1   **Title:** Expected differences in diversity and rarity between communities containing sexually

2   versus asexually reproducing taxa

3   **Running Title:** Reproductive method alters diversity, rarity

4   **Authors:** Cristina M. Herren[1, 2, 3]

5   **Author affiliations:**

6   [1]Harvard Data Science Initiative, Harvard University, Cambridge, Massachusetts, USA

7   [2]Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston,

8   Massachusetts, USA

9   [3]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA

10   Email: cristina_herren@hms.harvard.edu

11   310 Countway Medical Library, 10 Shattuck St, Boston MA 02115

12

12 **Abstract**

13   Recent studies evaluating the community structures of microorganisms and macro-

14 organisms have found greater diversity and rarity within micro-scale communities, compared to

15 macro-scale communities. However, reproductive method has been a confounding factor in these

16 comparisons; the microbes considered generally reproduce asexually, while the macro-organisms

17 considered generally reproduce sexually. Sexual reproduction imposes the constraint of mate

18 finding, which can have significant demographic consequences by depressing birth rates at low

19 population sizes. Here, I examine theoretically how the effects of mate finding in sexual

20 populations translate to the emergent community properties of diversity, rarity, and dominance.

21 Using continuous-time Markov chain models, I compare communities with and without

22 constraints of mate limitation. In mate-limited sexual populations, the decreased growth rates at

23 low population densities translates to a much higher extinction rate. In communities consisting of

24 sexually reproducing taxa, the increased extinction rate due to mate limitation decreases

25 expected diversity. Furthermore, mate limitation has a disproportionately strong effect on taxa

26 with low population density, leading to fewer rare taxa. These shifts in community structure

27 mirror recent empirical studies of micro versus macro communities, indicating that reproductive

28 method may contribute to observed differences in emergent properties between communities at

29 these two scales.

30   **Introduction**

31   Ecologists have historically been fascinated by the diversity of microbial communities

32   (Hutchinson 1961), and several recent studies have indeed demonstrated differences in

33   community structure between microbes and larger "macro" organisms (Nemergut et al. 2013,

34   Hansen and Carey 2015, Locey and Lennon 2016, Shoemaker et al. 2017, Meyer et al. 2018).

35   Generally, microbial communities have higher diversity that results in part from the large

36   number of rare taxa (Neufeld and Lynch 2015). But, other properties, such as abundance of the

37   most dominant taxon, are indistinguishable between communities at the two different scales

38   (Locey and Lennon 2016). Despite increasing data on which to base these comparisons, the

39   mechanisms generating these patterns of population distributions within and between

40   communities are poorly understood (Shade 2017). One prominent additional difference between

41   many of the microbial populations and macro populations in prior comparative studies is

42   reproductive method; the microbial populations considered (bacteria, archaea, and most

43   phytoplankton) reproduce asexually, while most macro populations considered have sexual

44   reproduction. Here, I examine theoretically whether reproductive method can contribute to

45   observed differences in community structure between asexually reproducing microorganisms and

46   sexually reproducing macro organisms.

47   Individuals in sexually reproducing populations must encounter a mate before

48   reproducing, whereas asexual individuals do not have this constraint. Mate finding and its

49   consequences on population dynamics have been extensively studied in the theoretical literature

50   (beginning with Volterra 1938), in part because it is one mechanism that causes Allee effects

51   (reviewed in Gascoigne et al. 2009). An Allee effect is defined as positive density dependence

52   within a population, meaning that individual-level growth rates increase as population density

53   increases (Odum and Allee 1954). When an Allee effect is present, the benefit of encountering

54   another individual from the population outweighs negative interactions, such as competition, and

55   individuals become more reproductively successful as density increases (Courchamp et al. 1999).

56   In populations with sexual reproduction, sparse populations are slow growing due to the inability

57   to find a mate. Mate encounters become more frequent as the population grows, such that per-

58   capita fitness increases as density increases. The effects of mate finding on population growth

59   are prominent when population sizes are small, but decrease when population size is large and

60   mates are no longer limiting (Dennis 1989).

61   Many previous theoretical models have considered mate finding and Allee effects using

62   differential or difference equations describing the population growth rate (Odum and Allee 1954,

63   Dennis 1989, Boukal and Berec 2002). Strong reductions in birth rates due to mate limitation can

64   cause population declines at low abundance, effectively setting a "critical density" below which

65   the population becomes extinct (Gerritsen 1980). When the population size is greater than the

66   critical density, the population continues to grow until reaching a stable equilibrium at its

67   carrying capacity (Stephens et al. 1999). However, a major drawback of deterministic models is

68   the inability to consider time to extinction for populations with a positive stable equilibrium;

69   with deterministic equations, any population with a positive stable state will persist indefinitely.

70   This result conflicts with the empirical observation that smaller populations are more vulnerable

71   to extinction (Purvis et al. 2000).

72   Stochastic models are promising for studying demographic consequences of mate

73   finding, because they allow for extinction in populations that would otherwise reach a positive

74   carrying capacity (Lande 1993). Whereas the persistence of populations in deterministic

75   equations is governed by local population growth rates around an equilibrium, persistence in

76   stochastic models depends on the growth rates at every density (Assaf and Meerson 2010). In

77   other words, the chance of extinction in real populations is related to population growth rates

78   near zero, which are important in stochastic models but rarely considered in deterministic

79   models. Several forms of stochastic populations models have been used to study populations

80   with Allee effects, often with discrete time models (Stephan and Wissel 1994, Allen et al. 2005,

81   Sun 2016). These studies have concluded that diminished growth rates at low population

82   densities can substantially decrease expected time to extinction (Stephan and Wissel 1994,

83   Dennis 2002). However, it is computationally difficult to model multiple interdependent

84   populations or populations with overlapping generations in discrete time models (Allen and

85   Allen 2003), which is often a prohibitive barrier to such studies.

86   Here, I compare population and community dynamics between communities that must

87   find mates before reproducing and communities where populations have no mate limitation (a

88   case equivalent to asexual reproduction). I use stochastic models to evaluate demographic

89   consequences of mate finding. First, I use continuous-time Markov chain models (CTMCs) to

