a community with highly covarying populations, we expect the total community size to be low, and the expected evenness generated by a prey-switching predator will yield higher-than-average n_1^- . Increases in K will be tied to decreases in the variance of both n_1^- (due to the apostatic death function behaving like something between K^{-1} and K^{1-z} for $z > 1$; see equation S19) and Ψ_1 (by moving the entire multivariate density away from the boundaries where singletons occur), and thus increases in K should be associated with decreases in $\text{Cov}[n_1^-, \Psi_1]$.

Existence of ρ^* and K^*

We are now equipped to see how K modulates the effect of predation on species richness in our model. The existence of ρ^* is ensured by sufficiently large increases of $E[n_1^-]$, $E[\Psi_1]$, and $\text{Cov}[n_1^-, \Psi_1]$ with ρ beyond the point of maximal species richness. However, the existence of ρ^* does not automatically imply the existence of a critical carrying capacity K^* , above which a prey-switching predator improves species richness and below which suppresses species richness relative to a predator free community. In order for K^* to exist, species richness must increase more rapidly with K in the vicinity of $\rho = \rho^*$ than at $\rho = 0$.

To see why this occurs, we first note that for $z > 1$, n_1^- decreases with K for $\rho > 0$ but not for $\rho = 0$ since the singleton death rate in absence of predation is simply δ . As long as Ψ_1 and/or the covariance $\text{Cov}[n_1^-, \Psi_1]$ do not increase with K , this result ensures that some predation regimes will have opposite effects on richness depending on the value of K . In fact, in our models Ψ_1 and $\text{Cov}[n_1^-, \Psi_1]$ enhance the effect further.

In our models, variation in n_1^- is shallow over $K \in [500, 2000]$ and cannot account for the full effect (Figure S8). Instead, much of the effect can be attributed to the decreasing of the number of singletons Ψ_1 and the decreasing covariance $\text{Cov}[n_1^-, \Psi_1]$ with increasing K (Figure 5a). In the predator-free communities ($\rho = 0$), $\text{Cov}[n_1^-, \Psi_1] = 0$ since n_1^- is independent of the community state, and Ψ_1 varies little across the full range of carrying capacities considered. However, in the presence of prey-switching predators ($\rho > 0$ and $z > 1$), smaller communities have higher covariance $\text{Cov}[n_1^-, \Psi_1]$. This probably results from an increase in the between-prey covariance $\sigma_{i,j}$, which causes

singletons to occur when community sizes are below average and, consequently, n_1^- above average.

Variance in relative abundance as a measurable, heuristic tool for assessing impacts of stabilizing forces on species richness

Rates of prey-switching predation on a focal species increase as a function of that species' relative abundance, and thus fluctuations in relative abundance may be more informative than fluctuations in numerical abundance for understanding the impacts of a prey-switching predator (or other frequency-dependent stabilizing mechanism) on the statistics of the population sizes of its prey and community species richness. Fluctuations in community size with fixed relative abundance are visualized as spread of the multivariate prey distribution along the major axis of the ellipses pictured in Figure S9 (see also Figure 5b). Fluctuations in relative abundance in a fixed community size, on the other hand, are fluctuations on a simplex (the minor axis of the ellipses).

A Taylor expansion for the moments of the relative abundance yields an approximation for the variance in relative abundance. This allows us to appreciate the relationship between the relative abundance and statistics of the distribution of prey population size. For a function $f(X, Y)$ of random variables X and Y with well defined means and variance, its variance can be approximated by performing a Taylor expansion about μ_X and μ_Y , the means of X and Y , respectively

$$\begin{aligned} \text{Var}[f(X, Y)] &\approx \text{Var} \left[f(\mu_X, \mu_Y) + (X - \mu_X) \frac{\partial f}{\partial X} + (Y - \mu_Y) \frac{\partial f}{\partial Y} \right] \quad (\text{S25}) \\ &\approx \left(\frac{\partial f}{\partial X} \right)^2 \sigma_X^2 + \left(\frac{\partial f}{\partial Y} \right)^2 \sigma_Y^2 + 2 \left(\frac{\partial f}{\partial X} \right) \left(\frac{\partial f}{\partial Y} \right) \text{Cov}[X, Y] \end{aligned}$$

where the derivatives are evaluated at $(X, Y) = (\mu_X, \mu_Y)$. Performing this expansion for the relative abundance R of some population N in a community of size C ($R = N/C$), we obtain

$$\text{Var}[R(N, C)] \approx \frac{\sigma_N^2}{\mu_C^2} + \frac{\mu_N^2}{\mu_C^4} \sigma_C^2 - 2 \frac{\mu_N}{\mu_C^3} \text{Cov}[N, C] \quad (\text{S26})$$

where μ_N is the expected population size, μ_C the expected community size, σ_N^2 the variance of a prey species' population size, σ_C^2 the variance in community size, and $\text{Cov}[N, C]$ the covariance between a focal prey population and the community size. Assuming all M prey species are neutral, we have that $\mu_C = M\mu_N$, $\text{Cov}[N, C] = \sigma_N^2 + (M - 1)\sigma_{i,j}$, where $\sigma_{i,j}$ is the covariance between two populations. Substituting these into equation 25, we get an expression for the

variance in relative abundance, σ_R^2

$$\begin{aligned} \sigma_R^2 \approx & \frac{\sigma_N^2}{M^2 \mu_N^2} + \frac{1}{M^4 \mu_N^2} [M \sigma_N^2 + M(M-1) \sigma_{i,j}] \\ & - \frac{2}{M^3 \mu_N^2} (\sigma_N^2 + (M-1) \sigma_{i,j}). \end{aligned} \quad (\text{S27})$$

This simplifies to

$$\sigma_R^2 \approx \frac{M-1}{M^3} \left[\left(\frac{\sigma_N}{\mu_N} \right)^2 - \frac{\sigma_{i,j}}{\mu_N^2} \right] \quad (\text{S28})$$

or

$$\sigma_R^2 \propto \theta^2 - \zeta^2 \quad (\text{S29})$$

where $\theta = \frac{\sigma_N}{\mu_N}$ is the coefficient of variation and $\zeta = \frac{\sqrt{\sigma_{i,j}}}{\mu_N}$ the coefficient of covariation. From this approximation, we see that changes in the variance in relative abundance, the coefficient of variation, and the covariance are intimately tied. In the case where prey-switching is independent of overall predation rates (as in our models), prey-switching predators increase the covariance between prey populations and decrease the variance within prey populations. The overall effect is to decrease the variance in relative abundance.

It turns out that this decomposition of population fluctuations in terms of fluctuations in community size and fluctuations on simplices of constant community size may lead scale-free statistics of population fluctuations that are useful for understanding how stabilizing forces impact extinction risk. Both the coefficient of variation and the coefficient of covariation scale with K ; we can approximate them as $\theta \approx \bar{\theta} K^{-\frac{1}{2}}$ and $\zeta \approx \bar{\zeta} K^{-\frac{1}{2}}$ where $\bar{\theta}$ and $\bar{\zeta}$ do not depend on K . We define a re-scaled compositional variance $\bar{\sigma}_R^2 = \sigma_R^2 K^{-1}$, and we have

$$\bar{\sigma}_R^2 \propto \frac{1}{K^2} [\bar{\theta}^2 - \bar{\zeta}^2] \quad (\text{S30})$$

which combines information from the means, variances and covariances in prey populations to yield a useful metric of fluctuations in relative abundance - precisely the fluctuations that are dampened by many stabilizing mechanisms (including our prey-switching predators). We observe from our models that for a given value of z , $\bar{\sigma}_R^2$ is a good metric of extinction rates across a range of K and ρ (Figure S13). Note, however, that additional information about z is necessary to predict species richness from empirical measurements of $\bar{\sigma}_R^2$, because systems without prey-switching predators (the diamonds in Figure S13) or with different values of z fall on a different lines from one another.

We do not know the exact reason for the close relationship between $\bar{\sigma}_R^2$ and species richness, but it may be related to the relationship between σ_R^2 and the expected angular displacement of the community from its mean (Figure S14). Defining

$$\phi = 2 \cos^{-1} \left(\frac{\sum_i N_i}{\sqrt{\sum_i N_i^2}} \right)$$

we can re-write ϕ as a function of relative abundances, $R_i = \frac{N_i}{\sum_j N_j}$ and perform a Taylor expansion of ϕ about the mean relative abundances $E[\mathbf{R}]$ and note that all species are neutral to approximate

$$E[\phi] \approx M\sigma_R^2 \left[\frac{\partial^2}{\partial R_i^2} \cos^{-1} \left[\left(\sum_i R_i^2 \right)^{-\frac{1}{2}} \right] \right]_{E[\mathbf{R}]} + 2M(M-1)\text{Cov}[R_i, R_j] \left[\frac{\partial^2}{\partial R_i \partial R_j} \cos^{-1} \left[\left(\sum_i R_i^2 \right)^{-\frac{1}{2}} \right] \right]_{E[\mathbf{R}]}$$

and thus

$$E[\phi] \propto \sigma_R^2 + a\text{Cov}[R_i, R_j]$$

where a is some constant. In the special case of a 2-species community (or any grouping of a community into two groups that are equivalent both in the eyes of the researcher and of the predator) we have relative abundances R_1 and $R_2 = 1 - R_1$, and $\text{Cov}[R_i, R_j] = -\sigma_R^2$ thus our approximations imply a direct proportionality between the expected angular deviation and the variance in relative abundance $E[\phi] \propto \sigma_R^2$ (Figure S14). Intuition for the utility of the variance in relative abundance towards making scale-free statistic for estimating extinction rates and equilibrium species richness comes from its proportional relationship between the angular deviation of a community from its mean, as these radial coordinates are closer to a scale-free indicator of the proximity of a community to the edges where extinctions occur. More thorough analysis is warranted, as are tests of whether or not the statistics σ_R^2 , $\bar{\sigma}_R^2$ and ϕ are useful in asymmetric communities (perhaps with re-scaled axes to ensure the central tendency of the rescaled community is centered in the positive orthant).

