

6-14-2012

Classification of Phase Transition Behavior in a Model of Evolutionary Dynamics

Dawn Michelle King

University of Missouri-St. Louis, dmkxtd@mail.umsl.edu

Follow this and additional works at: <http://irl.umsl.edu/thesis>

Recommended Citation

King, Dawn Michelle, "Classification of Phase Transition Behavior in a Model of Evolutionary Dynamics" (2012). *Theses*. 274.
<http://irl.umsl.edu/thesis/274>

This Thesis is brought to you for free and open access by the Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Theses by an authorized administrator of IRL @ UMSL. For more information, please contact marvinh@umsl.edu.

Classification of Phase Transition Behavior in a
Model of Evolutionary Dynamics

by

Dawn M. King

B.S., Biophysics, University of Southern Indiana, 2009

A Thesis

Submitted to the Graduate School of the

University of Missouri – St. Louis

In partial fulfillment of the requirements for the degree

Master of Science

in

Physics

June, 2012

Advisory Committee

Sonya Bahar
Chairperson

Ricardo Flores

Nevena Marić

Abstract

Amongst the scientific community, there is consensus that evolution has occurred; however, there is much disagreement about how evolution happens. In particular, how do we explain biodiversity and the speciation process? Computational models aid in this study, for they allow us to observe a speciation process within time scales we would not otherwise be able to observe in our lifetime. Previous work has shown phase transition behavior in an assortative mating model as the control parameter of maximum mutation size (μ) is varied. This behavior has been shown to exist on landscapes with variable fitness (Dees and Bahar, 2010), and is recently presented in the work of Scott et al. (submitted) on a completely neutral landscape, for bacterial-like fission as well as for assortative mating. Here I investigate another dimension of the phase transition. In order to achieve an appropriate 'null' hypothesis and make the model mathematically tractable, the random death process was changed so each individual has the same probability of death in each generation. Thus both the birth and death processes in each simulation are now 'neutral': every organism has not only the same number of offspring, but also the same probability of being randomly killed. Results show a continuous nonequilibrium phase transition for the order parameters of the population size and the number of clusters (analogue of species) as the random death control parameter δ is varied for three different mutation sizes of the system. For small values of μ , the transition to the active state of survival happens at a small critical value of δ ; in contrast, for larger μ , the transition happens later – suggesting a robustness of the system with increased mutation ability.

Introduction

Phase transitions are an inherent characteristic of nature. The most familiar examples occur in the realm of physics with the classical thermodynamic equations of state and their relation to the physical changes of a substance. Pressure, volume, and temperature define a substance as a gas, liquid, solid, or plasma; while, energy, entropy, and enthalpy define the processes that bring a system into equilibrium with its environment. Any changes to the state of the system, such as increasing or decreasing the temperature, can lead to a sharp change in the physical properties of the system, which is characteristic of a phase transition. Even more, the equations of state for these systems are reversible. This dynamic reversibility allows for these transitions to be classified as *equilibrium* phase transitions. Equilibrium phase transitions are not to be confused with a system that reaches a steady state with its environment; rather, they are systems that can seamlessly transition from one state to another no matter what the direction of travel. For example, water can freeze, and ice can melt; hence there is an intrinsic reversibility of the two states, and the system obeys the so-called *principle of detailed balance* (Henkel et al., 2008).

Statistical mechanics is another way to view natural phase transitions. Instead of the continuous equations of state, statistical mechanics examines the macroscopic behavior that occurs from interactions among the individual microscopic components. The famous Ising model is probably the most studied model in the literature of the statistical mechanics of phase transitions. It was developed by Ernst Ising in 1925 and describes ferromagnetism as the energy of interaction between adjacent spins on a line.

It was almost another twenty years before the two dimensional model was solved exactly by Lars Onsager in 1944 -- this was a great feat, since the three dimensional case has yet to be solved (Ódor, 2002; Solé, 2011), and the 1-D case does not exhibit phase transition behavior (Cipra, 1987). The phase transition behavior of the 2-D model is shown to occur between two qualitative states of magnetization where disorder and order amongst the spins represent the unmagnetized and magnetized states, respectively. The critical point of this phase transition marks the coexistence of the two states, i.e., 'ordered structures exist at every length scale' (Yeomans, 1992).

Phase transition behavior can be characterized by the discontinuities or divergences of mean field parameters at the critical point. In a first order phase transition, there are discontinuities in the first derivatives of a variable describing some property of the system at the critical point, while a second order transition has a discontinuity/divergence in the second derivative. Consider the thermodynamic first law for a magnetic system:

$$dU = TdS - MdH - PdV, \quad (1)$$

where dU , dS , dH , and dV represent changes in the internal energy, entropy, magnetic field, and volume, respectively, and T , M , and P are temperature, magnetization, and pressure, respectively. The change in free energy of a system is:

$$d\mathcal{F} = dU - TdS - SdT. \quad (2)$$

Substituting (2) into (1) and assuming the volume and temperature are held constant, the free energy then becomes

$$d\mathcal{F} = -MdH, \quad (3)$$

and the magnetization is then

$$M = - \left(\frac{\partial \mathcal{F}}{\partial H} \right)_{V,T}. \quad (4)$$

The second derivative of the free energy is equal to the isothermal susceptibility

$$\chi_T = - \left(\frac{\partial^2 \mathcal{F}}{\partial H^2} \right)_{V,T}. \quad (5)$$

The magnetization M in this case would serve as the order parameter of the system, while the magnetic field would serve as the control parameter since it is the parameter that is varied. The critical point (H_c) marks the transition between the two different phases of the system – in this case, magnetized for $H > H_c$ and unmagnetized for $H < H_c$. If there exists a discontinuity in the first derivative of the free energy that describes the order parameter M (Eq. 4), then the system is said to be of first order or exhibit a first order phase transition (a discontinuous jump in M at H_c). If the discontinuity exists in the second derivative of the free energy (Eq. 5), the system is said to be of second order, and the magnetization M will exhibit a continuous phase transition. In this case, a fluctuating state of the order parameter rather than a discontinuous jump as the control parameter is varied marks H_c . At H_c scale-free behavior – or, as mentioned previously, the existence of order at all length scales – of properties, such as the size of clusters

created by adjacent spins that line up together in space, will occur. But what happens when we leave this well-defined domain of classical thermodynamics? How do we classify all the complex dynamic phase transitions that occur in nature?

As noted by Yeomans (1992), the terminology of 'second order phase transition' is a relic of Ehrenfest, who classified transitions based on discontinuities in their derivatives rather than divergences. It is proposed to use continuous, higher order, or critical to describe anything other than a first order transition (Yeomans, 1992). This distinction is important because, while it is true that derivatives can be taken from the equations of state or from the partition function describing many systems in statistical mechanics, this cannot be done for many models that exhibit nonequilibrium phase transition behavior. For many nonequilibrium transitions do not emerge from continuous equations where derivatives exist; rather, models of the nonequilibrium type tend to be phenomenological, agent-based, and involve Markov chains and/or random walk processes (Henkel et al., 2008). Irreversibility is characteristic of nonequilibrium phase transitions because of the so-called 'absorbing states' the systems can fall into. An absorbing state is one from which the system cannot escape; thus, these transitions are irreversible in such a way that the principle of detailed balance is not obeyed. For example, if we consider a phase transition between a surviving and an extinction state of a population, the population can never reverse back to the surviving state once it has gone extinct; extinction is thus an absorbing state. For this reason nonequilibrium phase transitions are sometimes called absorbing phase transitions. The terms are interchangeable.

Characterization of critical phase transition behavior continues with the emergence of complexity and universality, for scientists are recognizing that ‘the road from disorder to order is maintained by powerful forces of self-organization ... paved with power laws’ (Barabási, 2003). Both equilibrium and nonequilibrium continuous phase transitions can be characterized by the *critical* exponents that define the scaling behavior of the system near the critical point. The scaling behavior near the critical point is described by a power law distribution. These unique exponents can be used to determine the *universality class* of a phase transition (Ódor, 2002). Events distributed according to a power law are said to be scale free because many tiny events occur with only a few large ones; there is no characteristic scale. For example, since, on average, there are approximately 1,000,000 earthquakes of magnitude less than 3 on the Richter scale annually, and typically only about one above a Richter scale value of eight, it is thought that the current state of the tectonic plates exists at criticality (Buchanan, 2000). Other systems for which continuous phase transitions have been characterized include (but are not limited to): catalytic chemical reactions, mutation rates in viruses, epidemic spreading, changes in vegetation due to climate (Solé, 2011), social networks, stock market crashes, the world wide web (Barabási, 2003), earthquakes, solar flare occurrences, and the evolution of biological systems (Ódor, 2002). Thus, the study of phase transition behavior gives the ability to group a broad range of systems into a particular universality class based on the specific behavior of the system at the critical point.

Many systems which exhibit continuous phase transitions are also complex systems, to the extent that they involve the study of the phenomena which emerge from a collection of objects (Johnson, 2007). In other words, it is from the manner in which individual objects interact that collective behavior emerges, with the ensemble of individuals exhibiting behavior as a whole unit. For example, the dynamics of a traffic jam are heavily dependent on the individual choices of people and the space provided on the freeway. During rush hour, individuals make independent choices to either drive on the freeway or to take the side streets. If enough individuals choose to take the side streets, the space on the freeway does not completely fill, and a traffic jam will not occur; however, if enough individuals make the choice of the freeway, and those choices surpass the critical threshold of space available on the road, the whole system slows or stops and there is a traffic jam. The traffic jam cannot be predicted because it is driven (no pun intended) by individual people's driving choices; it is an emergent phenomenon, resulting from the collective behavior of the individual drivers (Johnson, 2007).

With the recent rise in the study of emergent phenomena, and the seemingly eloquent description of nature it inspires, it seems only natural to look at evolution from such a bottom-up, collective approach. But where is the bottom? If we look at the biological classification scheme for taxonomic ranks of life, species are at the bottom (Campbell, 2005). But speciation is only a snapshot of the evolutionary history of life, for the time line of evolution began when the first replicator emerged from the primordial soup (Dawkins, 2006). These replicators developed protective coats, coated replicators emerged as cells, cells formed organisms, and then organisms grouped into

species. It is only after the emergence of multicellular individuals that one can begin to think of traditional Darwinian natural selection and speciation, let alone the important and controversial concept of group selection.