90   study how mate limitation alters time to extinction for single populations. These models use a

91   computationally efficient simulation algorithm, which allows for simulation of multiple

92   coexisting populations. Such models have been extensively used for simulating chemical

93   reaction networks (Gillespie 2007), but can also be used for modeling population dynamics

94   (Dobramysl et al. 2018). After obtaining the mean times to extinction obtained from these

95   models, I use the island biogeography framework to evaluate how varying extinction times

96   translate to changes in community diversity. The island biogeography framework posits that the

97   expected long-term community diversity can be calculated by identifying the number of taxa

98   where immigration and extinction rates are equal (MacArthur and Wilson 1967). In these

99   models, I assume identical immigration rates between the various communities, but extinction

100   rates are a function of mate limitation. Finally, I simulate communities consisting of populations

101   with differing growth rates to evaluate how consequences of mate limitation scale to

102   heterogeneous communities. I show that the constraint of mate search decreases diversity,

103   primarily by excluding rare taxa, whereas dominance of the most abundant population is

104   unaffected.

105

106   **Methods**

107   *Single population dynamics*

108   First, I studied the effects of mate limitation on the time to extinction for single

109   populations. I used CTMCs to evaluate time to extinction, implemented with the Gillespie

110   algorithm (Gillespie 1977). Briefly, these models record births and deaths in a population as

111   events that occur with varying frequency, depending on population size. Births are marked by the

112   addition of a single individual to the population, whereas deaths remove a single individual. The

113   overall rate at which any event (birth or death) occurs is the sum of the birth and death rates. The

114   time until the next event is exponentially distributed with a parameter equal to the summed event

115   rates. Therefore, as event rates increase, waiting time until the next event decreases. After

116   drawing a random value from the exponential distribution for the time increment, the magnitudes

117   of the instantaneous birth and death rates indicate whether a birth (add one individual) or death

118   (remove one individual) is more likely to occur. Another random number is generated to

119   determine whether a birth or death event transpires. After an individual is added or removed

120   from the population, birth and death rates are updated based on the new population size, and the

121   steps repeat. Extinction occurs at the first time point where the population equaled zero.

122   Throughout this study, I consider populations that are self limiting. In deterministic

123   models, self-limiting populations experiencing logistic growth reach a stable carrying capacity

124   determined by the intrinsic birth rate (b) and the density-dependent death rate (d) (Eq. 1).

125

$$\frac{dN}{dt} = bN - dN^2$$

126   Eq. 1

127

128   In the stochastic model formulation, births and deaths are modeled as discrete events,

129   also referred to as "reactions" (Anderson and Kurtz 2011). The equivalent birth event rate for

130   logistic growth model is equal to bN (Eq. 2):

131

132    $$Event: N \rightarrow N+1 \quad Rate: bN$$ Eq. 2

133

134   To study effects of mate limitation, I modified the birth event rate to include mate search.

135   Previous work has yielded an equation governing the encounter rate between one individual and

136   other individuals when searching in three-dimensional environments (Gerritsen and Strickler

137   1977). Results for the two-dimensional case yield qualitatively equivalent results and are shown

138   in the supplementary material. The mate encounter rate is dependent upon the speed at which

139   individuals move (V) and the radius at which they can detect a mate (R). Here, I assume that

140   males and females move at the same speed and that there is a 1:1 male to female ratio.

141   Multiplying the intrinsic birth rate (b) by the probability that at least 1 mate will be encountered

142   yields the following birth event rate for mate-limited populations (Eq. 3):

143

144 $$Event: N \rightarrow N+1 \quad Rate: bN \bullet \left(1 - e^{-2\pi R^2 VN/3}\right)$$ Eq. 3

145

146 The death rate functions used for the two cases were identical (Eq. 4):

147

148 $$Event: N \rightarrow N-1 \quad Rate: dN^2$$ Eq. 4

149

150   As an illustration of the difference between deterministic and stochastic models, I

151 investigated cases where long-term population dynamics of the various populations would be

152 equivalent in the deterministic case; both mate-limited and non-limited populations would have

153 nearly identical carrying capacities if these dynamics were translated to deterministic models. I

154 simulated population trajectories and evaluated time to extinction using CTMCs. I used the two

155 birth rates expressions above (Eq. 2 and 3) for the scenarios with and without mate limitation,

156 and Eq. 4 for the death rate in all models. I recorded the mean time to extinction (MTE) for 1000

157 simulated chains (populations).

158

159 *Evaluating diversity with island biogeography theory*

160   The theory of island biogeography formalized the concept that long-term community

161 diversity is governed by the rate at which taxa enter the community (i.e. immigration) and the

162 rate that taxa leave the community (i.e. extinction). In island biogeography models, the

163 immigration rate and extinction rate of taxa within a community change as a function of the

164 number of taxa currently present in a community (MacArthur and Wilson 1967). Thus, the

165 expected diversity (defined here as the number of coexisting taxa) of the community is identified

166 by finding the number of taxa where the immigration and extinction rates are equal. To evaluate

167   the accuracy of this analytical approximation for these simulations, I calculated the expected

168   long-term diversity for a community consisting of populations with identical birth and death

169   rates (and, therefore, identical MTEs).

170   I compared results of the analytical estimates of diversity to simulations of diversity in a

171   stochastic reaction network model (coupled, simultaneous CTMCs) explicitly tracking each

172   population. In the stochastic reaction network, the community-level immigration event rate was a

173   function of current diversity. Immigration events were modeled as a population increasing from

174   0 to a small population size, in this case of 2 individuals (Eq. 5):

175

176         $$Event : 0 \rightarrow 2 \quad Rate : (100 - Diversity) \bullet i \qquad\qquad \text{Eq. 5}$$

177

178   The immigration rate decreased linearly as a function of current community diversity, and

179   reached zero when 100 taxa were present. I used an immigration constant of $i = 0.001$, and

180   conducted these simulations across different parameters governing mate finding.