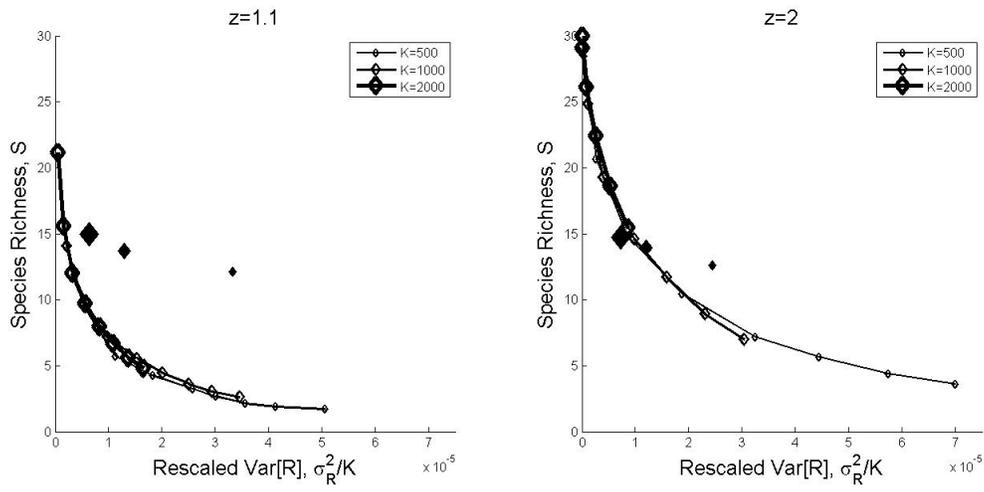


Figure S13: Species richness vs. σ_R^2 (equation S30) plotted for $K \in \{500, 1000, 2000\}$, along lines of increasing $\rho \in [0, 350]$. σ_R^2 may be a useful proxy for extinction risk given a particular mechanism of stabilization. More information is required to predict species richness for different stabilizing mechanisms, as systems without prey-switching predators (solid diamonds) or systems with different degrees of prey-switching stabilization (the two subplots) appear to fall on significantly different curves.

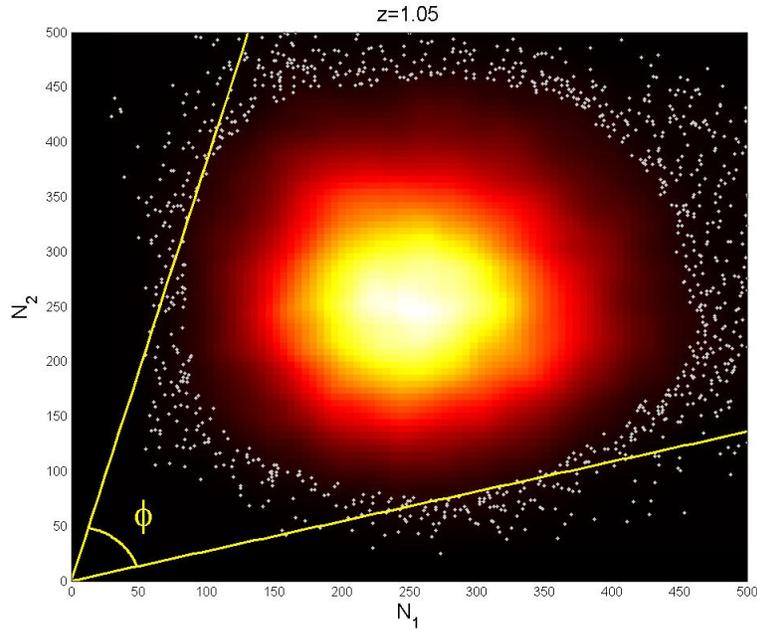


Figure S14: The expected angular displacement of a community from its mean, ϕ , is linearly related to the variance in relative abundance. Since the variance in relative abundance, σ_R^2 , is dampened by frequency-dependent stabilizing mechanisms, ϕ and σ_R^2 may lead to useful scale-free statistics for assessing extinction risk in finite communities with frequency-dependent stabilization such as the prey-switching predators we've considered.

6 Appendix VI: Guide to MATLAB simulation code

We provide all of our MATLAB code for simulation in the online supplement. The logic of the simulations is described in detail both in the main paper and in annotations embedded within the attached MATLAB files. Some minor differences exist between the variable and parameter names used in our paper and those used in the code. Where they involve parameters, these differences are noted in Table S1. Where they involve variables, these differences are noted in the annotated code.

The MATLAB file `PredSim.m`, available with the online supplementary materials, defines a function `PredSim` that takes as inputs the control parameters in our models, as well as a text string that specifies the functional form of predation to be used. `PredSim` provides as outputs snapshots of the predator/prey community after simulations have run for a user-specified number of timesteps. The file `PredSim_Demo.m` contains the parameter values used in our simula-

tions, and a few short analyses that call PredSim and provide outputs.

References

- Avens, L., J. C. Taylor, L. R. Goshe, T. T. Jones, and M. Hastings. 2009. Use of skeletochronological analysis to estimate the age of leatherback sea turtles *Dermochelys coriacea* in the western North Atlantic. *Endangered Species Research* 8:165–177.
- Barnes, C., D. Maxwell, D. C. Reuman, and S. Jennings. 2010. Global patterns in predator-prey size relationships reveal size dependency of trophic transfer efficiency. *Ecology* 91:222–232.
- Brose, U., T. Jonsson, E. L. Berlow, P. Warren, C. Banasek-Richter, L.-F. Bersier, J. L. Blanchard, T. Brey, S. R. Carpenter, and M.-F. C. Blandinier. 2006. Consumer-resource body-size relationships in natural food webs. *Ecology* 87:2411–2417.
- Brown, C. J. D., and R. C. Ball. 1943. A fish population study of third sister lake. *Transactions of the American Fisheries Society* 72:177–186.
- Calder, W. A. 1990. Avian longevity and aging. Pages 185–204 *in* D. E. Harrison, ed. *Genetic Effects on Aging*. The Telford Press, Caldwell, NJ.
- Calow, P. 1977. Conversion efficiencies in heterotrophic organisms. *Biological Reviews* 52:385–409.
- Clapp, R. B., and M. K. Klimkiewicz. 1982. Longevity records of North American birds: Gaviidae through Alcidae. *Journal of Field Ornithology* 53:81–124.
- Clarke, A. 1976. Some observations on krill (*Euphausia superba* Dana) maintained alive in the laboratory. *Br. Antarct. Surv. Bull* 43:111–118.
- Condit, R., P. S. Ashton, N. Manokaran, J. V. LaFrankie, S. Hubbell, and R. B. Foster. 1999. Dynamics of the forest communities at Pasoh and Barro Colorado: comparing two 50-ha plots. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 354:1739–1748.
- González-Romero, A., A. Ortega, and R. Barbault. 1989. Habitat partitioning and spatial organization in a lizard community of the Sonoran desert, Mexico. *Amphibia-Reptilia* 10:1–11.
- Härkönen, T., and M. P. Heide-Jørgensen. 1990. Comparative life histories of East Atlantic and other harbour seal populations. *Ophelia* 32:211–235.
- Holmes, R. T., T. W. Sherry, and F. W. Sturges. 1986. Bird community dynamics in a temperate deciduous forest: long-term trends at Hubbard Brook. *Ecological Monographs* 56:201–220.

- Honěk, A. 1993. Intraspecific variation in body size and fecundity in insects: a general relationship. *Oikos* pages 483–492.
- Jennings, S., K. J. Warr, and S. Mackinson. 2002. Use of size-based production and stable isotope analyses to predict trophic transfer efficiencies and predator-prey body mass ratios in food webs. *Marine Ecology Progress Series* 240:11–20.
- Jordan, D. S. 1905. A guide to the study of fishes, vol. 1. Henry Holt and Company, New York.
- Kilpatrick, A., and A. Ives. 2003. Species interactions can explain Taylor's power law for ecological time series. *Nature* 422:65–68.
- Lawton, J. H., M. P. Hassell, and J. R. Beddington. 1975. Prey death rates and rate of increase of arthropod predator populations. *Nature* 255:60–62.
- Li, C. H., B. L. Ma, and T. Q. Zhang. 2002. Soil bulk density effects on soil microbial populations and enzyme activities during the growth of maize (*Zea mays L.*) planted in large pots under field exposure. *Canadian Journal of Soil Science* 82:147–154.
- MacGinitie, G. E. 1934. The egg-laying activities of the sea hare, *Tethys californicus* (Cooper). *Biological Bulletin* pages 300–303.
- Madsen, T., and R. Shine. 1994. Components of lifetime reproductive success in adders, *Vipera berus*. *The Journal of Animal Ecology* pages 561–568.
- Martin, T. E. 1995. Avian life history evolution in relation to nest sites, nest predation, and food. *Ecological Monographs* 65:101–127.
- Morozov, A., and S. Petrovskii. 2013. Feeding on multiple sources: Towards a universal parameterization of the functional response of a generalist predator allowing for switching. *PloS one* 8:e74586.
- Nisbet, R. M., and W. S. C. Gurney. 1982. Modelling fluctuating populations. Wiley Chichester ; New York.
- Price, E. R., B. P. Wallace, R. D. Reina, J. R. Spotila, F. V. Paladino, R. Piedra, and E. Vélez. 2004. Size, growth, and reproductive output of adult female leatherback turtles *Dermochelys coriacea*. *Endangered Species Research* 5:1–8.
- Rohwer, F. C. 1988. Inter-and intraspecific relationships between egg size and clutch size in waterfowl. *The Auk* pages 161–176.
- Sacher, G. A. 1959. Relation of Lifespan to Brain Weight and Body Weight in Mammals. Pages 115–141 *in* G. Wolstenholme and M. O'Conner, eds. *Ciba Foundation Symposium - The Lifespan of Animals (Colloquia on Ageing)*. John Wiley & Sons, Ltd, Chichester, UK.