Thus, when thinking in terms of “how bottom-up to go” in the study of evolution, the ‘unit’ of selection (what is actually being selected for) is important. According to Richard Dawkins, organisms are NOT the unit of selection (Dawkins, 2006); rather, selection occurs at the level of the genes (the replicators inside the cell). Other scientists such as Leo Buss and Stephen Hubbell generally agree that there are multiple levels of selection, yet Buss (1987) focuses on multicellular evolution while Hubbell (2001) focuses on the emergence of patterns of biodiversity. Mikhailov and Calenbuhr (2002) address the ability to see evolution on multiple levels by saying, “Fortunately, in most cases the elements interact not fully expressing their complexity. Therefore, they can be described as *automata* with a limited repertoire of responses and relatively simple effective internal dynamics.” This suggests that it is sufficient to understand how the individual components drive the system to its collective behavior rather than include all the internal complexities of each individual; from this point of view, one can take a physics-based approach, and deal only with the simplest possible canonical organisms in order to investigate the emergence of collective behavior. For example, one could investigate the behavior of immensely complex organisms that are reduced to characterization by only a few simple rules, such as how they reproduce, mutate, and die. This has been done recently with models of collective animal behavior, but a similar

approach can be taken in a simple evolutionary model with regard to the formation of clustering of organisms into "species".

While evolution by natural selection (on individual organisms) is the standard view of Darwinian evolution, 'neutral' evolution is a relatively new idea that inspires much debate. Natural selection has three main tenets:

1. Competition for resources in the natural environment.
2. Variation of traits.
3. Heritability of traits.

Having variation of traits means that some individuals will be able to compete better for resources, and thus have a better chance of surviving and passing along their traits to the next generation. Therefore, individuals that are more 'fit' have traits that are well adapted to the environment. In biology, fitness is a measure to describe reproductive success. Thus, natural selection is predicated on the assumption that organisms will have different fitnesses based on adaptability to the environment, and that those with greater fitness will have a greater chance of survival. But what if criteria 1, 2 and 3 are present and variation doesn't buy the organisms any improvement in fitness? Essentially, this is a null condition. In this case, would populations still survive and speciate? This type of null model was first proposed by Motoo Kimura (1968, 1983) who suggested the occurrence of speciation due to random genetic drift with his 'Neutral Theory of Molecular Evolution.' A different aspect of this null condition was introduced nearly eighteen years later with an investigation of ecological drift by Steven Hubbell (2001). As implied in the title of Hubbell's book, 'The Unified Neutral Theory of

Biodiversity and Biogeography,' ecological drift occurs under *neutral* conditions, in which every individual in the population experiences the same conditions. Neutral theory is still the subject of much controversy within the ecological community because it implies that every individual is just as fit as the next. Even if neutral conditions rarely occur in nature, as some scientists who strongly disagree with Hubbell maintain, neutral theory can still serve as a useful 'null hypothesis', which is how Hubbell intended neutral theory to be used (Hubbell, 2001). It is from that perspective that a neutral model is presented here, in the context of an agent-based simulation of evolutionary dynamics.

Initial inspiration for this work comes from an earlier implementation of the model (Dees and Bahar, 2010), where each organism had a variable 'fitness' defined by a randomized, rugged fitness landscape. In the traditional Darwinian idea of evolution, the higher the fitness, the higher the reproductive success, and thus the more natural selection favors that particular organism's survival. (Or, in Richard Dawkins's "gene's-eye" view, the more natural selection favors the genes which lead to the expression of the particular phenotypic trait as the basis for which the organism experiences selection.) The Dees and Bahar model defined individual organisms by their position on a two dimensional landscape with the axes representing arbitrary phenotypic traits. A phenotype is a trait (such as hair or eye color) resulting from the expression of a gene or a collection of genes. So in essence, the landscape represents a phenospace and not a physical space. Since the landscape does not pose any geographical barriers between individuals of a population, speciation is said to occur in sympatry – without geographical separation. Organisms reproduced based on an assortative mating

algorithm, selecting mates nearby in the phenotype space. Phase transition-like behavior was shown as a control parameter, mutability (μ), was varied for both the population size and the number of clusters – with a cluster being the analogue of a species. The mutability μ represents how far an offspring can be from its parent and can be considered as biologically relevant since no offspring is an identical copy of its parent.

The model was further developed by Scott et al. (submitted), who took the neutral approach to fitness, in which each individual produces the same number of offspring in each generation. In this case, each parent produces two offspring. This version of the model also investigated two new mating schemes: a control case of random mating, and the reproductive strategy of bacteria-like asexual fission. One of the most striking results is that phase transition behavior still exists as μ varied for the assortative mating reproduction scheme, even without the noise of natural selection. The fission reproductive scheme also showed phase transition behavior; however, the random mating did not, typically yielding only one giant component or cluster throughout the simulation. This is consistent with the Erdős-Renyi model of network theory which predicts the emergence of a giant component (or cluster) from a randomly connected network (Barabási, 2003).

In both realizations above, after the populations reproduce, the offspring go through a series of removal/death processes. There is an overpopulation limit set to eliminate any offspring that are too phenotypically close, a uniformly distributed random elimination of individuals of up to 70 percent, and removal of any organism

produced outside the boundary of the landscape. The first major change I have implemented in the present work was to make the model more mathematically tractable so that particular properties of the phase transition behavior may be parsed out. Here, instead of removing a random number of individuals chosen at random from the entire population, in each generation, *each individual in the population has the same probability of death*. Effectively, this allows for both the death process as well as the birth process to be truly deemed 'neutral': every organism has not only the same number of offspring, but also the same probability of being randomly killed. This simplifies the model compared to the earlier versions, in which the percentage of organisms killed varied from one generation to another (with a maximum death rate of 70%), so that an organism might have a different probability of survival from its parent(s). Secondly, previous versions investigated the transitions occurring for the order parameters of population size and number of clusters by means of varying the control parameter μ . In this work, I investigate *the phase transition which ensues as another parameter, the individual death probability δ , is varied*. The results developed below clearly show the presence of a continuous phase transition as δ is changed, in addition to the continuous transition already demonstrated along the dimension of the mutability μ .

Methods

The simulated environment is a two dimensional phenospace, or morphospace (these terms are used interchangeably), which is not to be confused with a physical

space. The phenospace simply utilizes a description of individuals based on phenotypic traits rather than where they exist in a physical or geographical space. Thus, the location of each organism in the phenospace can be loosely interpreted as specifying its phenotypic traits (external characteristics). In the simulations shown here, the phenospace was a finite, continuous, two-dimensional space, with 45 arbitrary units along each axis. Each simulation started with an initial population of 300 individuals and was run for 2000 generations, unless the population became extinct.

In contrast to previous work, the new dynamics incorporated here include an individual probability δ of dying, rather than the entire population being subject to some percent chance of dying, with that percentage varying randomly from one generation to another. The simulations take place on a neutral fitness landscape, where each organism produces the same number of offspring – in this case the fitness is two. The three main steps of the simulation involve: 1. A reproduction scheme – random, assortative, or asexual fission. 2. Production of offspring – to be dispersed based on mutability μ . 3. Death processes – which include the removal of parents, the imposition of an overpopulation limit, random probability of death, and boundary conditions.

Mating and Reproductive Strategy

Three mating schemes (random, assortative, and fission) were compared. For each generation, every individual in the population chooses a mate, except for the case of bacterial-like fission where the individual simply splits within a defined space. The

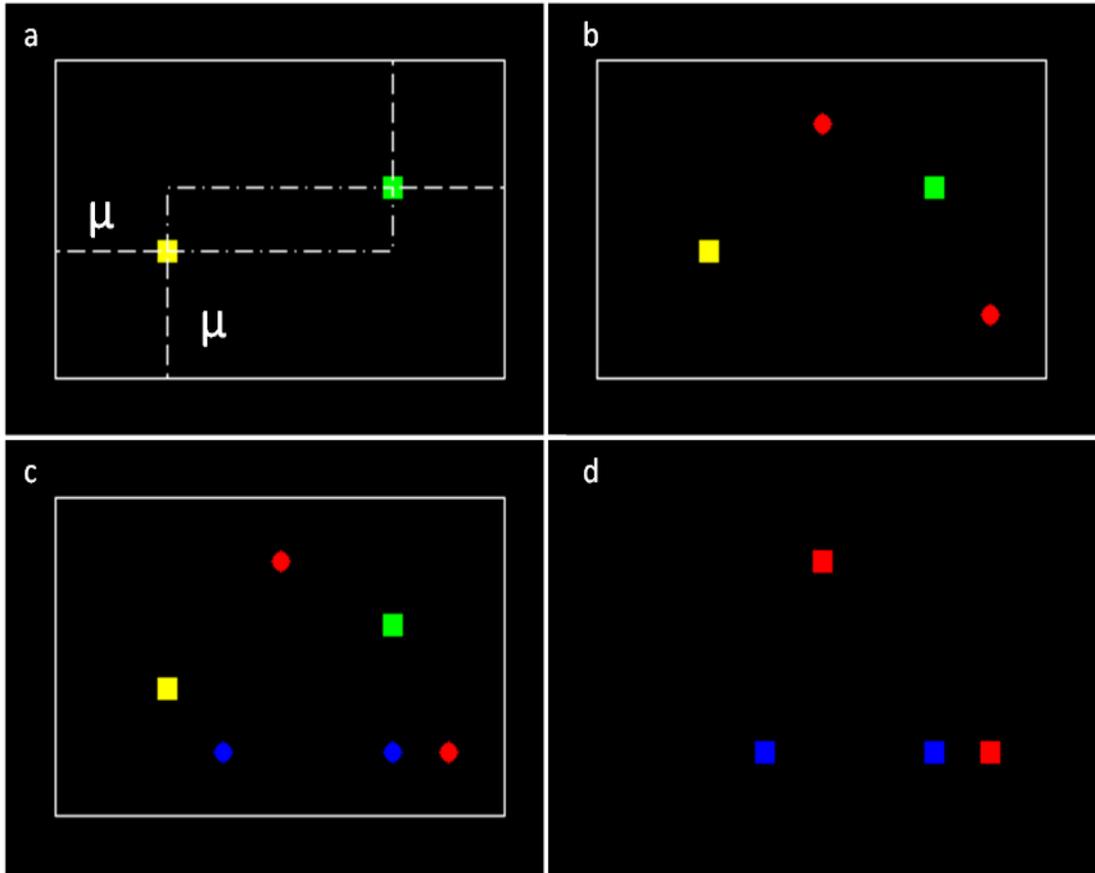


Figure 1 Schematic diagram for assortative mating. Parents are labeled as squares and offspring as circles. (a) A reference organism (yellow) selects its nearest neighbor (green) as a mate. The offspring are distributed in an area defined by the locations of the two parent organisms, extended by the mutability μ . (2) Generation of yellow's offspring organisms. (3) Generation of green's offspring. (This assumes a case in which yellow parent organism is also the nearest neighbor of the green organism. Note that this will not always be the case, and thus mating pairs will not necessarily be "monogamous"). (4) After every parent has mated (each acting once as the reference organism), all parents are removed, leaving their offspring to act as parents for the next generation.

inherent difference between random and assortative mating is the spatial restriction imposed by the assortative scheme. The biological rules of assortative mating (as mentioned in the Introduction) are followed by calculating the shortest Euclidean distance between two individuals in the population; thus, the most phenotypically similar individuals will always mate. The individuals using the random mating strategy

will choose mates at random. This leads to a variable distribution of phenotypic distance between mates. For each mating strategy, every individual produces two offspring. The placement of offspring in phenospace depends on μ , which defines an area around the parent(s) in which the offspring can be placed, and then distributes them within that area at random (illustrated in Figure 1 for assortative mating). For assortative and random mating, μ extends along the x and y axis for each parent, thus creating a reproduction area that is representative of both parents. For fission, reproduction occurs in an area of $2\mu \times 2\mu$, with the parent organism at the center.