181

182   *Simulation of community structure with heterogeneous taxa*

183   To study how demographic consequences of mate finding scale to communities with

184   heterogeneous populations, I simulated communities where populations had varying intrinsic

185   growth rates (b). I evaluated how changes in mate searching parameters affected the diversity,

186   rarity, and dominance of taxa within the simulated communities. Taxa within each community

187   experienced the same degree of mate limitation, which was determined by changing the values of

188   R and V over the range of 0.5 to 1.2. Across these combinations of search radius and search

189   speed, mate finding is limiting to population growth in small populations, but not limiting in

190   large populations (those with 20 or more individuals).

191   For each combination of search radius and search speed for the mate limited populations,

192   I simulated a stochastic reaction network where intrinsic birth parameters (*b*) were randomly

193   drawn from a lognormal distribution with a mean of the underlying normal distribution was 0

194   and the standard deviation was 0.25. The death parameter (*d*) was fixed at 0.1 for all populations.

195   I used a lognormal distribution of growth rates, because a lognormal distribution provides a good

196   fit to the observed abundance distributions of microbial taxa (Shoemaker et al. 2017).

197   Additionally, populations defined by these growth rates routinely become extinct within

198   computationally tractable time scales. However, I verified that simulation results were

199   qualitatively similar when using a normal distribution of birth rates. Again, immigration was a

200   linearly decreasing function of current diversity, where an immigration event was modeled as a

201   change in population size from 0 to 2 (Eq. 5). Immigrant populations were assigned a new birth

202   parameter from the lognormal distribution. After a burn in period of 5 000 000 events, I recorded

203   instantaneous measurements of diversity, dominance (largest population size), and mean

204   population size every 200 000 events. I compared results of simulations where mates were

205   limiting to results of simulations where mates were not limiting.

206

207   **Results**

208   *Single population dynamics*

209   First, I compared time to extinction for non-limited populations (i.e. asexual reproducers)

210   and for sexual populations where mates must be encountered. For sexual populations, I evaluated

211   two combinations of search radius and search speed. One scenario indicates a poor searcher with

212   low search radius and speed (R and $V = 0.62$), and the other was a more effective searcher with

213   higher search radius and speed (R and $V = 0.8$). I chose parameter values that would yield

214   equivalent long-term population dynamics if these populations were modeled deterministically;

215   all three scenarios have nearly identical population densities where the birth rate equals the death

216   rate, indicating equal carrying capacities in the absence of stochasticity (Fig. 1b). The per-capita

217   birth rate is much higher in small populations for the asexually reproducing populations than for

218   sexually reproducing populations (Fig. 1A). However, the birth rate in the sexual populations

219   increased as individuals became more effective at finding mates. Multiplying individual birth

220   rates by population size yields population-level birth minus death rates (Fig. 1B). Effects of mate

221   limitation are prominent at small populations, but negligible as population size increases. With

222   CTMC models, it is also possible to calculate the probability that the next event in the model will

223   be a birth or a death. In models with asexual reproduction, it is highly unlikely that a death will

224   occur in a small population. This probability of population decline at low population sizes is

225   increased when mate limitation is present (Fig. 1C).

Fig. 1
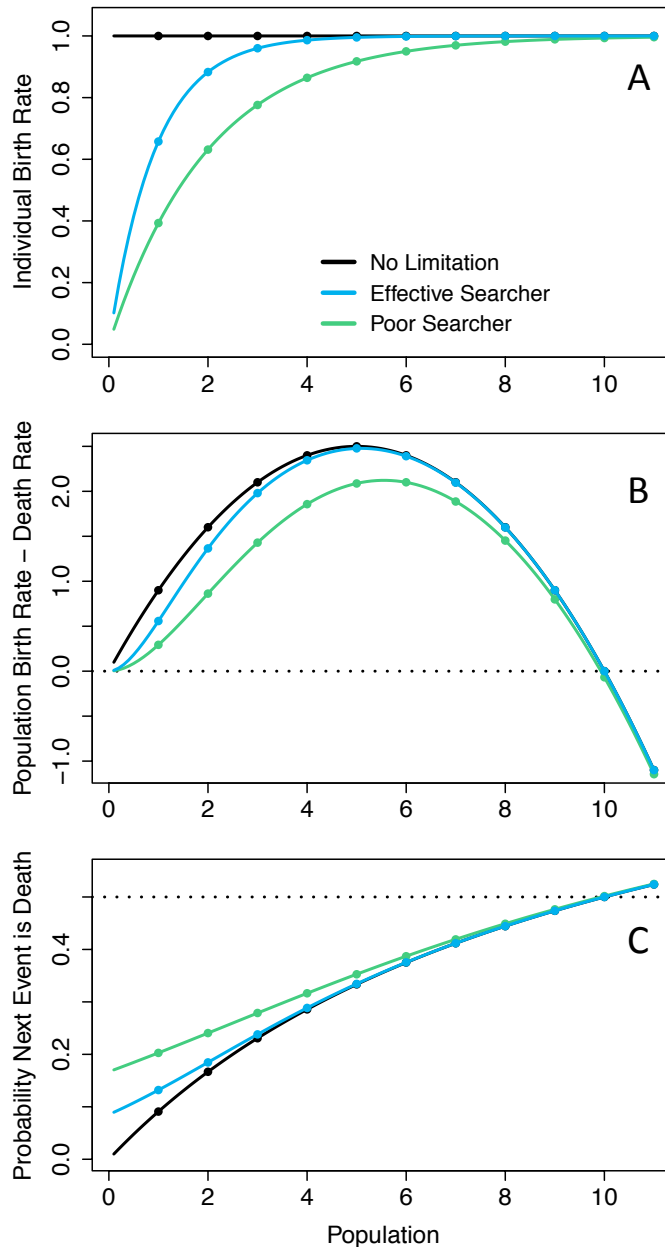
Figure 1: Mate limitation decreases the individual level birth rate at low population density (panel A), which influences both the population-level growth rate (panel B) and the probability that the next event in the model will be a death (panel C). Effects of mate limitation on population growth become negligible as population sizes increase, as visible by the convergence

231   of the three scenarios at larger populations. The effect of mate limitation is the difference

232   between the aseuxally reproducing populations (black line) and the sexually reproducing

233   populations (blue and green lines). Population growth rates are suppressed more strongly in poor

234   mate searchers (green line) than effective mate searchers (blue line). The dashed line in C

235   indicates a probability of 0.5, where a birth and death are equally likely.