- Smout, S., C. Asseburg, J. Matthiopoulos, C. Fernández, S. Redpath, S. Thirgood, and J. Harwood. 2010. The functional response of a generalist predator. *PLoS One* 5:e10761.
- Smout, S. C., and U. Lindstrom. 2007. Multispecies functional response of the minke whale *Balaenoptera acutorostrata* based on small-scale foraging studies. *Marine Ecology Progress Series* .
- Spiller, D. A., and T. W. Schoener. 1998. Lizards reduce spider species richness by excluding rare species. *Ecology* 79:503–516.
- Stroock, D. W., and S. S. Varadhan. 1979. *Multidimensional diffusion processes*, vol. 233. Springer Science & Business Media.
- Taylor, L., and I. Woiwod. 1980. Temporal stability as a density-dependent species characteristic. *The Journal of Animal Ecology* pages 209–224.
- Terborgh, J., L. Lopez, P. Núñez, M. Rao, G. Shahabuddin, G. Orihuela, M. Riveros, R. Ascanio, G. H. Adler, T. D. Lambert, and L. Balbas. 2001. Ecological meltdown in predator-free forest fragments. *Science* 294:1923–1926.
- Terborgh, J., S. K. Robinson, T. A. Parker, III, C. A. Munn, and N. Pierpont. 1990. Structure and organization of an amazonian forest bird community. *Ecological Monographs* 60:213–238.
- Timon, V. M., and E. J. Eisen. 1970. Comparisons of ad libitum and restricted feeding of mice selected and unselected for postweaning gain. I. Growth, feed consumption and feed efficiency. *Genetics* 64:41.
- Tomilin, A. G. 1959. *Mammals of USSR and adjacent countries*, vol. 9. Izdatel'stvo Akademi Nauk SSSR.
- Turchin, P., and R. S. Ostfeld. 1997. Effects of density and season on the population rate of change in the meadow vole. *Oikos* pages 355–361.
- van Kampen, N. G. 1981. *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam ; New York : New York.

**Chapter 3: Using Neutral Theory
(Wright-Fisher Process) As a Null Model
for Ecological and Economic Time-Series**

Alex Washburne

Introduction

Economies, ecosystems, and other complex-adaptive systems [11, 17, 1] change with time as the relative abundances of various companies, species, or strategies go up and down by a mix of forces we can predict and whose effects we understand and forces we can't predict and/or whose effects we don't understand. Reliable forecasts of these systems has implications for asset pricing, investment, and the management of socio-ecological systems.

A large subset of complex-adaptive systems are constrained by finite, limiting resources that may create an approximately zero-sum game for the agents (species, companies etc.) in competition over those resources. Examples are the cover of canopy trees on an island [12], the land owned by various groups (e.g. Sunni, Shiite, Kurds & others in Iraq) or the market share of companies in constant markets. For these zero-sum communities, the Wright-Fisher Process (WFP) serves as a mechanistically-based null model for time-series data against which one can look for competitively superior trees, groups or companies, or more complicated features of the data such as covariances between groups, clustering of different groups, percent variance covered by the first principal component, etc. Rejection of the WFP can motivate further construction of more detailed models for the stochastic evolution of complex-adaptive systems.

The Wright-Fisher Process is an Ito stochastic differential equation (SDE) originally proposed as a model for the neutral drift of alleles in a population [10, 24]. It was later adapted by Kimura [16] as a null model for the distribution of the frequency of molecular polymorphisms in a population, and extended by Hubbell [12] for species-abundance distributions and species-area relationships. It has been shown that models of volatility-stabilized markets [9] have market weights that follow the WFP [20]. The Wright-Fisher Process has also been proposed as an approximation to the Langevin SDE for the random opening and closing of ion channels in neurons [6]. Tests of neutrality have become cornerstone in population genetics [22, 7] and making neutrality tests available for time-series datasets may accelerate our understanding of non-neutral patterns in ecological, economic, sociological and other systems for which the WFP is a suitable null model.

The broad use of the Wright-Fisher process stems from its parsimonious assumptions about the underlying mechanisms of random birth, death, and dispersal. At each timestep, an organism is drawn at random to die. Immediately after death, another organism is drawn at random to replace the organism that died. With probability $1 - m$, an organism is randomly drawn from the local community to reproduce otherwise, with probability m , an organism is randomly drawn from an infinitely large metacommunity to migrate in, thereby maintaining a constant community size. This serves as a useful null model of neutral competition as all species have the same per-capita fitness - there are no niche differences, frequency-dependent stabilizing mechanisms, or other trophic structure stabilizing competition. For time-series datasets of communities of constant size, one might want to make a solid epistemological statement that a feature observed in their model is not explained by neutral birth, death and

dispersal. Features of interest could include the covariance between two species, the percent of variance explained by the 3rd principal component, or some feature of the Fourier spectrum and other complex features for which the null distributions of test statistics under neutral competition are not known or for which statistical null hypotheses detached from ecological mechanisms are less relevant and less conducive to the construction of alternative models.

The combined broad applicability of the Wright-Fisher process and recent emergence of long-term ecological research (LTER) datasets and other longitudinal studies of fast-changing communities such as microbial communities on the surface of the skin or tongue [2] motivates making the WFP of neutral competition accessible for hypothesis testing in time-series datasets. Some recent work has tested some analytically tractable features of populations within a neutral community [3, 15], but neutrality is not a population-level phenomenon: species are neutral with respect to one another, and so community-based features such as covariances and other more global structures of the data may illuminate general stabilizing mechanisms or non-neutral trophic structure in communities. The null distributions of such features of interest may be beyond the scope of current analytical tools; there is a need for the ability to simulate surrogate datasets of neutral drift against which features of time-series data can be compared. However, simulation of the WFP is fraught with pitfalls in numerical stability and computational speed; a fast, numerically stable method for accurate simulation of the WFP would greatly enable hypothetico-deductive progress in our understanding of the mechanisms underlying the dynamics of complex-adaptive systems.

In this chapter, I present a series of tools aimed at making the WFP accessible as a null model for time-series datasets. The first tool presented is a set of volatility-stabilizing transformations which can be combined with constant-volatility tests to test the quadratic covariation of the WFP; these multiple volatility-stabilizing transformations can be combined in a multiple-hypothesis-testing framework for a global test of whether an entire community is drifting according to a WFP. Second, I develop a numerically stable and computationally efficient scheme for simulating the WFP founded on two key calculations: I find a numerically stable decomposition of the covariation matrix which ensures a conservation of relative abundance and real-valued output, and I use the invertible log-odds-ratio transform to simulate a WFP on an unbounded state space which helps ensure trajectories don't hop out of bounds. Third, I provide calculations for parameter estimation so one can generate surrogate WFP datasets corresponding to their data. Finally, to illustrate the utility of the tools developed here, I analyze datasets from the human microbiome [2] and the S&P 500 and illustrate some ways forward towards the construction of alternative models that explain non-neutral phenomena in the data. I'm happy to make my MATLAB code available to accelerate progress on inferring non-neutral features in the dynamics of complex-adaptive systems.

The Model

The Wright-Fisher Process is derived as a continuous-state approximation of an urn process where balls are drawn out at random for death and then instantaneously replaced with either a birth from a randomly drawn member of the urn or a migration from a randomly drawn member of an external, infinitely large metacommunity [10, 24]. For a community of size J , this translates into the following transition probabilities for a population of size N_t^i at time t

$$\begin{aligned} P\{N_t^i \rightarrow N_t^i + 1\} &= \left(\frac{J - N_t^i}{J}\right) \left(mp + (1 - m)\frac{N_t^i}{J - 1}\right) \\ P\{N_t^i \rightarrow N_t^i - 1\} &= \left(\frac{N_t^i}{J}\right) \left(m(1 - p) + (1 - m)\frac{J - N_t^i}{J - 1}\right) \\ P\{N_t^i \rightarrow N_t^i\} &= \left(\frac{J - N_t^i}{J}\right) \left(m(1 - p) + (1 - m)\frac{J - 1 - N_t^i}{J - 1}\right) \\ &\quad + \left(\frac{N_t^i}{J}\right) \left(mp + (1 - m)\frac{N_t^i - 1}{J - 1}\right) \end{aligned}$$

where $P\{N_t^i \rightarrow N_t^i + 1\}$ is the probability that the population size of species i increases by 1 individual at time step t , m is the probability of migration and p_i the metacommunity relative abundance of the species. The defining assumptions of this one-step process are the random birth/death/migration and the constant community size creating zero-sum neutral drift.

The discrete process can be simulated directly. However, many systems are extremely large - there are 10^{14} bacteria in the gut, and $10^{12} - 10^{14}$ dollars in some markets - and the simulation of WFP surrogate datasets for these systems requires either computationally intensive direct simulation of the discrete process or more computationally efficient simulation of a similar, continuous process. This continuous process is obtained by considering relative abundances, $X_t^i = \frac{N_t^i}{C}$. In the limit of infinitely large system size [23] or by setting up a martingale problem and keeping constant the product of the probability of migration and the community size [21] one arrives at the Wright-Fisher SDE describing the trajectories of the relative abundances of all n species in the community:

$$d\mathbf{X}_t = \lambda(\mathbf{p} - \mathbf{X}_t) dt + \sigma(\mathbf{X}_t) d\mathbf{W}_t \quad (1)$$

where $\mathbf{X}_t \in \Delta^n \forall t$ is the vector of relative abundances at time t lying in the $n - 1$ dimensional simplex, Δ^n , in the positive orthant of \mathbb{R}^n , $\mathbf{p} \in \Delta^n$ is the relative abundance of the n species in the metacommunity, $\lambda = mJ$ is the rate of migration from the metacommunity into the local community, and $\frac{1}{2}\sigma\sigma^T = \Sigma$ is the covariation matrix where

$$\Sigma_{ij} = \begin{cases} X_t^i(1 - X_t^i) & i = j \\ -X_t^i X_t^j & i \neq j. \end{cases} \quad (2)$$

One can verify that the process stays in the simplex by confirming that $d\mathbf{X}_t^T \mathbf{1} = 0$.