Elimination

After each generation reproduces, the parent generation is eliminated (Figure 1d), and the offspring undergo three further elimination processes that occur in the order presented. The first controls how phenotypically close organisms can be to each other (in other words, an overpopulation limit) and removes one of any two individuals within a measured distance of 0.25 units on the phenospace. Death due to an overpopulation limit can be mathematically represented as a coalescence process, and can be considered as biologically relevant because effectively it prevents hybridization between two reproducing individuals. The overpopulation limit can also be viewed as a schematic representation of the competition for resources that would occur between phenotypically similar organisms (birds with the same size beak competing for the same food resources) located near each other in a physical space. The second process is the random removal of offspring, implemented by giving every individual in

the population the same probability of removal δ , hence a neutral death process. The final process is the elimination of any individual who exceeds the boundary of the phenotype space. After these death processes have been applied, the remaining offspring serve as parents for the future generation.

Clustering

Clusters were determined in accordance with the "biological species concept", i.e., species defined by reproductive isolation. A cluster seed was made by a closed

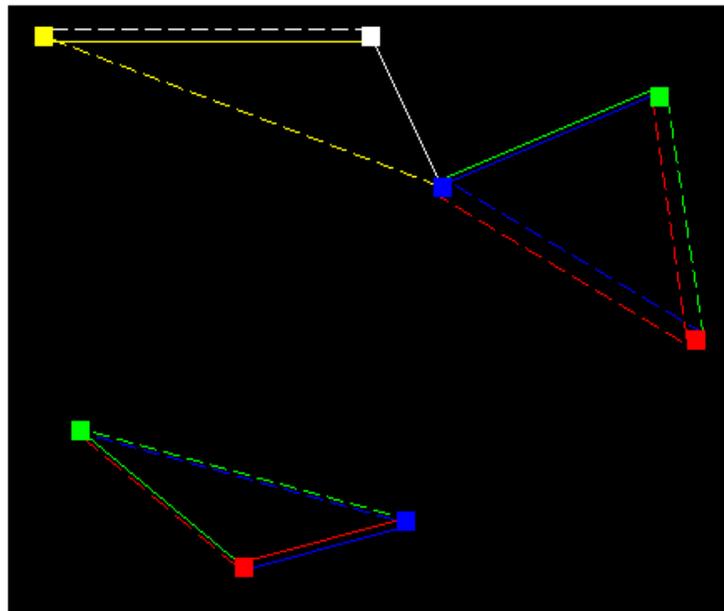


Figure 2 Schematic representation of the formation of reproductively isolated clusters. This algorithm is used for both the assortative mating and the fission model. The nearest organism to a reference organism is its mate (solid lines). The second nearest organism to the reference organism is its alternate (dashed lines). Lines are colored to indicate the mate and alternate mate of the correspondingly colored reference organism; for example, the white organism's mate is the blue organism, and its alternate is the yellow organism.

group of three organisms – a reference organism, its mate, and its second nearest

neighbor, also described as its “alternate” mate. An iterative process determined whether organisms within a cluster seed were listed in other cluster seeds, which led to the formation of groups of organisms composed of connected cluster seeds that formed a closed group. This closed set is analogous to a species, as mentioned previously, defined by reproductive isolation. The clustering algorithm is represented schematically in Figure 2. The fission model used the same algorithm as the assortative mating scheme, but slightly modified so that the previously defined “mate” was the most phenotypically similar organism. Likewise, the second nearest neighbor was the second most phenotypically similar to the reference organism. Clustering in the random mating model was determined by first identifying a cluster seed, as in the assortative mating model, but instead of a second nearest neighbor as an "alternate" mate, the alternate was, as with mate selection in this model, chosen at random.

Results

Examples from typical runs of the simulation are illustrated as snapshots at 2000 generations in Figures 3 and 4, for assortative mating and bacteria-like fission, respectively. The dots indicate the general population of individuals on the phenospace, while representative clusters are shown in red, white, yellow, purple and blue. The values of $\delta = 0.23, 0.38, \text{ and } 0.43$ for the assortative case, and $\delta = 0.26, 0.40, \text{ and } 0.44$ for the fission case, were chosen because they represent an approximation of the critical value, δ_c , at which the transition from extinction to survival occurs, measured using population size as an order parameter, at $\mu = 0.30, 0.60, \text{ and } 0.90$ respectively. $\delta = 0.20$ is representative of a survival state for each μ value shown. Figures 5 and 6 show

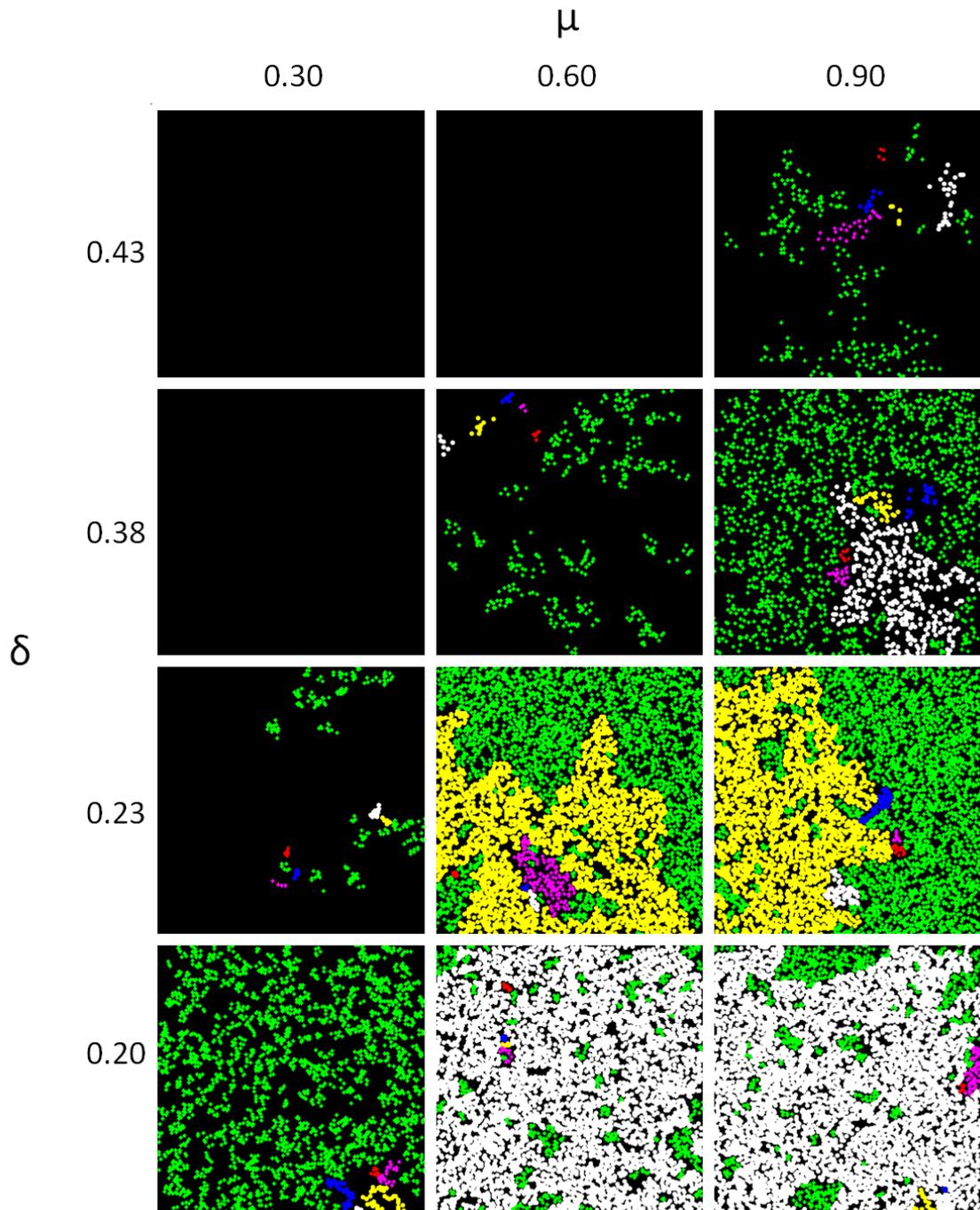


Figure 3 Clustering for assortative mating on a 45 x 45 landscape at 2000 generations. Individuals are represented by dots with example clusters highlighted in red, white, yellow, purple and blue. Approximate critical values of δ are 0.23, 0.38, 0.43 for $\mu = 0.30, 0.60, 0.90$, respectively. $\delta = 0.20$ lies within the survival regime of each μ .