236

237   I recorded the time to extinction for 1000 simulated populations parameterized with the

238   three birth rate scenarios shown in Fig 1. All populations had equivalent death rates. To evaluate

239   the effect of initial conditions, I used initial population sizes of 10 and 2. Asexual populations

240   persisted longest of the three scenarios, with a MTE of 2792 for an initial population size of 10

241   and 2664 for an initial population size of 2. In mate-limited populations, MTEs for effective

242   searchers were 1579 and 1557 (for initial size of 10 or 2 individuals), while for poor searchers

243   MTEs were 515 and 477.

244   Across all populations, rapid extinction (very short MTE) was more common than when

245   the initial population size was 2, rather than 10. The decrease in MTEs between populations with

246   an initial population of 10 and an initial population of 2 is partially due to this higher frequency

247   of very short times to extinction (Fig. 2).

248

Fig. 2
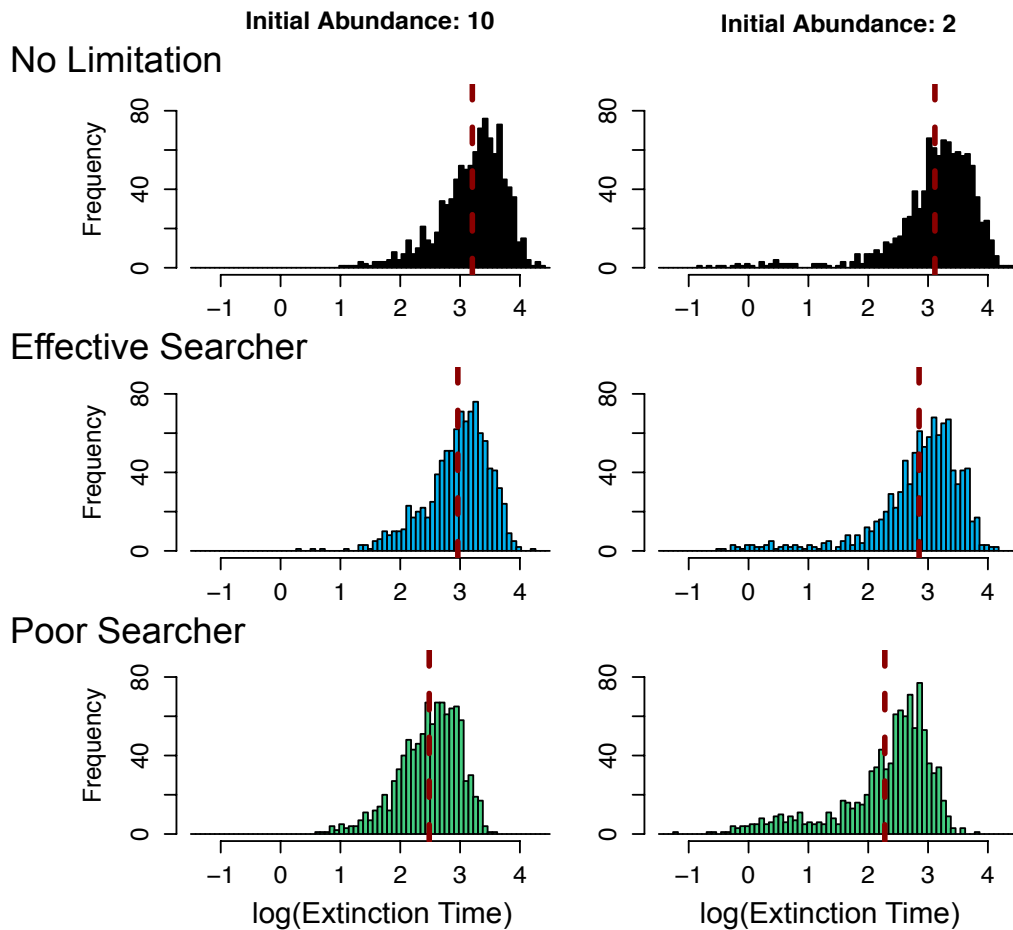
Figure 2: Times to extinction from 1000 simulated populations for communities with no mate limitation (top), mate limitation with effective searchers (midde), and mate limitation with poor searchers (bottom). Mean time to extinction (dashed red lines) decreases as the probability of encountering a mate decreases, and is therefore lowest for poor mate searchers. Time to extintion is also affected by initial population size, and decreases when the population is initiated with 2 individuals (right column) as compared to 10 individuals (left column).

Evaluating diversity with island biogeography theory

258   Assuming that a community consisted of populations with identical birth and death rates,

259   I calculated the estimated long-term diversity for the three birth rate scenarios from the

260   associated extinction rates (shown in Fig. 2). The extinction rate for a single population is

261   1/MTE, meaning the extinction rate for a community of *m* taxa is *m*/MTE. I used the same rate of

262   immigration in each scenario. The immigration rate was a linearly decreasing function of current

263   diversity, and reached 0 when 100 taxa were present (Fig. 3). Thus, no more than 100 taxa could

264   exist in the community. An approximation of long-term diversity under these assumptions can be

265   found using the formula (Eq. 6). For the expected diversity calculations and associated

266   simulations, I used an immigration constant $i = 0.001$, which determines the slope and intercept

267   of the immigration function. However, the stochastic nature of the simulations means that these

268   calculations will be inexact, because the populations never reach equilibrium.

269

$$Expected\ Diversity = \frac{100\ i}{\frac{1}{MTE} + i}$$

270   Eq. 6

271

272   Eq. 6 shows that the long-term diversity is a function of mean time to extinction. As MTE

273   approaches infinity, the expected diversity approaches the diversity level where immigration is

274   zero (in this case, 100). Conversely, as MTE approaches zero, expected diversity also approaches

275   zero. I evaluated the accuracy of this approximation using explicit simulations of simultaneously

276   coexisting populations using the same parameters. The two estimates of diversity were within

277   one unit (taxon). Approximations using Eq. 6 yielded expected long-term diversities of 72.6,

278   60.9, and 32.3 for the scenarios of no limitation, effective searching, and poor searching, while

279   simulations yielded long-term average diversities of 73.1, 61.8, and 32.9.