Variance-Stabilizing Transformation

One promising avenue for hypothesis testing of the WFP is to find certain features of the process in equation 1 which are invariant for the WFP but need not be for other processes, allowing one to identify processes that are not well described by the WFP. Since the mean-reverting drift of the WFP, $\lambda(\mathbf{p} - \mathbf{X}_t) dt$, is linear, any process fluctuating about a stable node would have qualitatively similar drift near its equilibrium. One could conceivably test for whether different variables, X_t^i , have different rates of reversion towards their equilibrium, λ^i , by noting their autocorrelation time is $\tau^i = \frac{1}{\lambda^i}$. Alternatively, one could look for systems reverting to a spiral node by looking for periodic signals arising in the power spectrum or the autocovariance function.

Another approach to identifying non-neutral features of data would be to look at global features of the covariance structure and how they differ from that defined by Σ . There are many possibilities, but here we consider just one: a volatility stabilizing transformation, $g(\mathbf{X}_t)$ such that the volatility of g is constant, $\mathbb{V}[dg|\delta t] = \delta t$. The SDE for g will be

$$dg_t = \left(\lambda \nabla g^T (\mathbf{p} - \mathbf{X}_t) + \frac{1}{2} \nabla g^T H_{\mathbf{X}} \nabla g \right) dt + \nabla g^T \sigma(\mathbf{X}_t) d\mathbf{W}_t \quad (3)$$

and thus the volatility over some time, δt , will be

$$\mathbb{V}[dg_t] = \delta t (\nabla g^T \sigma \sigma^T \nabla g) \quad (4)$$

. Substituting $\sigma \sigma^T = 2\Sigma$ and setting the volatility $\mathbb{V}[dg_t] = \delta t$, we get that a constant-volatility transformation, g , will satisfy

$$\nabla g^T \Sigma \nabla g = \frac{1}{2}. \quad (5)$$

Using equation 2, this can be expanded as

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^n X^i X^j \left(\frac{\partial g}{\partial X^i} - \frac{\partial g}{\partial X^j} \right)^2 = \frac{1}{2} \quad (6)$$

In the 2-D case, this equation is solved by setting $X^1 = x$ and $X^2 = y$ (for aesthetic purposes), yielding

$$xy \left(\frac{\partial g}{\partial x} - \frac{\partial g}{\partial y} \right)^2 = \frac{1}{2} \quad (7)$$

and then making the substitution $w = x - y$ and using $y = 1 - x$ to obtain

$$\frac{\partial g}{\partial w} = \frac{1}{\sqrt{2(1-w^2)}} \quad (8)$$

which is solved by

$$g = \frac{1}{\sqrt{2}} \sin^{-1}(x - y) \quad (9)$$

. An empiricist worried about reproducibility might prefer a symmetric function that doesn't require choosing which species is 'x' and which species is 'y'. This can be obtained by noting that $(x-y)^2 + 4xy = 1$ and so the symmetric function

$$g_{sym} = \frac{1}{\sqrt{2}} \sin^{-1} \sqrt{4xy} \quad (10)$$

is also a solution to equation 7 as both g and g_{sym} can be interpreted as the non-right angles of a right triangle with sides of length $x-y$ and $\sqrt{4xy}$, and, since $x+y=1$, a hypotenuse of length 1. Thus, since $g = \frac{\pi}{2} - g_{sym}$, if one angle has constant volatility then so must the other.

By setting $X^3 = z$, the $n=3$ case for a constant-volatility transformation can be written beautifully

$$xy \left(\frac{\partial g}{\partial x} - \frac{\partial g}{\partial y} \right)^2 + xz \left(\frac{\partial g}{\partial x} - \frac{\partial g}{\partial z} \right)^2 + yz \left(\frac{\partial g}{\partial y} - \frac{\partial g}{\partial z} \right)^2 = const. \quad (11)$$

We see that the conditions for the constant-volatility transformation are always a special case of the Hamilton-Jacobi equation,

$$H(\mathbf{x}, \nabla f(\mathbf{x})) = 0$$

whose solutions can be found by integrating the Hamilton equations

$$\begin{aligned} \dot{x}_i &= \frac{\partial H}{\partial q_i} \\ \dot{q}_i &= -\frac{\partial H}{\partial x_i} \end{aligned}$$

for $i = 1, \dots, n$ where $q_i = \frac{\partial g}{\partial x_i}$. This connection to Hamilton-Jacobi theory is not utilized in this chapter, but Burby & Washburne (in prep) have used this approach to show that the equation defines a metric on the simplex with constant Gaussian curvature, allowing one to find Killing vectors that enable construction of a mapping of the simplex to the sphere that serves as a constant-volatility transformation. For instance, note that for the 2-D case setting $u = \sqrt{x}$ and $v = \sqrt{y}$, we have that $u^2 + v^2 = 1$ - a mapping from the 2-simplex to the 2-sphere that then enables construction of a constant-volatility transformation $\sin^{-1} \sqrt{4xy}$. In Burby & Washburne (in prep), we found that there are uncountably many constant-volatility transformations. Here, I present a small, intuitively useful subset of those solutions that were found without the Hamilton-Jacobi theory.

Solution to equation 6: 2^n volatility-stabilizing transformations for the WFP

One set of solutions to equation 6 is

$$g(\mathbf{X}_t) = \sin^{-1} \left(\sum_{i=1}^n a_i X_t^i \right) \quad (12)$$

where $a_i = \pm 1$. The coefficients $\{a_i\}_{i=1}^n$ define a “grouping” of species by pooling the n species into two groups, those with $a_i = 1$ and those with $a_i = -1$, and thus projects the n species WFP into a 2 species WFP and exploits the solution given above in equation 9 (similarly, equation 10 can be used to define another complimentary set of solutions). This defines 2^n non-superimposable solutions, all of which have constant volatility, $\lim_{\delta t \rightarrow 0} \frac{\mathbb{V}[g_{t+\delta t} - g_t]}{\delta t} = \text{const}$. Defining $\Delta g = g_{t+\delta t} - g_t$, we can then test an arbitrary time series dataset against a null hypothesis of neutral drift by testing heteroskedasticity of Δg versus g , \mathbf{X}_t , t , and other state variables of the system. The utility of this method is illustrated below in the section “Application to Real Data”.

Simulation of WFP

Many scientists confronted with a time-series of a complex adaptive system may want to explore beyond a few simple, analytically tractable features of the WFP such as the percent of variance accounted for by the n th principal component, the distance between two clusters in a diffusion map, the cluster validity index for a particular clustering algorithm, or myriad other features. To enable these scientists to have a baseline for comparison, it would be helpful to simulate surrogate datasets of a WFP such that the null distributions of the scientists’ feature of choice can be obtained from simulation. Two obstacles prevent the trivial simulation of the WFP. First, derivation of a matrix σ which is numerically stable (i.e. $\sigma \in \mathbb{R}^{n \times d}$ and $\mathbf{1}^T \sigma = \mathbf{0}$) and computationally efficient - simply using numerical methods for computing the Cholesky or eigen-decomposition of Σ yield complex-valued matrices σ and do not reliably satisfy $\mathbf{1}^T \sigma = \mathbf{0}$. Second, ensuring that the simulations remain within the state space, Δ^n . In our paper, we overcome the first obstacle with a decomposition of Σ which works for a large class of what we call “zero-sum matrices”. We overcome the second obstacle by simulating the log-odds-ratio transform of \mathbf{X}_t which maps Δ^n into an unbounded manifold in \mathbb{R}^n and which can be inverted to obtain trajectories of the Wright-Fisher process.

Decomposition of Covariation Matrix

We note that the covariation matrix, Σ , can be decomposed as $\Sigma = PVP^T$ where P is what we call a “zero-sum matrix” which ensures $\mathbf{1}^T \Sigma = \mathbf{0}$, and V is the “competition matrix” - a diagonal matrix whose diagonal elements denote the nature of competition in a zero-sum community; in the case of the WFP the diagonal elements of V are all possible unique pairwise combinations of $X_t^i X_t^j$. For instance, let $n = 3$ and for easier notation set $X_t^1 = x$, $X_t^2 = y$, and $X_t^3 = z$,

the decomposition is

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} xy & & \\ & xz & \\ & & yz \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} x(1-x) & -xy & -xz \\ -xy & y(1-y) & -yz \\ -xz & -yz & z(1-z) \end{pmatrix} \quad (13)$$

For $n = 4$ we use the same notation as above in addition to $X_t^4 = w$ to write the decomposition as

$$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} xy & & & & & \\ & xz & & & & \\ & & xw & & & \\ & & & yz & & \\ & & & & yw & \\ & & & & & zw \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} x(1-x) & -xy & -xz & -xw \\ -xy & y(1-y) & -yz & -yw \\ -xz & -yz & z(1-z) & -zw \\ -xw & -yw & -zw & w(1-w) \end{pmatrix} \quad (14)$$

which makes it easy to see the generalization to n -species communities. For an n -species community, $P \in \mathbb{R}^{n \times M}$ and $V \in \mathbb{R}^{M \times M}$ where $M = \frac{n(n-1)}{2}$. P can be constructed in blocks as

$$P = \begin{pmatrix} \mathbf{1}_{n-1}^T & \mathbf{0}_{n-2}^T & \cdots & 0 \\ & \mathbf{1}_{n-2}^T & \ddots & \vdots \\ & & \ddots & 0 \\ & & & 1 \\ -I_{n-1} & -I_{n-2} & \cdots & -1 \end{pmatrix} \quad (15)$$

where $\mathbf{1}_n \in \mathbb{R}^n$ is the vector of all 1's, $\mathbf{0}_n \in \mathbb{R}^n$ is the vector of all 0's, and $I_n \in \mathbb{R}^{n \times n}$ is the identity matrix. V is a diagonal matrix whose diagonal is formed by vectorizing, row-by-row, the upper-triangular elements of $\mathbf{X}_t \mathbf{X}_t^T$.