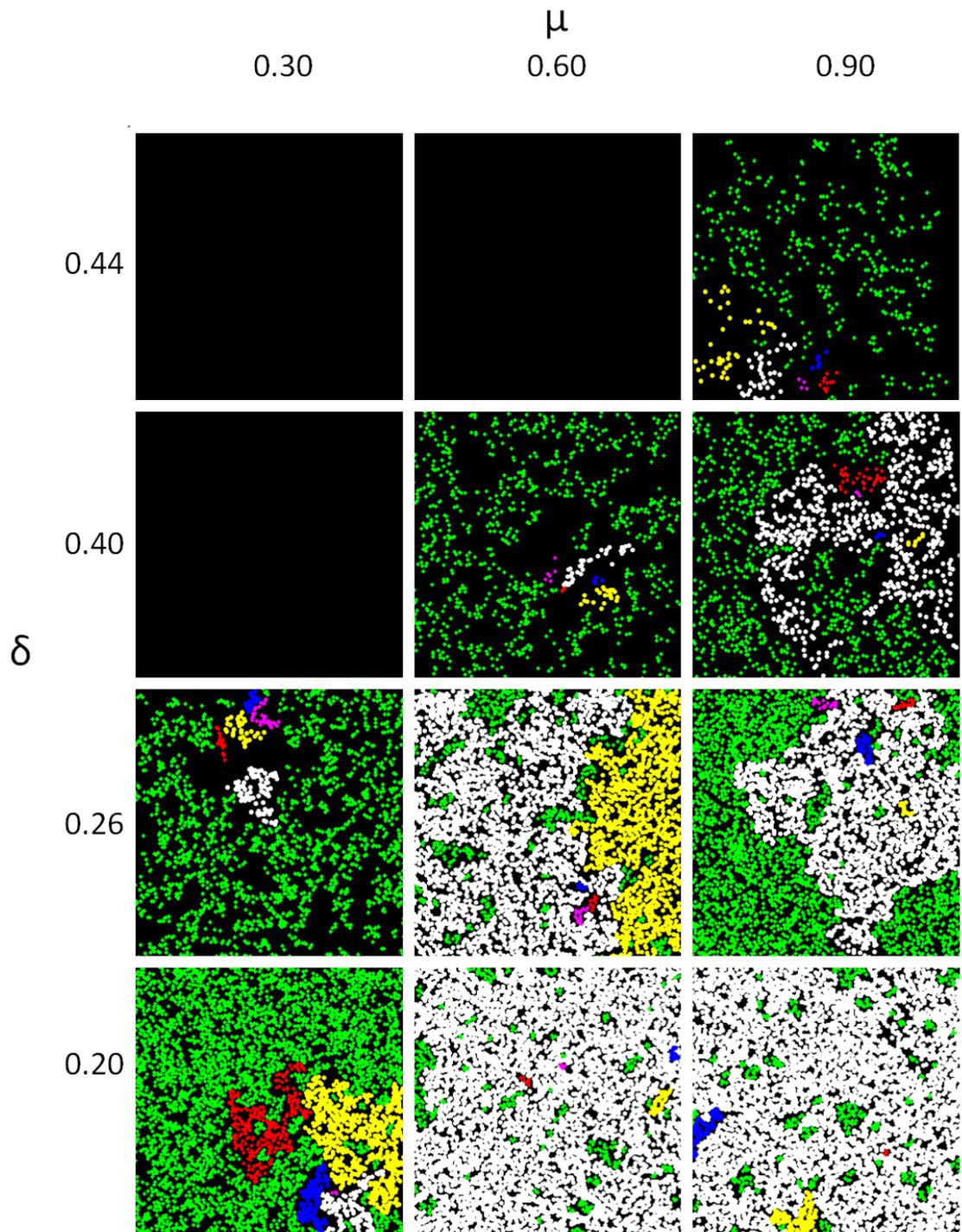


Figure 4 Clustering for bacteria-like fission on a 45×45 landscape at 2000 generations. Individuals are represented by dots with example clusters highlighted in red, white, yellow, purple, and blue. Approximate critical values of δ are 0.26, 0.40, 0.44 for $\mu = 0.30, 0.60, 0.90$, respectively. $\delta = 0.20$ lies within the survival regime of each μ .

Population		
μ	Assortative Mating δ_c	Fission δ_c
0.30	0.23	0.26
0.60	0.38	0.40
0.90	0.43	0.44

Number of Clusters		
μ	Assortative Mating δ_c	Fission δ_c
0.30	0.23	0.27
0.60	0.38	0.40
0.90	0.43	0.44

Table 1 Values of δ corresponding to the peak in standard deviation in Figures 5b - 8b. These values represent an estimate of δ_c for each mating scheme and value of μ . The values of δ_c match, with respect to each mating scheme and order parameter, except for $\mu = 0.30$ in the fission model.

phase transition curves of the population size as the control parameter δ is varied at $\mu = 0.03, 0.60, \text{ and } 0.90$; Figures 7 and 8 show phase transition curves for the number of clusters. Figures 5b-8b show a sharp rise in the standard deviation that indicates an estimated value of δ_c .

The estimated values of δ_c are shown in Table 1. The value of δ_c is the same for both order parameters (number of clusters and population size), with the exception of the fission model at $\mu = 0.30$, where the order parameter of population size has $\delta_c = 0.26$ and, for the number of clusters, $\delta_c = 0.27$. Furthermore, it can be seen from Figures 3 and 4, as well as Figures 5 and 6, that, as μ increases, the population is able to survive for larger values of δ , i.e., δ_c shifts as a function of μ . The population size transition becomes significantly less sharp as μ increases. This indicates that there might be

Assortative Mating

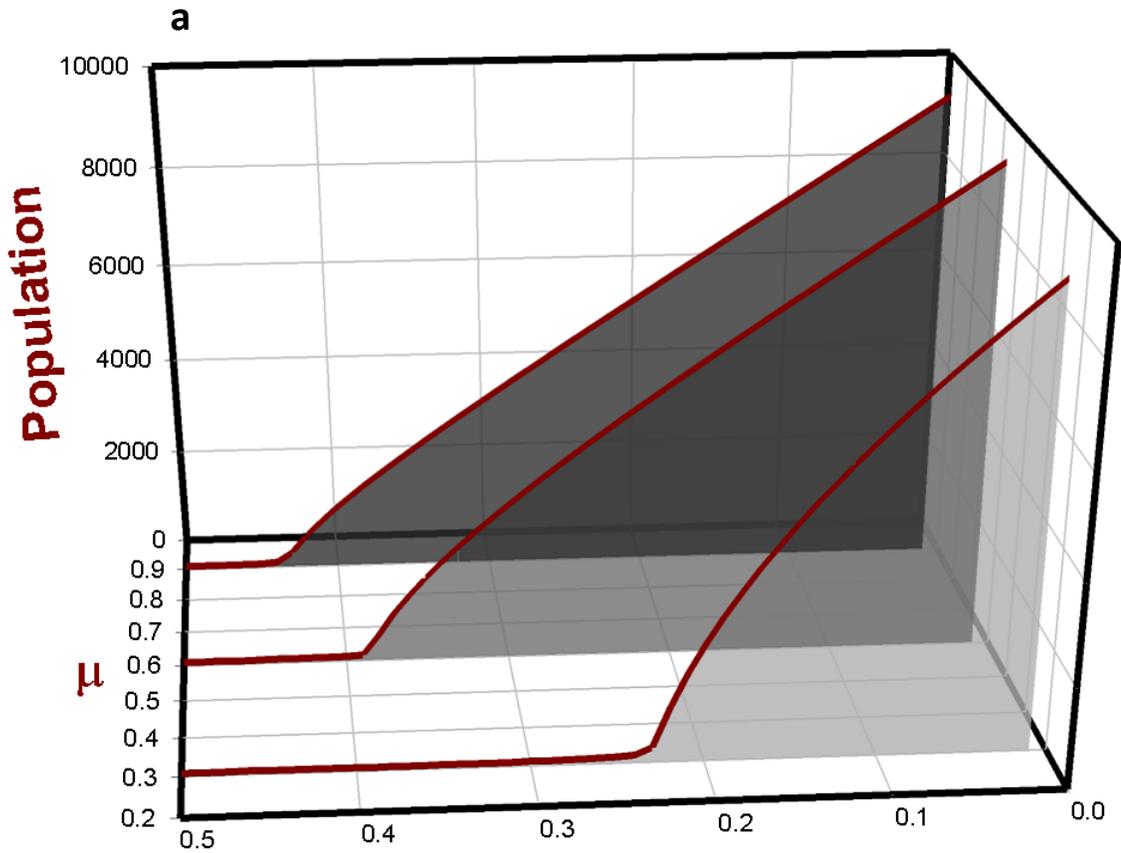
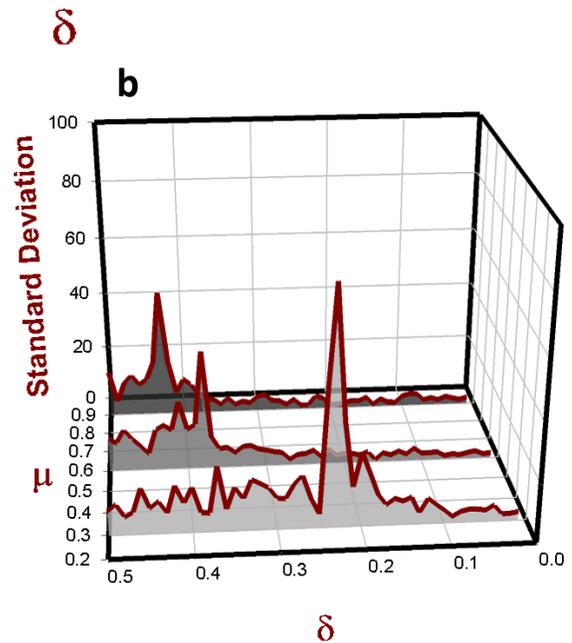


Figure 5 (a) Mean population for assortative mating as a function of the random death probability δ and mutability μ . Mean values are calculated over all surviving generations for each simulation, and then averaged over five different simulations at each value of δ and μ ; (b) Standard deviation over the five simulations. Each simulation ran for 2000 generations, unless extinction occurred.