280

Fig. 3
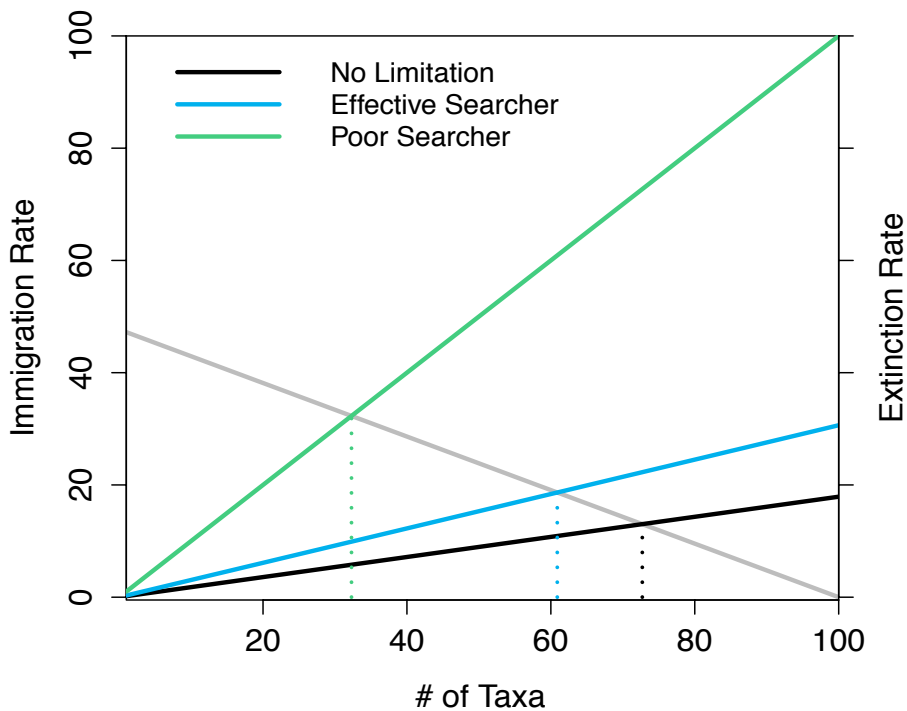
Figure 3: When using the same immigration function (grey line, left axis), mate limitation affects expected diversity by changing time to population extinction. A decreased time to extinction results in a greater slope in the community extinction rate (black, blue, and green lines on right axis). Expected diversity can be found by calculating the number of taxa where the immigration rate and extinction rate intersect (indicated by dashed lines). Communities are most diverse when there is no mate limitation (black lines). When populations are mate-limited, but individuals are effective at finding mates, there is a small decrease in expected diversity (blue lines). When individuals are poor searchers, there is a dramatic decline in diversity due to more rapid extinction (green lines). The time scale at which the immigration and extinction rates are shown here is the MTE of the shortest-lived populations (the poor searcher populations). Using a different time scale alters the y-axes, but does not change where lines intersect.

294   *Simulation of community structure with heterogeneous taxa*

295   Next, I simulated communities in which taxa had varied growth rates. I compared

296   diversity, mean population size, and dominance (population size of most abundant taxon) of

297   communities containing mate-limited taxa and those with non-limited taxa. When sexually

298   reproducing taxa were highly effective searchers due to high search radius and/or speed, mate

299   limitation had little effect on the effective birth rates of those taxa. Then, diversity and

300   population size converged to results from communities containing asexual populations (Fig. 4).

301   However, the abundance of the most dominant population was not affected by search efficacy or

302   reproductive method (Fig. 4). Cells in the figure panels are scaled by the values from the no-

303   limitation simulations (a value of 1 indicates equivalent results).

304

Fig. 4

305
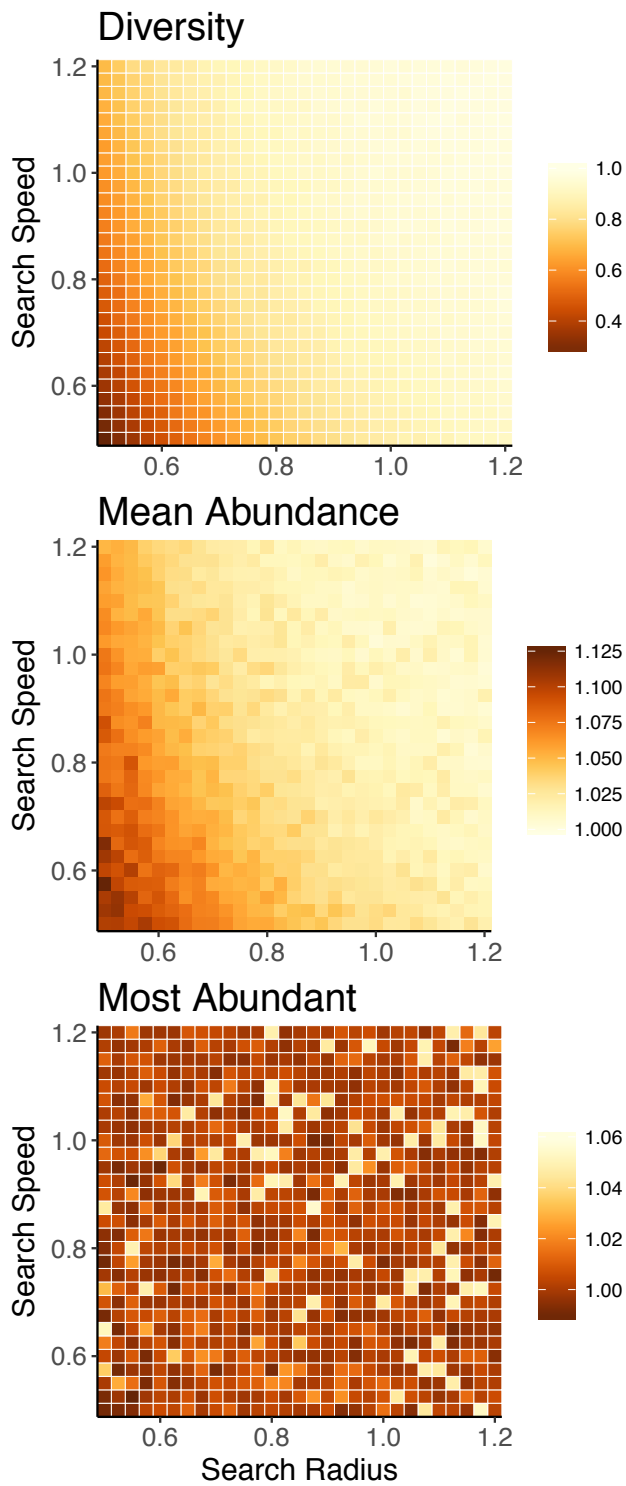
Figure 4: Heatmaps show average diversity (top), abundance (middle), and dominant population