This factorization allows us to compute σ directly through $\sigma = \sqrt{2}PV^{1/2}$ and thereby guarantee real-valued noise that could not be guaranteed by numerical Cholesky or eigen-decomposition of Σ . Furthermore, analytically the Wright-Fisher Process should have $\mathbf{1}^T \mathbf{X}_t = 1 \forall t$ and the matrix P , which clearly has $\mathbf{1}^T P = \mathbf{0}$, may reduce the numerical error causing increases or decreases in the total relative abundance.

We note that this factorization may be useful for other zero-sum covariance matrices where $\mathbf{1}^T \Sigma = \mathbf{0}$. In particular, this factorization works for any covariance matrix of simplex-valued processes with non-positive quadratic covariation,

$d[X_t^i, X_t^j] \leq 0$ for $i \neq j$. This zero-sum decomposition of the covariance matrix decomposes the volatility of the WFP in a meaningful way, where the columns of P containing a 1 and a -1 show that at every time step, δt , the transfer of relative abundance (or acquisition of market share) from species i to species j due to neutral drift is $\Delta_{ij} = \sqrt{2(\delta t) X_t^i X_t^j} Z$, where $Z \sim N(0, 1)$ is a standard normal random variable.

Log-odds-ratio transform

The next obstacle in simulating the WFP is the preservation of the boundary. Some work has been done on the boundary preservation of Wright-Fisher Processes: Milstein et al. [18] constructed a Balanced Implicit Method (BIM), Moro & Schurz [19] constructed a split-step method, and Dangerfield et al. [6] combined these two in a Balanced Implicit Split Step (BISS). The approach developed here seeks to use an invertible transformation to stretch the bounded state space of the WFP onto an unbounded state space, eliminating the need for the complex boundary-preserving numerical schemes.

For a process bounded on the simplex, we can use an invertible mapping of the simplex into an unbounded $n-1$ dimensional manifold in \mathbb{R}^n , simulating the process on the unbounded manifold, and then inverting to obtain our surrogate dataset. I define a new vector Ito process, \mathbf{f}_t , defined as the log-ratio transform of the WFP,

$$f_t^i = \log \left(\frac{X_t^i}{1 - X_t^i} \right). \quad (16)$$

Applying Ito's lemma, we get the SDE for \mathbf{f}_t

$$df_t^i = \left(e^{f_t^i/2} + e^{-f_t^i/2} \right)^2 \left[\left(\lambda p_i - \frac{1}{2} \right) + \frac{(1-\lambda)}{1 + e^{-f_t^i}} \right] dt + \nabla \mathbf{f}_t^T \sigma(\mathbf{f}_t) d\mathbf{W}_t \quad (17)$$

where $\sigma(\mathbf{f}_t)$ is $\sigma(\mathbf{X}_t) = PV^{\frac{1}{2}}$ converted to a function of \mathbf{f}_t by substituting $X_t^i = \left(1 + e^{-f_t^i} \right)^{-1}$ for all $i = 1, \dots, n$. However, the Euler-Maruyama integration and subsequent of inversion the process in equation 17 does not preserve the simplex (Figure 1). The warping of trajectories - and the final sum of relative abundances - is independent of the step-size of integration (Figure 1), suggesting that the error is not due to numerical error from the discretization of the continuous process. Interestingly, changing the number of species changes the degree of warping, and decreasing λ causes increased fluctuations about the long-term average of total relative abundance (Figure 1).

I found that incorporating a dimensionality correction, $-\frac{1}{2} \left(\frac{n-2}{n} \right) \left(e^{f_t^i/2} + e^{-f_t^i/2} \right)^2 dt$, improves the numerical stability of simulations in a reduced error of the total relative abundance, $\epsilon = \left| \sum_i X_t^i - 1 \right|$ without affecting the covariance matrix of

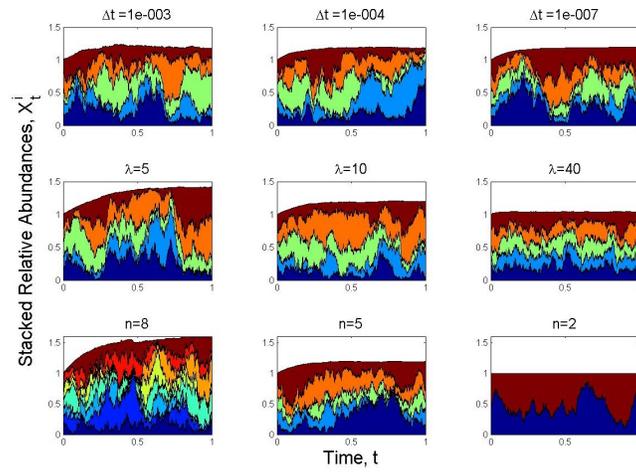


Figure 1: Stacked plots of the relative abundances, X_t^i , obtained by inverting numerical solutions of equation 17. The solutions leave the bounds of the simplex, $\sum_i X_t^i = 1$, not due to numerical error (first row), but instead due to some anomalous drift associated with the dimensionality of the system. Decreasing λ and increasing the number of species, n , increase the deviation of \mathbf{X}_t from the simplex. When $n = 2$, $\sum_i X_t^i \approx 1$, but when $n > 2$ $\sum_i X_t^i \gtrsim 1$ and increasingly so for decreasing λ . Unless the parameters are being varied as specified in the figure title, $n = 5$, $p = \frac{1}{n}$, $\lambda = 10$, and $\Delta t = 10^{-4}$.

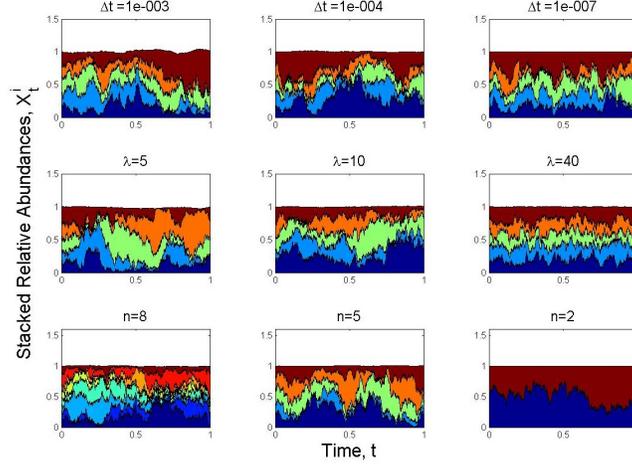


Figure 2: Adding a dimensionality correction to the drift of f_t^i decreases the error in total relative abundance, $\epsilon = |\sum_i X_t^i - 1|$. It is not yet proven that the solutions of the dimensionality-corrected process converge to solutions of the WFP.

\mathbf{X}_t obtained by inverting \mathbf{f}_t . (Figure 2). The SDE for f_t^i becomes

$$df_t^i = \left(e^{f_t^i/2} + e^{-f_t^i/2} \right)^2 \left[\left(\lambda p_i - \frac{1}{2} \right) - \frac{1}{2} \left(\frac{n-2}{n} \right) + \frac{(1-\lambda)}{1+e^{-f_t^i}} \right] dt + \nabla \mathbf{f}_t^T \sigma(\mathbf{f}_t) d\mathbf{W}_t \quad (18)$$

Parameter Estimation

The parameters λ and \mathbf{p} can be estimated by relating the mean and variance of the WFP. Considering a single species whose dynamics are covered by the univariate WFP

$$dX_t = \lambda(p - X_t) dt + \sqrt{2X_t(1-X_t)} dW_t \quad (19)$$

we can look at the dynamics of the first two moments, $m_t = E[X_t]$ and $s_t = E[X_t^2]$. Using Ito's lemma and the Ito Isometry one obtains the system of ODEs

$$\begin{aligned} \frac{dm_t}{dt} &= \lambda(p - m_t) \\ \frac{ds_t}{dt} &= 2(1 + \lambda p) m_t - 2(1 + \lambda) s_t \end{aligned} \quad (20)$$

which yields the stationary mean, $m_\infty = \lim_{t \rightarrow \infty} m_t$ and second moment $s_\infty =$

$\lim_{t \rightarrow \infty} s_t$

$$\begin{aligned} m_\infty &= p \\ s_\infty &= \frac{1 + \lambda p}{1 + \lambda} m_\infty. \end{aligned} \quad (21)$$

Thus, the stationary mean is $\mu = p$ and the stationary variance is given by $v^2 = s_\infty - m_\infty^2$, yielding $v^2 = \frac{p(1-p)}{1+\lambda}$.

These two equations give one way of estimating λ and \mathbf{p} from a time-series dataset: calculate the sample mean of each species $\{\hat{\mu}_i\}_{i=1}^n$ and then the sample variance of a focal species, \hat{v}_i^2 to get

$$\begin{aligned} \hat{p}_i &= \hat{\mu}_i \\ \hat{\lambda}_i &= \frac{\hat{\mu}_i(1 - \hat{\mu}_i)}{\hat{v}_i^2} - 1 \end{aligned} \quad (22)$$

or one can utilize their full data by getting the sample mean of each species and sample variance of each species, \hat{v}_i^2 and use the full dataset to estimate $\hat{\lambda}$

$$\hat{\lambda}_n = \sum_{i=1}^n \frac{\hat{\lambda}_i}{n} \quad (23)$$

Other methods can be developed, such as pooling species until the pool has an average relative abundance near 0.5, but for this paper I focus on $\hat{p}_i = \hat{\mu}_i$ and $\hat{\lambda}_n$ given in equation 23.