Fission

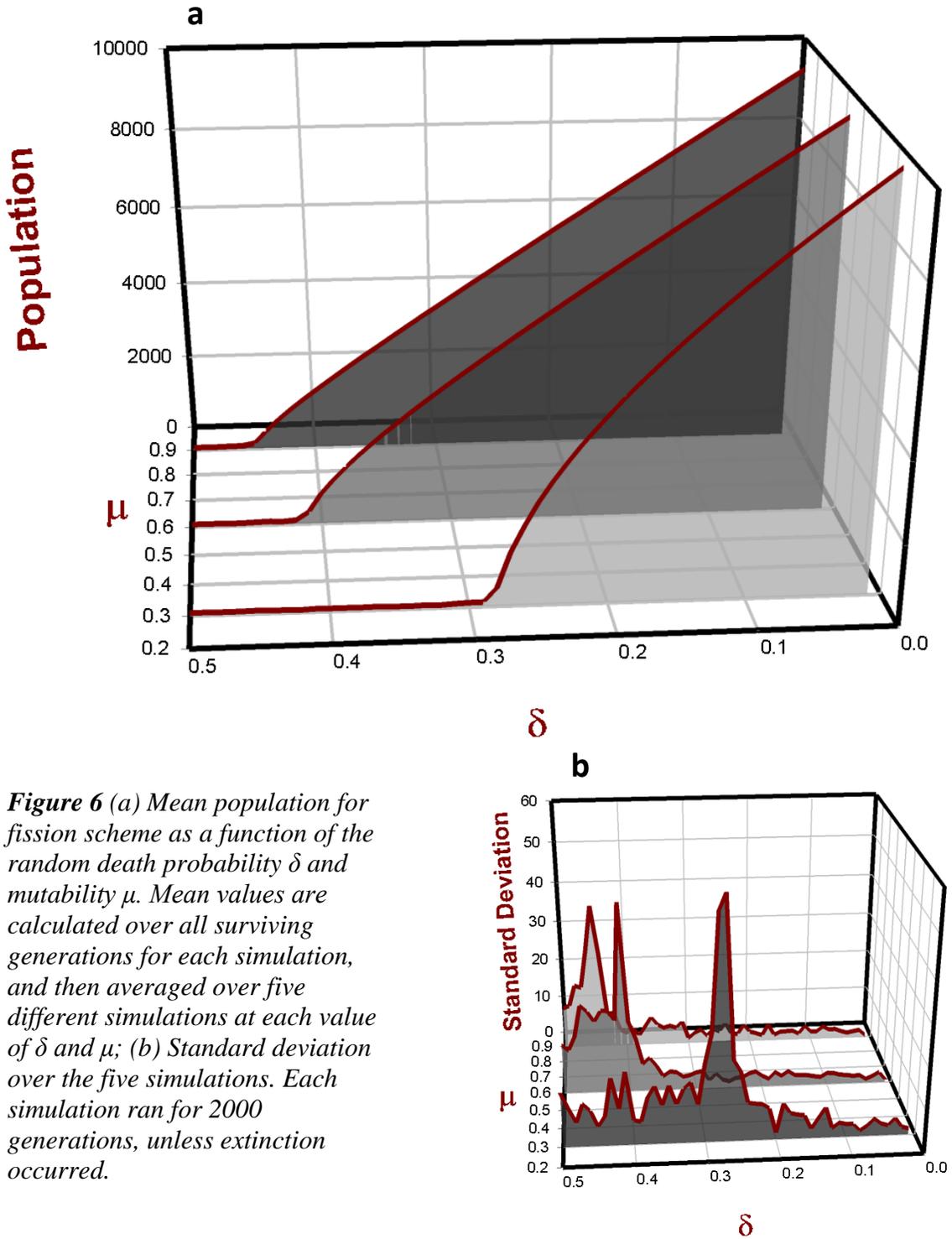


Figure 6 (a) Mean population for fission scheme as a function of the random death probability δ and mutability μ . Mean values are calculated over all surviving generations for each simulation, and then averaged over five different simulations at each value of δ and μ ; (b) Standard deviation over the five simulations. Each simulation ran for 2000 generations, unless extinction occurred.

Assortative Mating

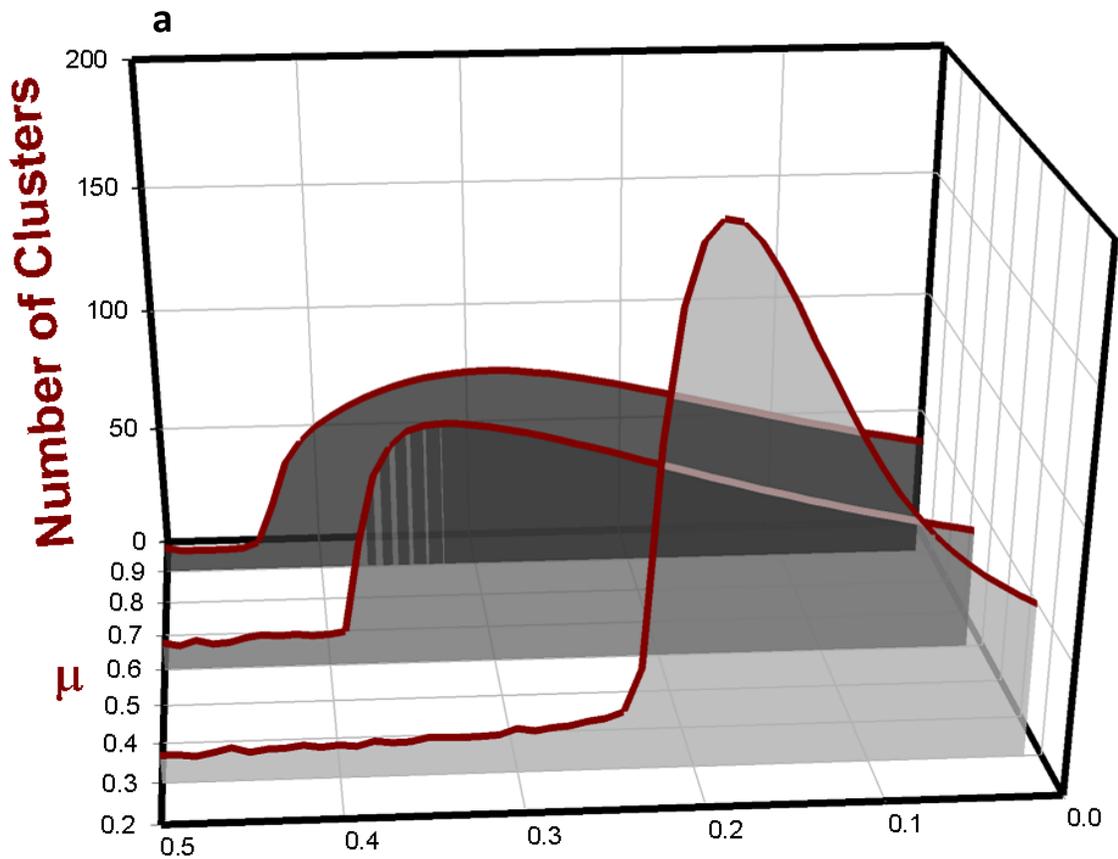
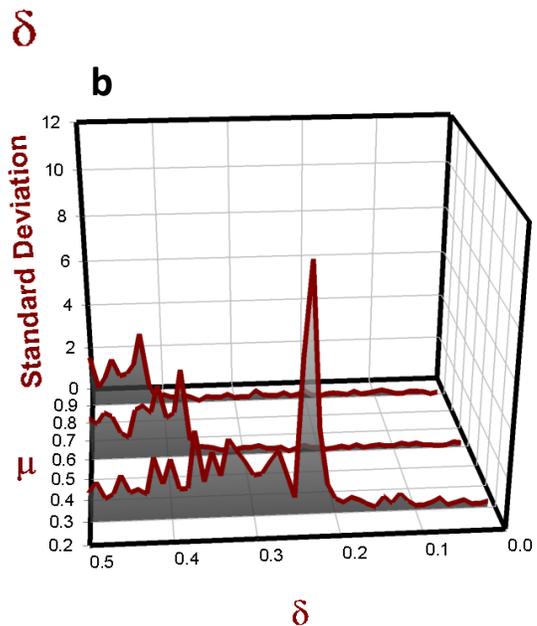


Figure 7 (a) Mean number of clusters for assortative mating as a function of the random death probability δ and mutability μ . Mean values are calculated over all surviving generations for each simulation, and then averaged over five different simulations at each value of δ and μ ; (b) Standard deviation over the five simulations. Each simulation ran for 2000 generations, unless extinction occurred.



Fission

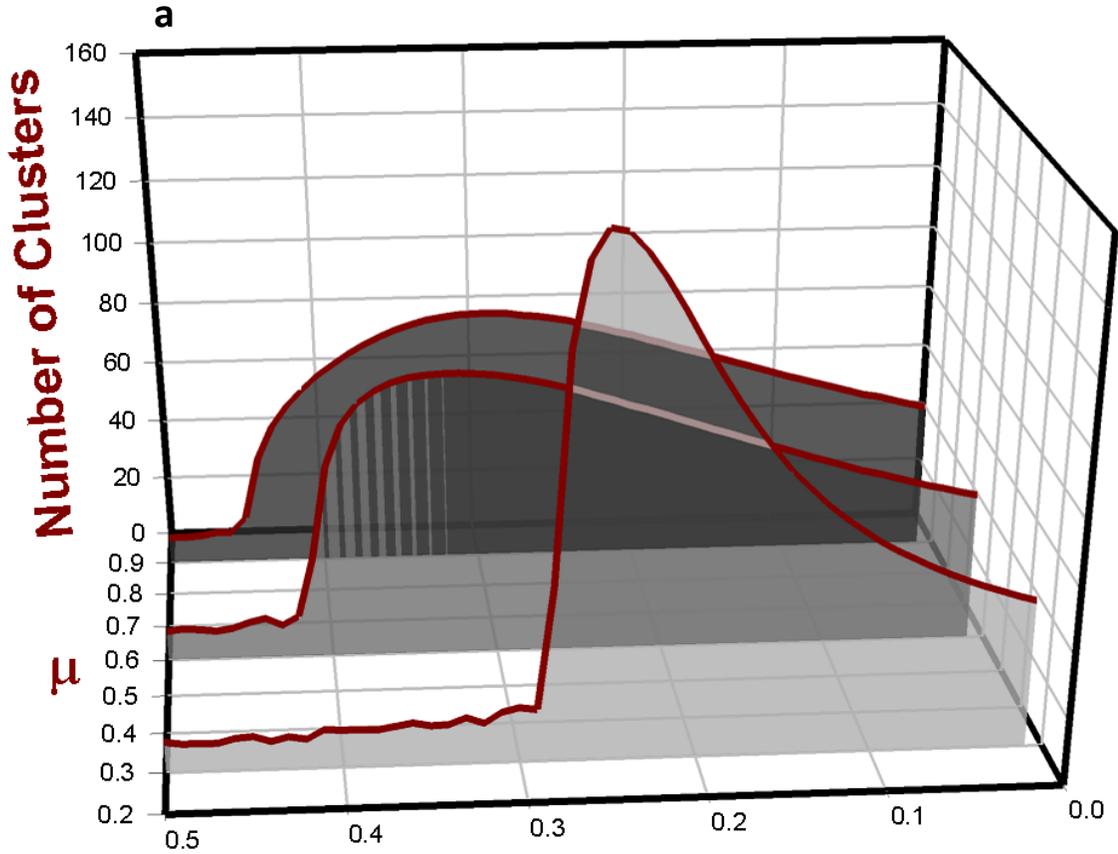
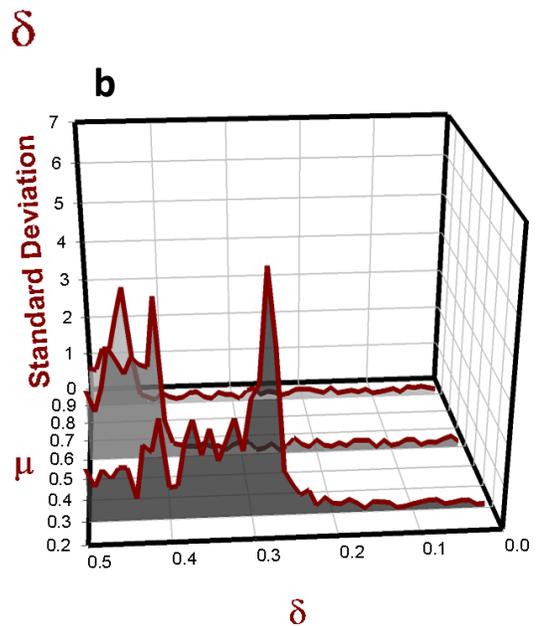