size (bottom) from simulations of mate-limited communities where populations have varying

308   search radius and search speed. Cells within each heatmap show the results for communities

309   consisting of populations with the given search radius and search speed. Cell values are scaled by

310   results from communities without mate limitation. Thus, a value of 1 idicates results equivalent

311   to those of non-limited communities. The diversity (top) and mean abundance size (middle)

312   within communities containing mate-limited populations change in response to mate limitation,

313   as stronger mate limitation corresponds to decreased diversity and larger average population size.

314   When the search radius and search speed are large, the probability of finding a mate approaches

315   1, and results from mate-limited and non-limited communities converge. However, the size of

316   the most abundant population (bottom) is not affected by mate limitation.

317

318   Communities containing the poorest mate searchers experienced the greatest declines in

319   diversity, in comparison to the communities with asexual populations. The communities with the

320   strongest mate search limitation (R = 0.5 and V = 0.5) had mean diversity of 17.7 taxa;

321   conversely, communities where mates were not limiting had mean diversity of 68.5 taxa.

322   Similarly, mean population size was 11.2 for the most mate-limited communities, but 10.2 in

323   non-limited communities. Another measurement of rareness, the skewness of population

324   abundances, showed a similar result (see supplementary material). Higher skewness indicates

325   more rare taxa. Average skewness in the distribution of population sizes was 0.18 in

326   communities with greatest mate limitation, and 0.57 in non-limited communities. However, the

327   abundance of the largest population was not consistently related to mate limitation. The

328   abundance of the dominant population in communities with mate-limited populations could be

329   higher or lower than the dominant population in the non-limited community.

330   Multiple regression analyses showed that mate search speed and search radius explained

331   approximately 90% of variation in diversity and mean abundance in the communities with mate-

332   limited sexual populations (see supplementary material). In contrast, search radius and search

333   speed explain only 1% of variation in dominance (maximum population size). In mate-searching

334   populations, the same degree of limitation could be generated with different combinations of

335   search radius and search speed. Any combination of R and V that produces a constant value of

336   VR^2 yields an equivalent probability of encountering a mate (Eq. 3).

337

338   **Discussion**

339   This study shows that the constraint of mate finding influences emergent community

340   properties, including diversity and average population size. Mate limitation strongly suppresses

341   birth rates when populations are small (Fig. 1), leading to a higher probability that sexually

342   reproducing populations will decline when rare. These discrepancies in birth rate lead to shorter

343   times to extinction in taxa that must find a mate, versus those that reproduce asexually (Fig. 2).

344   This effect is particularly strong when populations are introduced at low density, which is a

345   plausible scenario when considering newly established populations. In stochastic simulations,

346   communities consisting of asexual taxa maintained greater diversity due to a longer expected

347   persistence time of each population (Fig. 3). In the case where immigration is a linear function of

348   current diversity, expected diversity increases as MTE increases (Eq. 6). When these simulations

349   were extended to communities with heterogeneous taxa, differences in diversity were amplified,

350   because mate limitation had especially strong negative effects on taxa with already-low growth

351   rates (Fig. 4). Thus, mate limitation decreased the number of coexisting taxa, primarily by

352   excluding low abundance taxa. However, mate limitation has minimal consequences in larger

353   populations, and therefore the population size of the most abundant taxon was not related to

354   reproductive method or mate search efficacy.

355   The degree of mate limitation is a function of search ability, which is determined by

356   search radius and search speed. As either search variable (radius or speed) increases, the

357   probability of finding a mate approaches 1, indicating no limitation to the population birth rate.

358   In this case, simulation results of sexual populations with mate finding converge to those of

359   populations without limitation. This is also evident when looking at per capita and population

360   growth curves (Fig. 1). Birth rates asymptotically reach the no-limitation case as mate search

361   becomes more effective. Simulation results are qualitatively similar if instead considering a two-

362   dimensional search (see supplemental material). As in the three-dimensional scenario, mate

363   limitation led to lower diversity and rarity, but had no effect on maximum population size. When

364   searching for mates in three dimensions, the search radius has greater consequence of finding a

365   mate than the search speed. This observation suggests that a trait affecting search radius (such as

366   eyesight) might have a greater fitness effect in a 3D environment than a trait affecting search

367   speed (such as swimming velocity); an incremental increase in search radius would lead to a

368   greater effective birth rate than an increase in search speed. Finally, this study highlights the

369   utility of stochastic models for studying community structure. Deterministic models show

370   equivalent long-term dynamics for populations with the same carrying capacity, whereas these

371   stochastic models show pronounced differences. Thus, this study is concordant with prior models

372   showing that Allee effects increase extinction rates (Brassil 2001, Leibhold and Bascompte 2003,

373   Dennis et al. 2016), and further demonstrates that these population-level effects alter emergent

374   community properties.