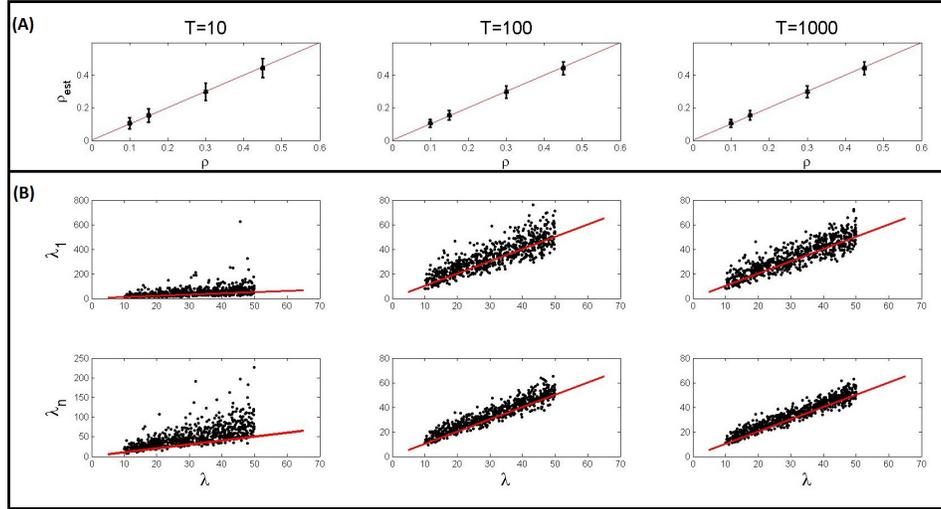


Figure 3: Parameter Estimation - plots of true values of the parameter versus their estimates for 700 solutions to equation 18 integrated over $t \in [0, 2]$ and sub-sampled for $T \in [10, 100, 1000]$ timesteps. $\mathbf{p} = [1, 1.5, 3, 4.5]$ for every solution and each solution had a unique λ drawn from a uniform distribution on $[10, 50]$. (A) The time-average provides an unbiased estimate of \mathbf{p} , shown here as plots of the mean estimator ± 2 standard deviations over the 700 replicates. (B) Equations 22 and 23 provide estimates of λ , but appear to be biased over-estimates of the true value of lambda, possibly due to the dimensionality-correction factor introduced in equation 18.

The last parameter that remains to be estimated is the length of time, T_{est} , to simulate our surrogate datasets. This can be obtained by noting that the true autocorrelation time of the WFP is $\tau_c = \frac{1}{\lambda}$, i.e. the autocorrelation $C(X_t, X_{t+\tau_c}) = e^{-1}$. For a dataset with M time points the autocorrelation function can be interpolated to find the time lag, τ with $0 \leq \tau \leq M$, where $C(X_t, X_{t+\tau}) = e^{-1}$. We want a surrogate dataset to be simulated for the same number of autocorrelation times as our empirical data, i.e. $\frac{T_{est}}{\tau_c} = \frac{M}{\tau}$, thus the time-length for the surrogate dataset will be

$$T_{est} = \frac{M}{\lambda\tau} \quad (24)$$

Figure 4 indicates that equation 24 produces a biased estimate of T_{est} , especially when using the true value of λ instead of its estimate λ_1 .

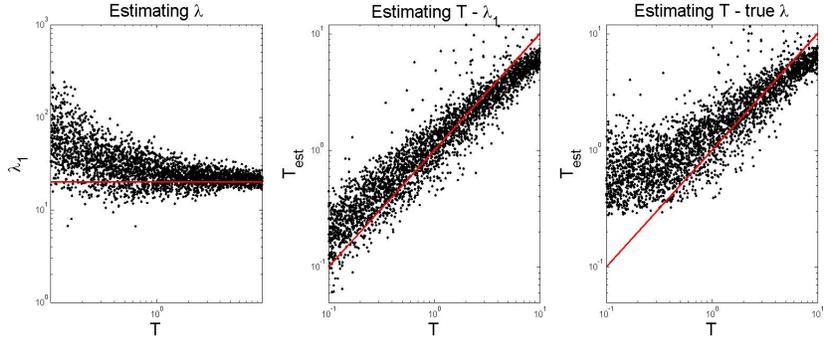


Figure 4: Estimating T and simultaneous inference of λ and T . 3000 values of T were drawn from a log-uniform distribution, $T \in [0.1, 10]$. Equation 18 with $\lambda = 20$, $n = 2$ and $\mathbf{p} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ was numerically integrated over $[0, T]$ for each value of T . The full trajectory was sub-sampled at $M = 300$ time points equally spaced between 0 and T , and from that dataset λ_1 was estimated through equation 22 and then T_{est} was obtained through equation 24.

Application to Real Data

To demonstrate a test of neutrality, I simulated a 15-species neutral community using the method described below in equation 18, a set of 15 geometric Brownian motions

$$d \log Y_t^i = \mu dt + \sigma dW_t^i \quad (25)$$

and a set of 15 mean-reverting geometric Brownian motions

$$d \log Z_t^i = \mu (\bar{z}_i - Z_t^i) dt + \sigma dW_t^i \quad (26)$$

with $\mu = 15$ and $\sigma = 30$. Solutions to these equation are later projected onto the simplex, $\mathbf{Y}_t^\Delta = \frac{\mathbf{Y}_t}{|\mathbf{Y}_t|}$, $\mathbf{Z}_t^\Delta = \frac{\mathbf{Z}_t}{|\mathbf{Z}_t|}$. Sample trajectories for these three simplex-valued Ito processes are plotted in figure 5 below.

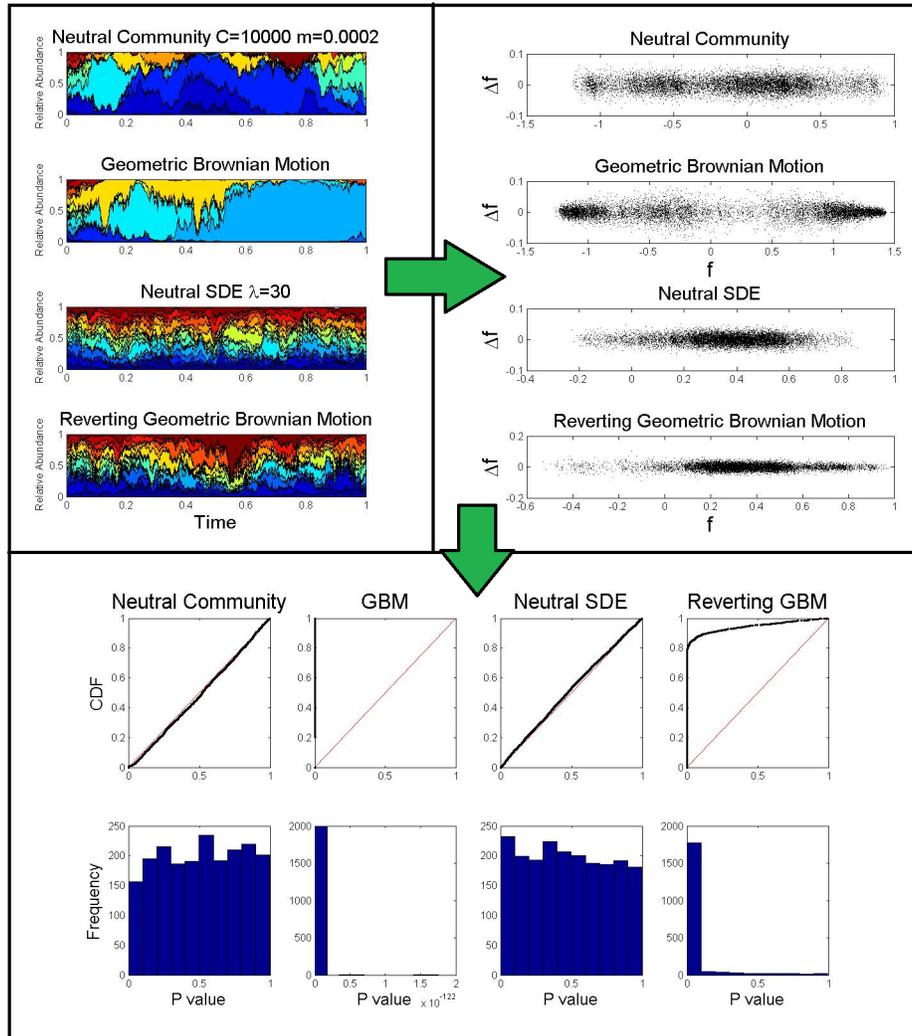


Figure 5: Testing Neutral Drift. The function $g(\mathbf{X}_t) : \Delta^n \rightarrow \mathbb{R}$ defined in equation 12 transforms a Wright-Fisher trajectory from equation 1 into an Ito process with constant volatility seen in the Gaussian distribution of Δg and the homoskedasticity and absence of a trend in scatterplots of Δg versus g . Processes with different covariance matrices, such as the simplex-projected geometric Brownian motion of equation 25 or the simplex-projected mean-reverting geometric brownian motion of equation 26, may have a combination of non-Gaussian distributions of Δg and trends or heteroskedasticity in Δg with respect to other state variables, g , \mathbf{X}_t , t . In the bottom panel, P-value distributions from White homoskedasticity tests of Δg vs. g are produced from 2,000 randomly drawn groupings, g . Tests of uniformity of this P-value distribution provide a test of the WFP as a null model for time-series datasets. The P-values are uniformly distributed when the process is a neutral community of size $C = 10,000$ or a WFP, and non-uniformly distributed (with many low P-values) for the geometric brownian motion and the mean-reverting geometric brownian motions from figure 1. The dependence of the P-values, however, causes a high type 1 error rate for a standard Kolmogorov-Smirnov test; correcting for this dependence - either by analytical calculation of the simulation of surrogate datasets to produce a null distribution of KS-statistics - will allow for more accurate statements of the error rates in this test.