Figure 8 (a) Mean number of clusters for fission scheme as a function of the random death probability δ and mutability μ . Mean values are calculated over all surviving generations for each simulation, and then over five different simulations at each value of δ and μ ; (b) Standard deviation over the five simulations. Each simulation ran for 2000 generations, unless extinction occurred.



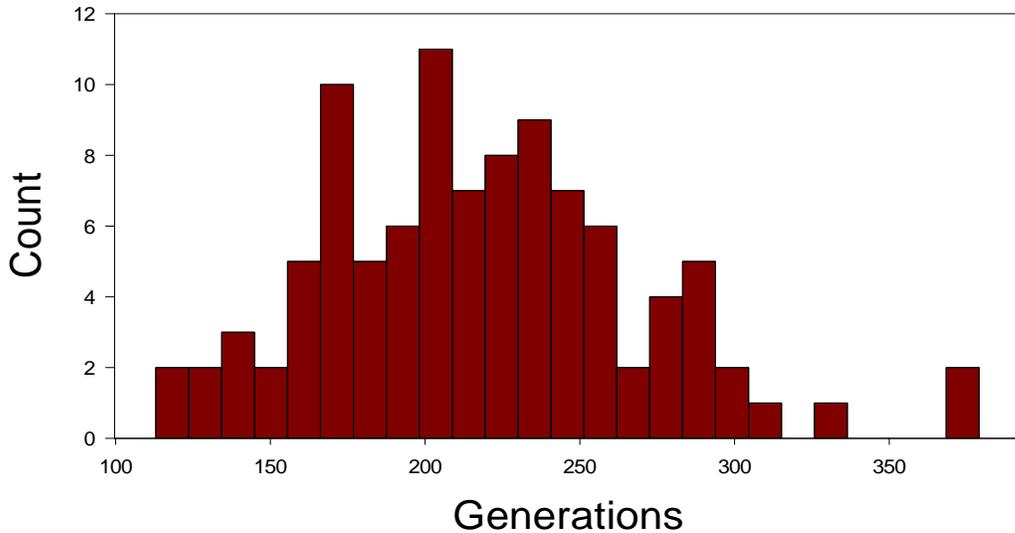
another transition, as μ increases beyond the values presented here, to a point where there is no phase transition at all. Similarly, Figures 7 and 8 show that δ_c increases with μ for the transition defined with the number of clusters serving as an order parameter; however, instead of the phase transition curve flattening out as μ increases, there now exists a sharp peak at $\mu=0.30$ that flattens as μ increases, suggesting again a qualitative change in behavior of the transition as μ is increased.

Figures 9 and 10 show the distributions of the number of generations a population survives for $\mu = 0.30$ and both mating schemes. The number of generations was set to one million, and one hundred simulations were run for each μ and δ presented. The trend from both figures demonstrate a more Gaussian-like distribution for values of δ in the absorbing state of extinction, and a more power law-like distribution in the neighborhood of δ_c . Note that had the fission simulation been allowed to continue beyond the one millionth generation, the tail of the distribution would have extended further. The six simulations that stopped at the millionth generation are not shown. After the approximated δ_c , all simulations ran to one million generations, thus indicating the system was in the active 'survival' state. Similar behavior occurred for $\mu = 0.60, 0.90$ (data not shown).

While Figures 9 and 10 show increasingly critical behavior of the system lifetime, Figures 11-16 suggest the emergence of power law scaling of cluster sizes. In these figures, the abundance of clusters of a given size (measured as individuals/cluster) are shown on a log-log scale. Here, for each value of μ and both mating schemes, there

Fission

$$\delta = 0.31$$



$$\delta = 0.28$$

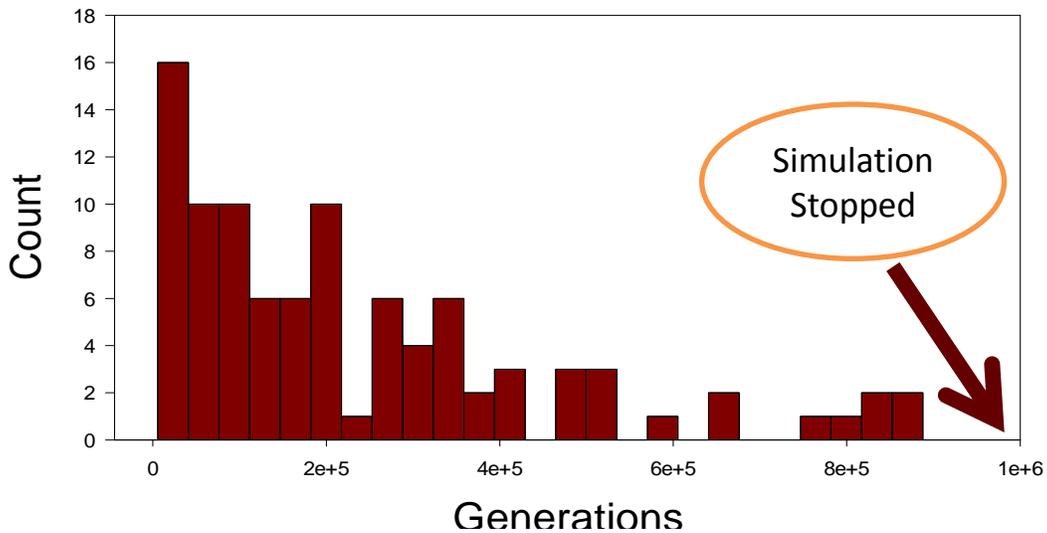
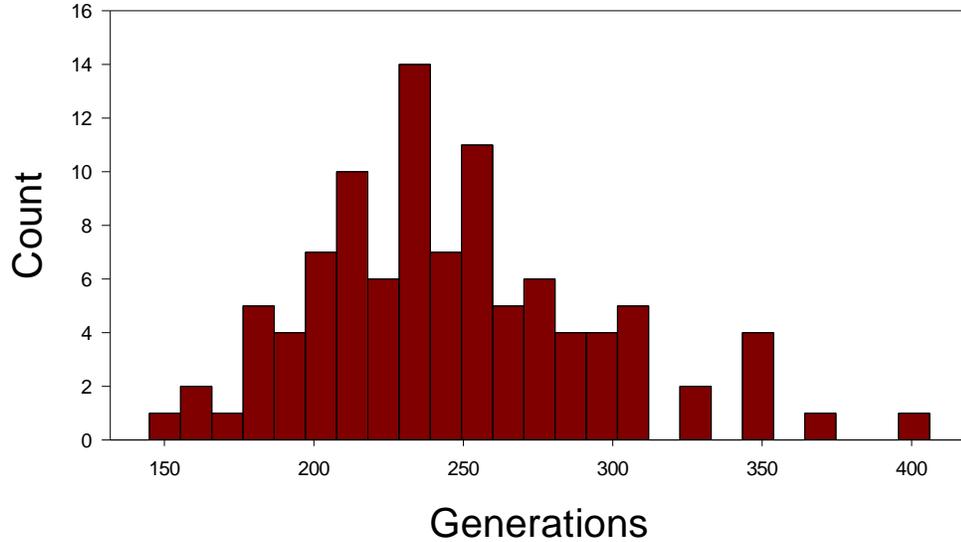


Figure 9 Lifetime distributions from 100 simulations show the number of generations a population survived for $\mu = 0.30$ in the fission model (horizontal axis) vs. the number of simulations which survived for that many generations (vertical axis). The top distribution shows a value δ within the extinction regime. The bottom shows a value just below the value of δ_c obtained from the maximal standard deviation (Figures 6b and 8b). Simulations were allowed to run for 1,000,000 unless extinction occurred first. Note the different horizontal axis scales.

Assortative Mating

$$\delta = 0.27$$



$$\delta = 0.23$$

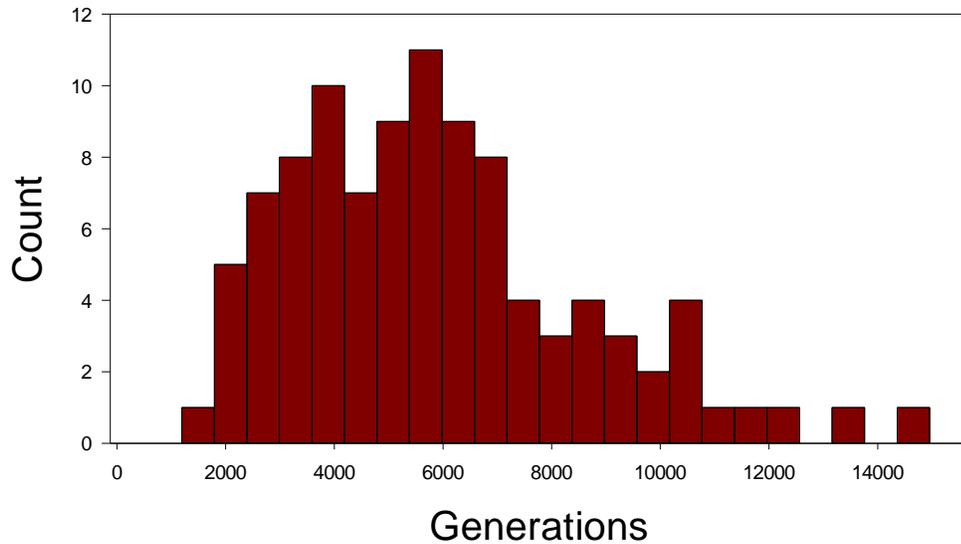


Figure 10 Lifetime distributions from 100 simulations show the number of generations a population survived for $\mu = 0.30$ of the assortative mating model (horizontal axis) vs. the number of simulations which survived for that many generations (vertical axis). The top distribution shows a value δ within the extinction regime. The bottom shows a value just below the value of δ_c obtained from the maximal standard deviation (Figures 5b and 7b). Simulations were allowed to run for 1,000,000 unless extinction occurred first. Note the different horizontal axis scales.