375   Results from these models mirror empirical findings that microbial populations (with

376   asexual reproduction) tend to be high in diversity and rarity, although not distinct from other

377   communities in the dominance of abundant taxa (Locey and Lennon 2016). In these simulated

378   communities, eliminating the constraint of mate finding translated to greater diversity with a

379   higher frequency of low-abundance taxa, while the population size of the most abundant taxon

380   was unaffected (Fig. 4). Thus, allowing for mate finding generates a parsimonious explanation

381   for the community-level patterns observed in comparisons of micro and macro ecological

382   communities. Using empirical data to probe the hypothesis that mate limitation constrains

383   diversity and rarity illustrates the plausibility of this explanation. For example, one study used

384   very deep 16S amplicon sequencing to evaluate whether marine bacterial populations that

385   appeared to only be present seasonally were, instead, consistently present at abundances below

386   the usual detection limit (Caporaso et al. 2012). At a depth of approximately 11 million

387   sequences, 48% of sequences appeared only once (Caporaso et al. 2012). Similarly, another

388   study used deep sequencing of human gut samples to generate rarefaction curves illustrating how

389   many taxa were observed in response to sequencing depth. New taxa continued to be identified

390   after one million sequences were recovered (Turnbaugh et al. 2010). Even if many of the

391   observed rare taxa are the product of sequencing errors, these findings suggest persistence of

392   extremely low abundance taxa (fewer than 1 individual per million). For sexually reproducing

393   populations, these relative abundances could be prohibitively low for individuals to find mates

394   within a lifetime. Finally, asexual reproduction in microbes makes it possible for single

395   individuals to establish populations in new environments. Given the demonstrated plausibility of

396   immigration from microbial seed banks (Lennon and Jones 2011, Caporaso et al. 2012), the

397   growth rate of small populations is especially relevant for the persistence of microbial taxa.

398   Diversity is a common outcome variable in ecological studies, though there is ongoing

399   debate about how diversity is related to community function (Shade 2017). In the models studied

400   here, diversity is a byproduct of population demographics, including birth rate and mate search

401   ability. More generally, these models show that diversity can be affected by neutral and

402   stochastic processes. Subsequent empirical studies using diversity as an outcome variable might

403   also collect information about immigration and extinction rates to determine whether diversity

404   reflects these processes. For example, surveys of human-associated microbial communities have

405   found variation in diversity across body sites (Caporaso et al. 2011, Huttenhower et al. 2012).

406   Gut and oral bacterial communities are especially diverse (Huttenhower at el. 2012), but these

407   habitats could conceivably have higher immigration rates than other body sites due to daily

408   introduction of bacteria within food (David et al. 2014). Similarly, a recent study found little

409   evidence that fungi can persist within the healthy human gut, but still identified hundreds of

410   fungal taxa in stool samples (Auchtung et al. 2018). The high diversity of fungi in the human gut,

411   in spite of their inability to colonize this habitat, was attributed to persistent immigration of fungi

412   on ingested foods (Auchtung et al. 2018). These studies, coupled with the modeling results

413   presented here, demonstrate how diversity could change independently of community function.

414   Mathematical models serve as an unbiased line of inference for linking mechanisms to

415   emergent community properties. In the context of comparing community structures of microbial

416   communities and macro-organisms, the differing empirical methods used to study communities

417   at these two scales can generate spurious patterns. Thus, it is hard to discern which findings are

418   true distinctions between micro and macro communities, and which are artifacts of methodology.

419   For instance, DNA sequence similarity is often used to define microbial taxa, whereas macro

420   organisms are generally identified using direct observations. The differences in error rate and

421   detection limits between these two methods could also explain the higher diversity and rarity in

422   microbial communities. Several steps in the workflow of generating 16S amplicon data,

423   including variation in sample processing and sequencing errors, can generate observations of

424   artifactual rare taxa (Fouhy et al. 2016). Furthermore, macro-organisms are often identified using

425   morphological characteristics, but many more taxa can be differentiated if instead using DNA

426   sequencing methods (Fontaneto et al. 2009). Thus, if methodology is a confounding factor when

427   comparing communities, there is uncertainty about whether observed differences are spurious

428   due to sampling bias. Theoretical studies can therefore reinforce empirical findings by

429   determining, with unbiased methodology, whether an identified mechanism can reproduce

430   observed community structure. These modeling results indicate that there are expected

431   differences in diversity and average population size when comparing communities consisting of

432   taxa that reproduce sexually versus asexually.

433

Acknowledgements

439 **References**

440 Allen LJS, Allen EJ. A comparison of three different stochastic population models with regard to

441 persistence time. *Theoretical Population Biology* 2003; **64**: 439–449.

442 Allen LJS, Fagan JF, Högnäs G, Fagerholm H. Population extinction in discrete-time stochastic

443 population models with an Allee effect. *Journal of Difference Equations and*

444 *Applications* 2005, **11**, 273-293.

445 Anderson DF, Kurtz TG. Continuous Time Markov Chain Models for Chemical Reaction

446 Networks. *Design and Analysis of Biomolecular Circuits*. 2011. Springer, New York,

447 NY, pp 3–42.

448 Assaf M, Meerson B. Extinction of metastable stochastic populations. *Physical Reviews E* 2010;

449 **81**: 021116.

450 Auchtung TA, Fofanova TY, Stewart CJ, Nash AK, Wong MC, Gesell JR, et al. Investigating

451 Colonization of the Healthy Adult Gastrointestinal Tract by Fungi. *mSphere* 2018; **3**:

452 e00092-18.

453 Boukal DS, Berec L. Single-species Models of the Allee Effect: Extinction Boundaries, Sex

454 Ratios and Mate Encounters. *Journal of Theoretical Biology* 2002; **218**: 375–394.

455 Brassil CE. Mean time to extinction of a metapopulation with an Allee effect. *Ecological*

456 *Modelling* 2001; **143**: 9–16.

457 Caporaso JG, Lauber CL, Costello EK, Berg-Lyons D, Gonzalez A, Stombaugh J, et al. Moving

458 pictures of the human microbiome. *Genome Biology* 2011; **12**: R50.

459 Caporaso JG, Paszkiewicz K, Field D, Knight R, Gilbert JA. The Western English Channel

460 contains a persistent microbial seed bank. *The ISME Journal* 2012; **6**: 1089–1093.