The fact that there are 2^n solutions all with constant volatility can be exploited to make a test of neutrality that does not require subjective choice of grouping $\{a_i\}$. If a process is a WFP and the time-points are equally spaced, then the P-values for any test of normality or homoskedasticity or trends of Δg should be uniformly distributed. Thus, 2^n P-values can be generated from any test of constant volatility of the 2^n variance-stabilizing transformations and a uniform distribution of P-values fails to reject the null hypothesis of neutral drift. A left-skewed P-value distribution leads to the qualitative rejection of the WFP as a null model for the time series dataset. The bottom panel of Figure 5 shows the resulting P-value distributions from normality tests of Δg for 2,000 randomly sampled groupings, g , indicating a uniform distribution of P-values for the WFP and a correct rejection of the non-neutral drifts plotted above.

Next, I apply these same Δg homoskedasticity tests time-series datasets from 372 time points of a male's palm, tongue and gut microbiomes, 130 time points of a female's palm, tongue and gut microbiomes [2] and 1256 day-end prices and market capitalization of all companies continuously in the S&P 500 from January 1st 2000 to January 1st 2005 are tested for non-neutral dynamics (data obtained from CRSP). For these datasets, one could fit species-abundance distributions as in figure 6, but I hope to show that the species-abundance distribution fitting is a much weaker test than the time-series analysis presented here and are summarized in figure 7.

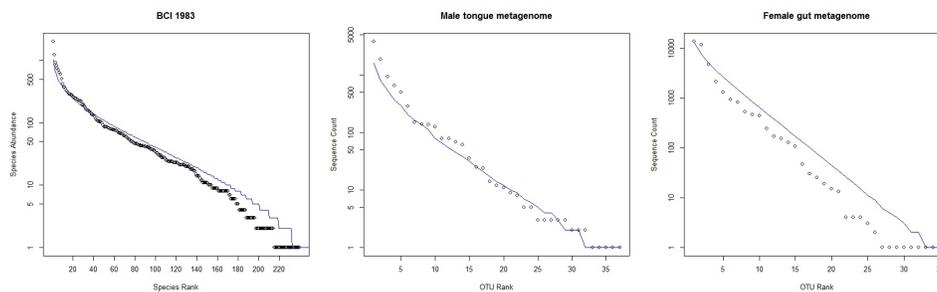


Figure 6: Species-abundance distributions for BCI [4], the male tongue and female gut metagenomes [2] fit using the “untb” and “sads” packages in R. Species-abundance distributions are just snapshots of communities fluctuating over time - the added information contained in time-series datasets can be exploited to develop more powerful tests of Neutral Theory.

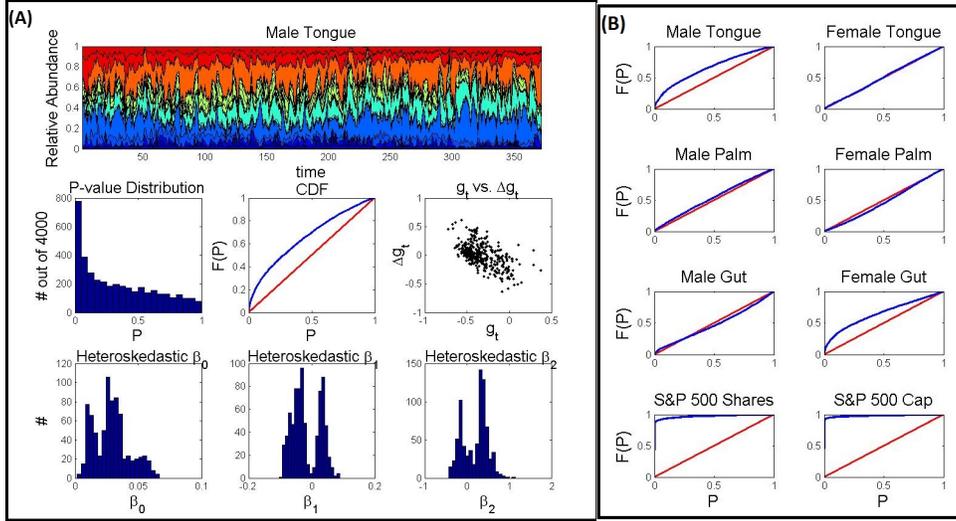


Figure 7: Testing neutral theory with ecological and economic time-series datasets. Time-series datasets from 372 time points of the human microbiome [2] and 1256 time points of share prices and market capitalization from all companies continuously in the S&P 500 from years 2000-2005 are tested for non-neutral dynamics by White tests on g vs. Δg for 4,000 randomly drawn variance-stabilizing transformations, g (using function TestHet in MATLAB by pure quadratic regression of residuals, $\epsilon^2 = \beta_0 + \beta_1 g_t + \beta_2 g_t^2 + e$, following linear regression of g on Δg , $\Delta g_t = \gamma_0 + \gamma_1 g_t + \epsilon$). (A) Metagenomic samples from a male's tongue qualitatively appear to have non-neutral dynamics based on the non-uniformity of the P-value distribution from White tests also seen clearly in the CDF. Alternative models or explanations for stochastic dynamics of this male's tongue microbiome may gain insight from inspection of scatter plots of g versus Δg and the coefficients, β for auxiliary regression in the significantly heteroskedastic cases ($P < 0.05$). The scatterplots of g versus Δg indicate a clear negative slope that persists across many choices of g (this is confirmed by looking at the dominance of negative values for γ_1) - it's possible that the long time-intervals between samples require a drift-correction, a hypothesis investigated further in figure 7 below. Looking for heteroskedasticity beyond the clear trend in Δg , we see that for the male tongue, 751 out of the 4,000 choices of g (18%) had $P < 0.05$ from the White test. 513 out of the 751 cases of significant heteroskedasticity had $\beta_2 > 0$, which correspond to a volatility smile of g - high variance of Δg at the extreme values of g . (B) Repeating a similar analysis for other microbiomes and for the S&P 500 data lead to various qualitative results ranging from clear rejection of the WFP for the movement of market weights and relative market cap to more uniformly-distributed P-values for a female's tongue microbiome. The dependence of the P-values on each other invalidate a Kolmogorov-Smirnov goodness-of-fit test for concise statements of acceptance/rejection of the WFP in these cases.

The microbiome datasets all appeared to yield a negative relationship between g_t and Δg_t . To test whether or not this could be caused by coarse sampling of a Wright-Fisher process, values of λ and T were estimated for the male tongue microbiome by grouping the first 200 operational taxonomic units (OTUs), yielding a mean relative abundance $\hat{\mu} = 0.5652$, and then applying

equations 22 and 24 described above. This yielded $\lambda = 18.0063$ and interpolating the first 5 values of the autocovariance to find where τ_c (in units days) such that $C(X_t, X_{t+\tau_c}) = e^{-1}$, we get an estimate $\tau_c = 1.5318$ and consequently $T_{est} = 13.4872$. 1000 surrogate datasets corresponding to this grouping of the first 200 OTUs was constructed by integrating the 2-species equations 18 and 17 (which are identical for $n = 2$) over $t \in [0, T_{est}]$ using the decomposed $\sigma(\mathbf{f}_t)$ discussed in equations 12-14 and an Euler-Maruyama integration scheme with $\Delta t = 10^{-5}$. No samples left the boundary of the simplex, so there was no need to re-sample trajectories. The full surrogate time-series was sub-sampled at 372 equally spaced time points and the results are summarized below in figure 8.

Discussion

Complex adaptive systems change with time and quantifying the uncertainty of their future states has major implications for investment in and management of these systems. Ecological communities are composed of many individual organisms competing over the same, finite, limiting resources, and these communities turn over as organisms reproduce, die, and migrate. Economic markets are composed of many companies competing over a finite amount of capital - measured in human capital, number of loyal customers, capital stock or investment capital - and the market shares of these companies changes with time. The birth & death process underlying the turnover of these ecological and economic systems has led many authors to hypothesize that ecosystems, economies, and other complex adaptive systems may be qualitatively similar [11, 17, 1]. The Wright-Fisher Process is a null model for the turnover of such systems with birth, death, dispersal, and zero-sum competition over finite resources. It can be used as a null model for populations of genes [16], communities of tropical trees [12] market weights [9, 20] and other systems, and it is a compelling null hypothesis as neutrality may be a likely occurrence because anything with a significant competitive disadvantage will be quickly driven extinct, leading to an evening of competitive abilities over evolutionary timescales [13].