Assortative Mating
 $\mu = 0.30$

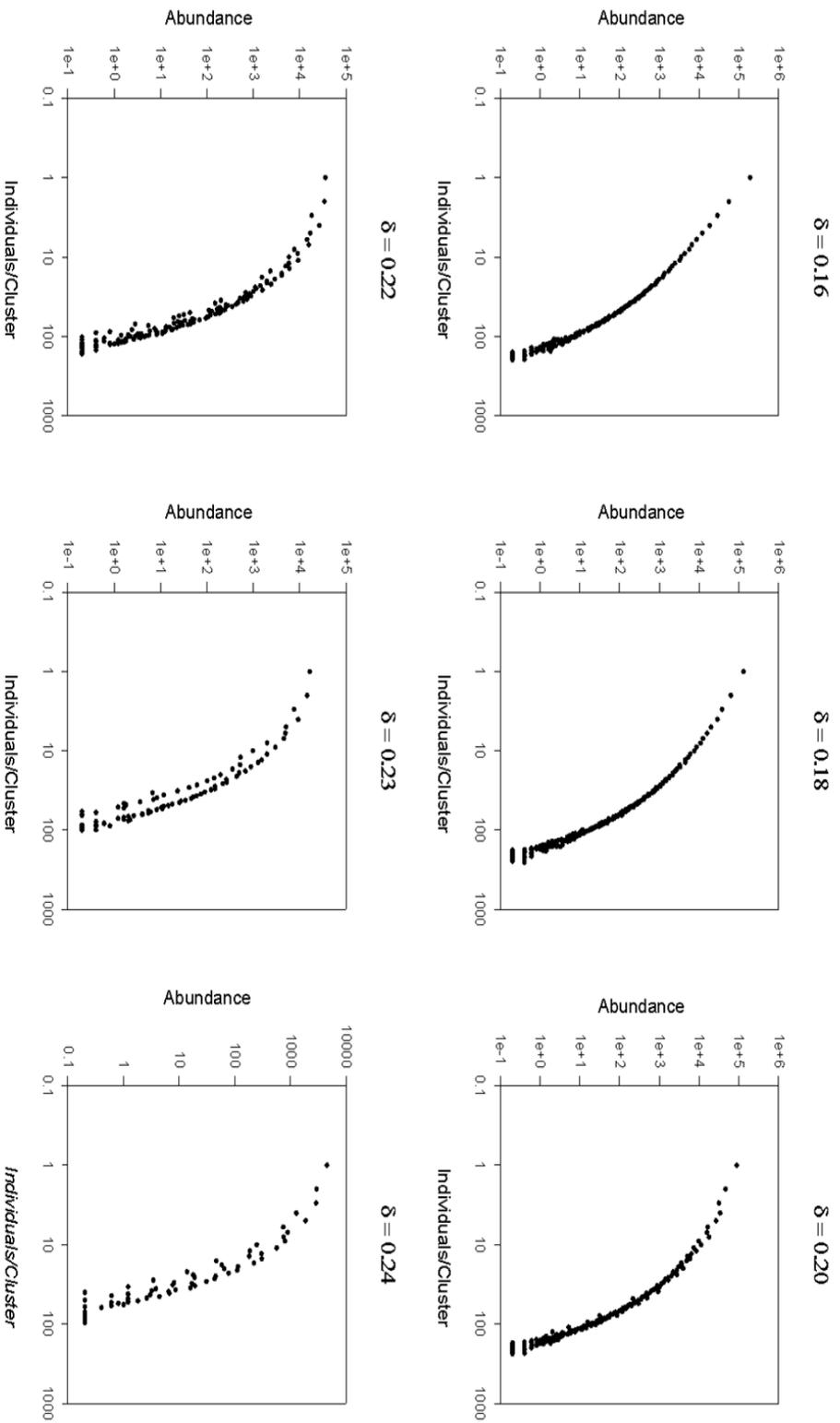


Figure 11 Distributions of number of individuals per cluster for the assortative mating model at $\mu = 0.30$. Values of δ read from left to right with δ_c in bottom row, middle column.

Assortative Mating
 $\mu = 0.60$

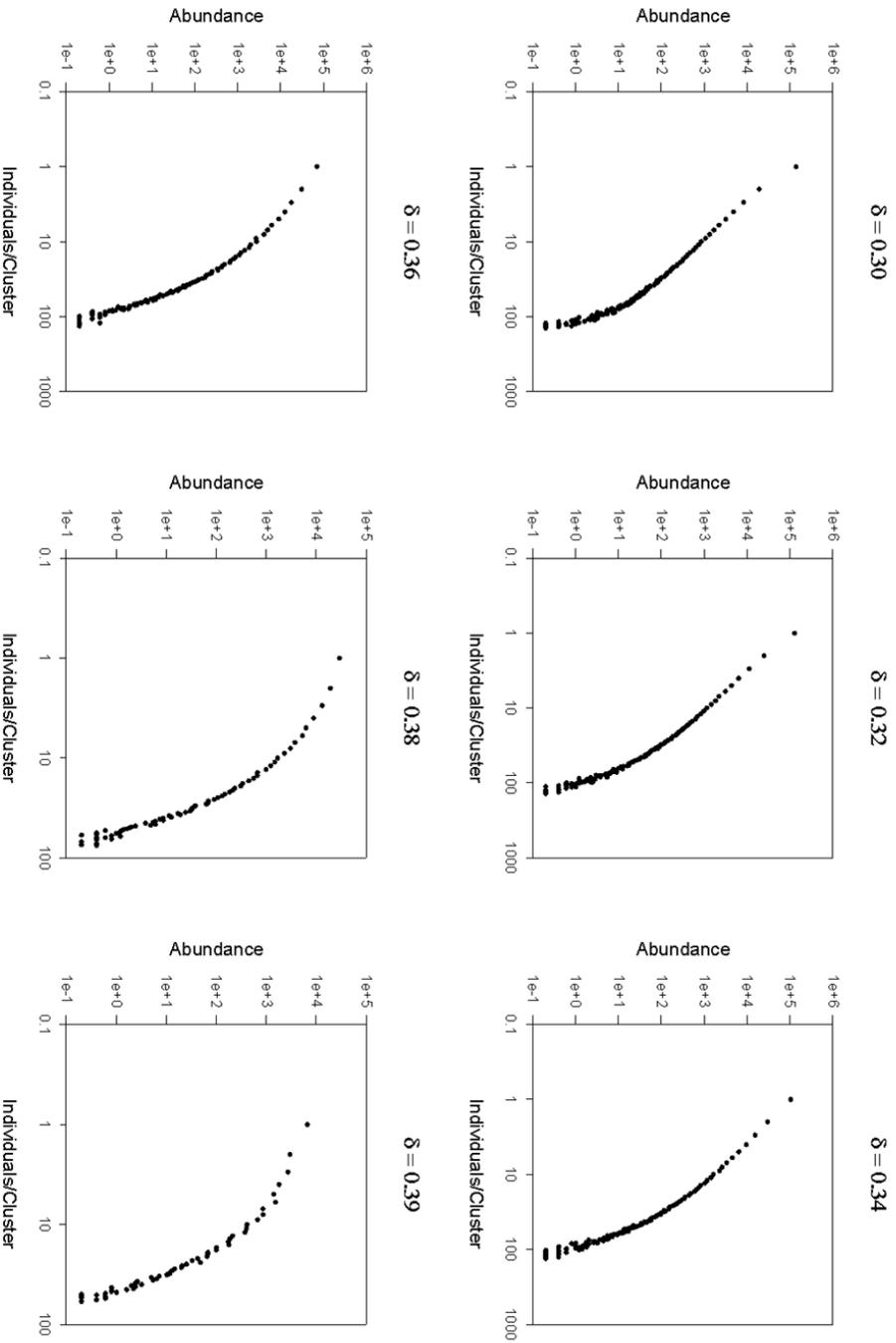


Figure 12 Distributions of number of individuals per cluster for the assortative mating model at $\mu = 0.60$. Values of δ read from left to right with δ_c in bottom row, middle column.