461   Consortium THMP, Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, et al.

462   Structure, function and diversity of the healthy human microbiome. *Nature* 2012; **486**:

463   207–214.

464   Courchamp F, Clutton-Brock T, Grenfell B. Inverse density dependence and the Allee effect.

465   *Trends in Ecology & Evolution* 1999; **14**: 405–410.

466   David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly

467   and reproducibly alters the human gut microbiome. *Nature* 2014; **505**: 559–563.

468   Dennis B. Allee effects: population growth, critical density, and the chance of extinction.

469   *Natural Resource Modeling* 1989; **3**: 481-538.

470   Dennis B. Allee Effects in Stochastic Populations. *Oikos* 2002; **96**: 389–401.

471   Dennis B, Assas L, Elaydi S, Kwessi E, Livadiotis G. Allee effects and resilience in stochastic

472   populations. *Theoretical Ecol* 2016; **9**: 323–335.

473   Dobramysl U, Mobilia M, Pleimling M, Täuber UC. Stochastic population dynamics in spatially

474   extended predator–prey systems. *Journal of Physics A: Mathematical and Theoretical*

475   2018; **51**: 063001.

476   Fontaneto D, Kaya M, Herniou EA, Barraclough TG. Extreme levels of hidden diversity in

477   microscopic animals (Rotifera) revealed by DNA taxonomy. *Molecular Phylogenetics*

478   *and Evolution* 2009; **53**: 182–189.

479   Fouhy F, Clooney AG, Stanton C, Claesson MJ, Cotter PD. 16S rRNA gene sequencing of mock

480   microbial populations- impact of DNA extraction method, primer choice and sequencing

481   platform. *BMC Microbiology* 2016; **16**: 123.

482   Gascoigne J, Berec L, Gregory S, Courchamp F. Dangerously few liaisons: a review of mate-

483   finding Allee effects. *Population Ecology* 2009; **51**: 355–372.

484   Gerritsen J. Sex and Parthenogenesis in Sparse Populations. *The American Naturalist* 1980; **115**:

485   718–742.

486   Gerritsen J, Strickler JR. Encounter Probabilities and Community Structure in Zooplankton: a

487   Mathematical Model. *Journal of the Fisheries Research Board of Canada* 1977; **34**: 73–

488   82.

489   Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *The journal of physical*

490   *chemistry*, 1977; **81**: 2340-2361.

491   Gillespie DT. Stochastic Simulation of Chemical Kinetics. *Annual Review of Physical Chemistry*

492   2007; **58**: 35–55.

493   Hansen GJA, Carey CC. Fish and Phytoplankton Exhibit Contrasting Temporal Species

494   Abundance Patterns in a Dynamic North Temperate Lake. *PLOS ONE* 2015; **10**:

495   e0115414.

496   Hutchinson GE. The paradox of the plankton. *The American Naturalist* 1961; **882**: 137-145.

497   Lande R. Risks of Population Extinction from Demographic and Environmental Stochasticity

498   and Random Catastrophes. *The American Naturalist* 1993; **142**: 911–927.

499   Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of

500   dormancy. *Nature Reviews Microbiology* 2011; **9**: 119–130.

501   Ley RE, Peterson DA, Gordon JI. Ecological and Evolutionary Forces Shaping Microbial

502   Diversity in the Human Intestine. *Cell* 2006; **124**: 837–848.

503   Liebhold Andrew, Bascompte Jordi. The Allee effect, stochastic dynamics and the eradication of

504   alien species. *Ecology Letters* 2003; **6**: 133–140.

505   Locey KJ, Lennon JT. Scaling laws predict global microbial diversity. *PNAS* 2016; **113**: 5970–

506   5975.

507   Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. *Nature Reviews*

508   *Microbiology* 2015; **13**: 217–229.

509   MacArthur RH, Wilson EO. The Theory of Island Biogeography. 1967. Princeton University

510   Press.

511   Meyer KM, Memiaghe H, Korte L, Kenfack F, Alonso A, Bohannan BJM. Why do microbes

512   exhibit weak biogeographic patterns? *The ISME Journal* 2018.

513   Nemergut DR, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, et al. Patterns and

514   Processes of Microbial Community Assembly. *Microbiology and Molecular Biology*

515   *Reviews* 2013; **77**: 342–356.

516   Odum HT, Allee WC. A Note on the Stable Point of Populations Showing Both Intraspecific

517   Cooperation and Disoperation. *Ecology* 1954; **35**: 95–97.

518   Purvis A, Gittleman JL, Cowlishaw G, Mace GM. Predicting extinction risk in declining species.

519   *Proceedings of the Royal Society of London B: Biological Sciences* 2000; **267**: 1947–

520   1952.

521   Shoemaker WR, Locey KJ, Lennon JT. A macroecological theory of microbial biodiversity.

522   *Nature Ecology & Evolution* 2017; **1**: 0107.

523   Shade A. Diversity is the question, not the answer. *The ISME Journal* 2017; **11**: 1–6.

524   Stephens PA, Sutherland WJ, Freckleton RP. What Is the Allee Effect? *Oikos* 1999; **87**: 185–

525   190.

526   Stephan T, Wissel C. Stochastic extinction models discrete in time. *Ecological Modelling* 1994;

527   **75–76**: 183–192.

528   Sun G-Q. Mathematical modeling of population dynamics with Allee effect. *Nonlinear*

529   *Dynamics* 2016; **85**: 1–12.

530   Turnbaugh PJ, Quince C, Faith JJ, McHardy AC, Yatsunenko T, Niazi F, et al. Organismal,

531   genetic, and transcriptional variation in the deeply sequenced gut microbiomes of

532   identical twins. *PNAS* 2010; **107**: 7503–7508.

533   Volterra V. Population growth, equilibria, and extinction under specified breeding conditions: a

534   development and extension of the theory of the logistic curve. *Human Biology* 1938; **10**:

535   1–11.

536   Xu Z, Knight R. Dietary effects on human gut microbiome diversity. *British Journal of Nutrition*

537   2015; **113**: S1–S5.