In ecology and population genetics, Neutral Theory has been developed for communities with zero-sum competition. Clear examples of a zero-sum competition exist in nature, such as competition over limiting space - as in percent cover of trees in the Amazon or the amount of land allocated to one of a set of land uses such as agricultural, residential, open space, national forest, etc. Some communities or markets may not be strictly zero-sum, but their approximately constant size might warrant zero-sum competition as an approximation. However, since Pal [20] has shown that one particular model of market weights can still be a Wright-Fisher process in its relative abundances despite not having zero-sum constraints on absolute abundances, the rejection of zero-sum community dynamics does not reject the possible neutrality of the species in the community. Some authors use species-abundance distributions to test for neutral drift, but since infinitely many models have the same stationary distributions such tests may not be the most powerful. In the presence of time-series datasets for eco-

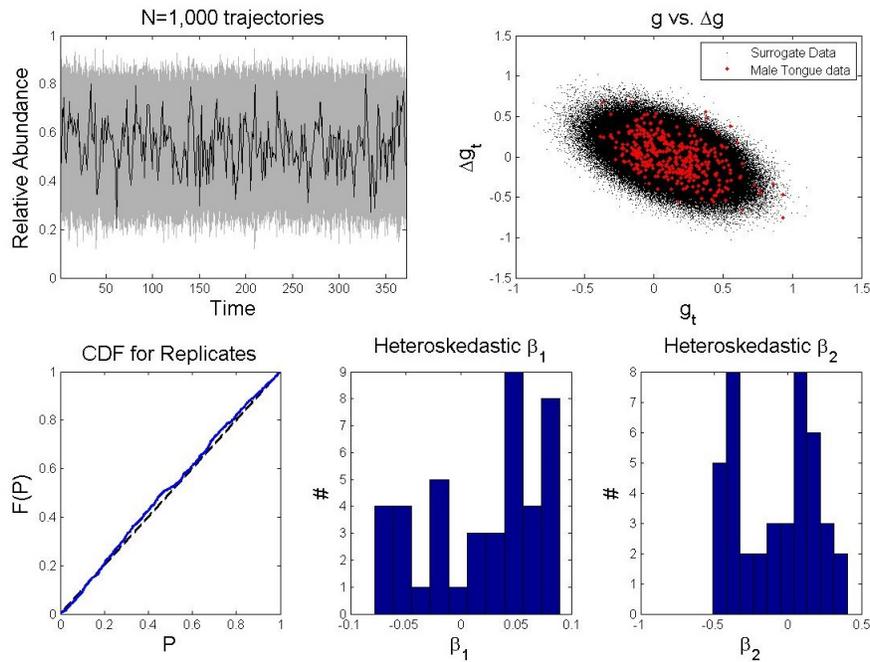


Figure 8: 1,000 surrogate trajectories were run using the parameter estimation and simulation tools provided here to investigate whether the negative slope of Δg_t versus g_t could arise in neutral simulations with long time-intervals between samples. In the top-right sub-panel, I've plotted Δg versus g for all 1,000 replicates (a total of 372,000 points), from which we see that under the null hypothesis of a WFP we can obtain a cloud of points with a negative slope of Δg vs. g if the time between samples is long (black dots are from surrogate datasets, red dots are from the grouping of the male tongue used to parameterize these simulations). Thus, the negative slope of Δg versus g seen in figure 7a is not the noteworthy departure from neutrality. However, the P-values from White homoskedasticity tests are still uniformly distributed even with longer time intervals between timepoints. Since the trajectories simulated here are independent, the P-values in the White tests are also independent (since there is only one informative volatility-stabilizing transformation for each trajectory), and a KS test can be used to suggest that the CDF is in fact uniformly distributed with $P = 0.3085$. Further analysis of the auxiliary regression shows that 20 out of 42 of the trajectories yielding significant heteroskedasticity had $\beta_2 > 0$. The 20 out of 42 $\beta_2 > 0$ in the WFP is significantly different from the 513 out of 751 in the male tongue dataset ($P = 9.9 \times 10^{-9}$). This provides more evidence that the non-neutrality of the male tongue may come from a tendency for the community to become more volatile when it is farther from its mean; this could be verified by constructing alternative models with hypervariable rare & abundant species or analogs of stochastic volatility models which try to explain a similar volatility smile in financial data.

logical or economic systems, one might hope to use the added information from the time series to construct stronger neutrality tests.

In this chapter, I provided two main tools for hypothesis-testing of neutral drift in time-series ecological and economic datasets: volatility-stabilizing transformations, and a means for the fast and reliable simulation of WFP surrogate datasets whose parameters can be estimated from the data. However, more work needs to be done to make these tools accessible to researchers as reliable and user-friendly packages in R and MATLAB. For the constant-volatility tests, the multiple-hypothesis-testing framework of testing the uniformity of P-values from homoskedasticity/constant-volatility tests for the multiple choices of g requires a test-statistic, such as the Kolmogorov-Smirnov test statistic, whose null distribution is known. For the simulations, the viability of higher-order integration schemes and adaptive algorithms for adjusting the time-window of integration could be explored to develop more reliable simulators. For parameter estimation, the clear bias in the estimates of λ and T must be corrected, which may be resolved using techniques from the literature on parameter estimation in stochastic differential equations.

Ultimately, though, the rejection of a null hypothesis is only as compelling as the null hypothesis. The WFP has been proposed as a null model for many systems, but repeated rejection of the WFP may render it obsolete. Time-series datasets can provide more powerful tests and enable the swift rejection of the WFP along with the construction of more suitable alternative models by careful assessment of how our data differ from neutrality - e.g. by having hyper- or hypo-variable rare species or particular taxonomic groupings (e.g. C3 or C4 grasses) which yield particularly non-neutral dynamics. Some alternative models already exist or are in development, such as the Atlas model [8] or urn processes with frequency-dependent stabilizing mechanisms analogous to Janzen-Connell effects ([14, 5], Socolar & Washburne, in prep), and comparison of patterns in g_t vs. Δg_t plots like figures 7 and 8 will be a stepping stone towards the development and motivation of more compelling alternative models. The production of reliable models of the stochastic evolution of complex-adaptive systems can enable better forecasts of species richness and community viability in ecology to enable better reserve design or fisheries management; it can enable the construction of portfolios to achieve particular goals balancing risk and return [8]; it can allow the development and clinical trial of more potent probiotics through a better understanding of the pharmacokinetics of probiotics.

Acknowledgments

I thank my colleagues D. Lacker for his patient instruction in all things stochastic and many interdisciplinary discussions on the interface between ecology and economics, S. Pacala for introducing me to the problem and for providing short but brilliant insights along the way, R. Chisholm for many useful discussions and feedback, E. Shertzer for help with parameter estimation, J. Burby and S. A. Levin for many hours of fun math working towards the variance-stabilizing

transformation, J. Squire for computational support and inspirational badassedness, and J. O'Dwyer for his help with species-abundance distributions and other null models for the dynamics of gut biota.

References

- [1] W Brian Arthur. Complexity and the economy. *science*, 284(5411):107–109, 1999.
- [2] J Gregory Caporaso, Christian L Lauber, Elizabeth K Costello, Donna Berg-Lyons, Antonio Gonzalez, Jesse Stombaugh, Dan Knights, Pawel Gajer, Jacques Ravel, Noah Fierer, et al. Moving pictures of the human microbiome. *Genome Biol*, 12(5):R50, 2011.
- [3] Ryan A Chisholm, Richard Condit, K Abd Rahman, Patrick J Baker, Sarayudh Bunyavejchewin, Yu-Yun Chen, George Chuyong, HS Dattaraja, Stuart Davies, Corneille EN Ewango, et al. Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. *Ecology letters*, 17(7):855–865, 2014.
- [4] Richard Condit. *Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a comparison with other plots*. Springer Science & Business Media, 1998.
- [5] Joseph H. Connell. On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In P. J. den Boer and G. Gradwell, editors, *Dynamics of Populations*. Pudoc, Washington, 1971.
- [6] Ciara Dangerfield, David Kay, and Kevin Burrage. Stochastic models and simulation of ion channel dynamics. *Procedia Computer Science*, 1(1):1587–1596, 2010.
- [7] Justin C Fay and Chung-I Wu. Hitchhiking under positive darwinian selection. *Genetics*, 155(3):1405–1413, 2000.
- [8] E Robert Fernholz. *Stochastic portfolio theory*. Springer, 2002.
- [9] Robert Fernholz and Ioannis Karatzas. Relative arbitrage in volatility-stabilized markets. *Annals of Finance*, 1(2):149–177, 2005.
- [10] Ronald Aylmer Fisher. *The genetical theory of natural selection*. Clarendon, 1930.
- [11] John H Holland. Complex adaptive systems. *Daedalus*, pages 17–30, 1992.
- [12] Stephen P Hubbell. *The unified neutral theory of biodiversity and biogeography (MPB-32)*, volume 32. Princeton University Press, 2001.

- [13] Stephen P Hubbell. Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional ecology*, 19(1):166–172, 2005.
- [14] Daniel H. Janzen. Herbivores and the number of tree species in tropical forests. *The American Naturalist*, 104(940):501–528, 1970.
- [15] Michael Kalyuzhny, Efrat Seri, Rachel Chocron, Curtis H Flather, Ronen Kadmon, and Nadav M Shnerb. Niche versus neutrality: A dynamical analysis. *The American Naturalist*, 184(4):439–446, 2014.
- [16] Motoo Kimura. *The neutral theory of molecular evolution*. Cambridge University Press, 1985.
- [17] Simon A Levin. Ecosystems and the biosphere as complex adaptive systems. *Ecosystems*, 1(5):431–436, 1998.
- [18] GN Milstein, Eckhard Platen, and Henri Schurz. Balanced implicit methods for stiff stochastic systems. *SIAM Journal on Numerical Analysis*, 35(3):1010–1019, 1998.
- [19] Esteban Moro and Henri Schurz. Boundary preserving semianalytic numerical algorithms for stochastic differential equations. *SIAM Journal on Scientific Computing*, 29(4):1525–1549, 2007.
- [20] Soumik Pal. Analysis of market weights under volatility-stabilized market models. *The Annals of Applied Probability*, 21(3):1180–1213, 2011.
- [21] Daniel W Stroock and SR Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233. Springer Science & Business Media, 1979.
- [22] Fumio Tajima. Statistical method for testing the neutral mutation hypothesis by dna polymorphism. *Genetics*, 123(3):585–595, 1989.
- [23] N. G. van Kampen. *Stochastic processes in physics and chemistry*. North-Holland, Amsterdam ; New York : New York, 1981.
- [24] Sewall Wright. Evolution in mendelian populations. *Genetics*, 16(2):97, 1931.