Assortative Mating $\mu=0.90$

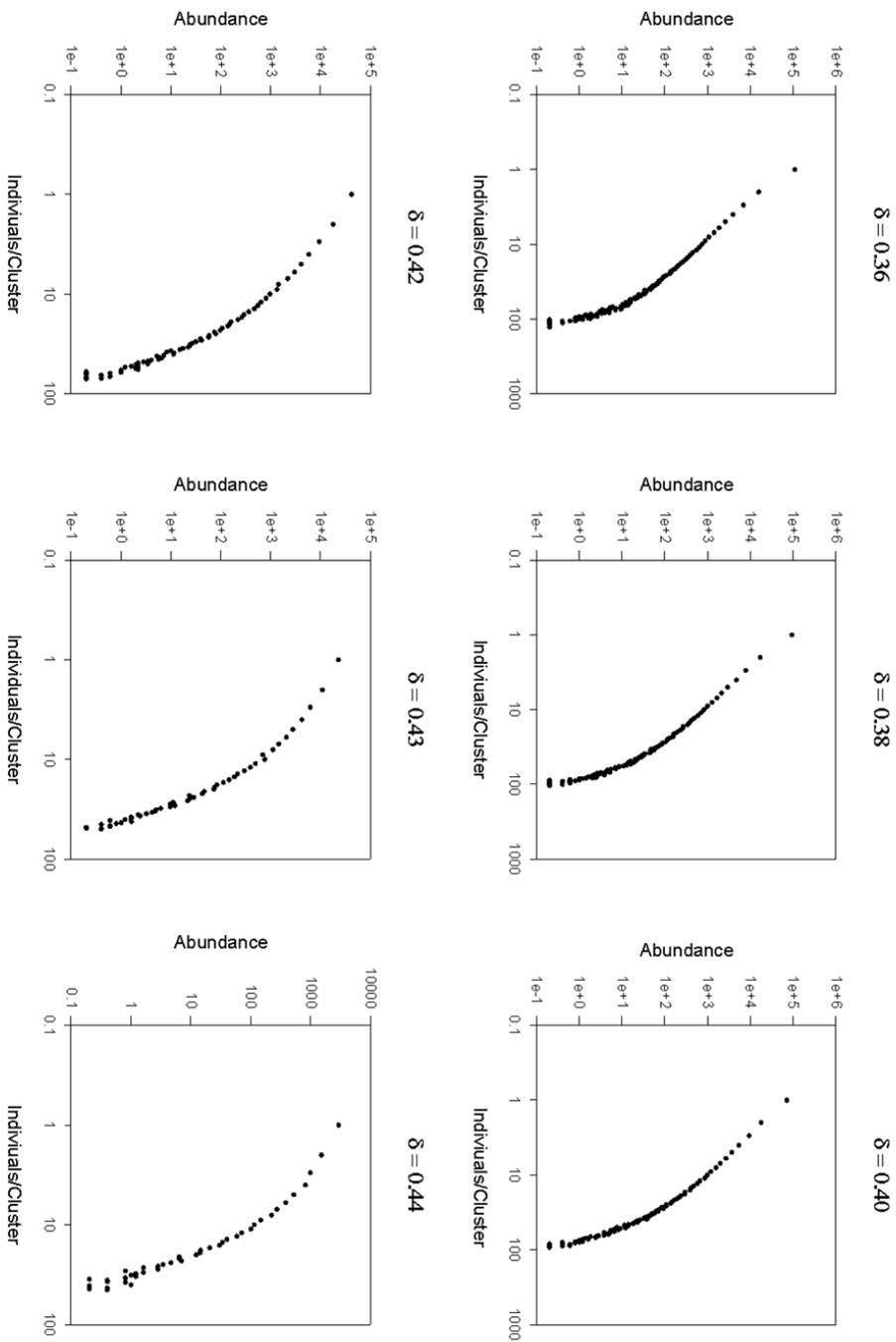


Figure 13 Distributions of number of individuals per cluster for the assortative mating model at $\mu = 0.90$. Values of δ read from left to right with δ_c in bottom row, middle column.

Fission
 $\mu = 0.30$

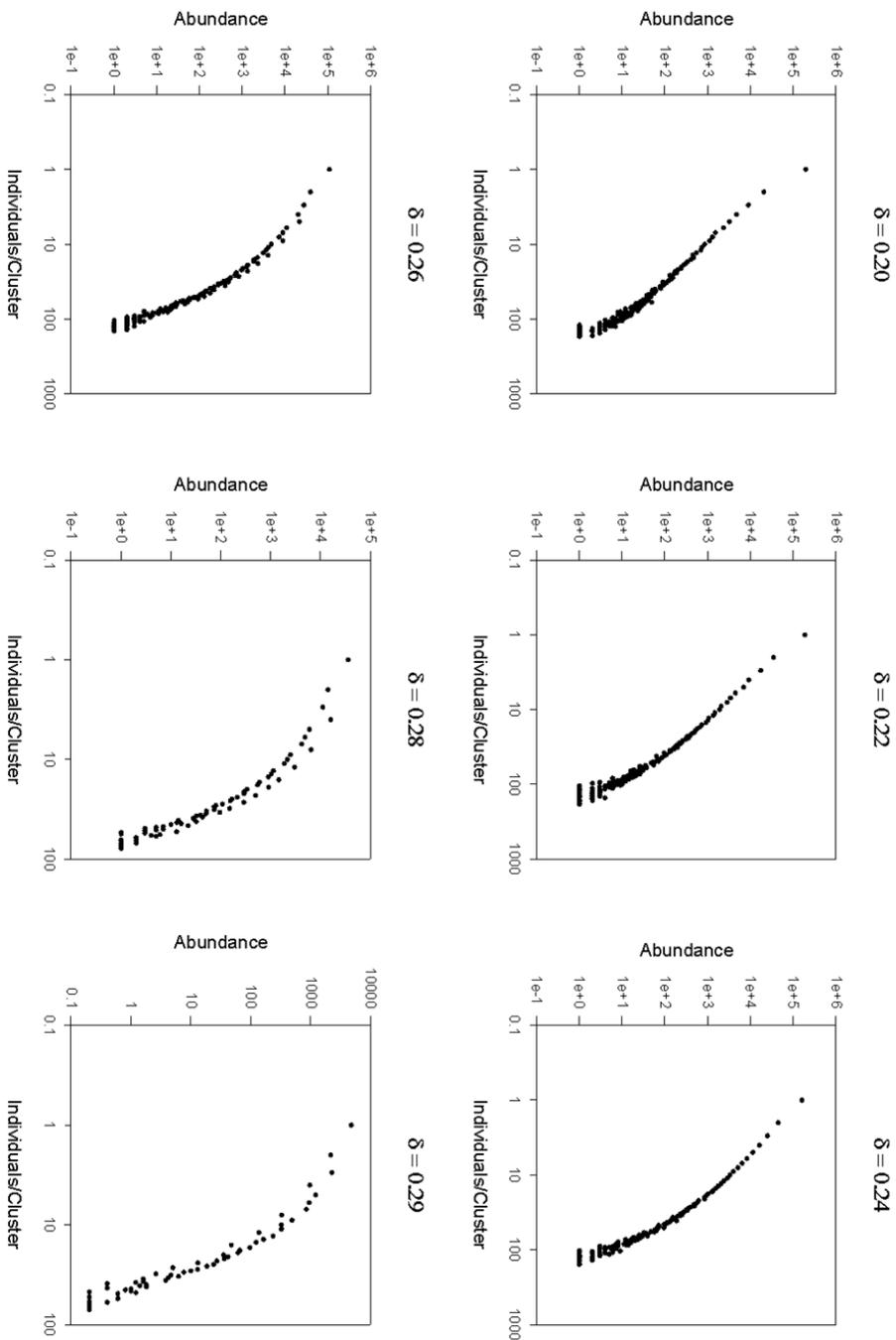


Figure 14 Distributions of number of individuals per cluster for the fission model at $\mu = 0.30$. Values of δ read from left to right with δ_c in bottom row, middle column.

Fission
 $\mu = 0.60$

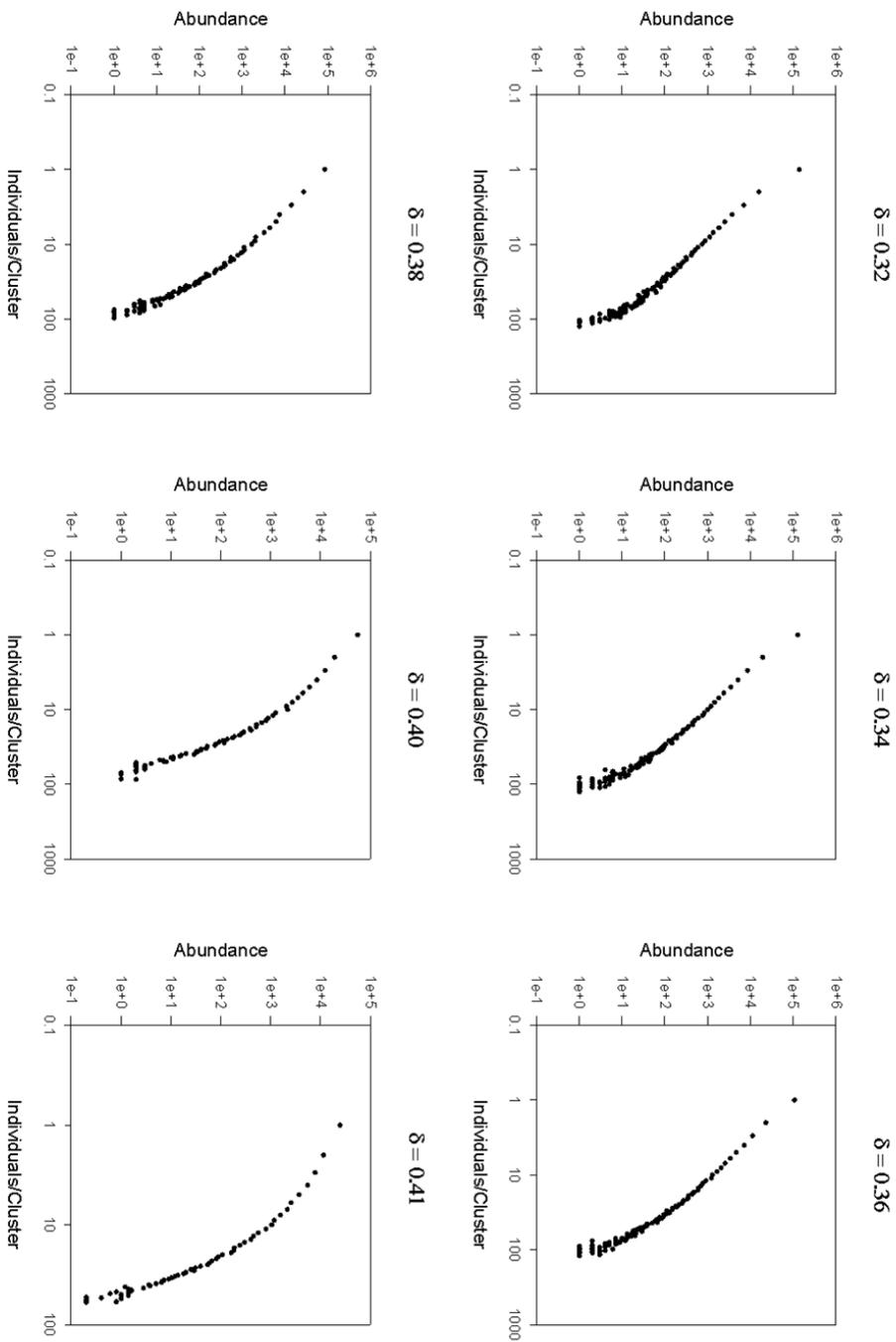


Figure 15 Distributions of number of individuals per cluster for the fission model at $\mu = 0.60$. Values of δ read from left to right with δ_c in bottom row, middle column.

Fission
 $\mu = 0.90$

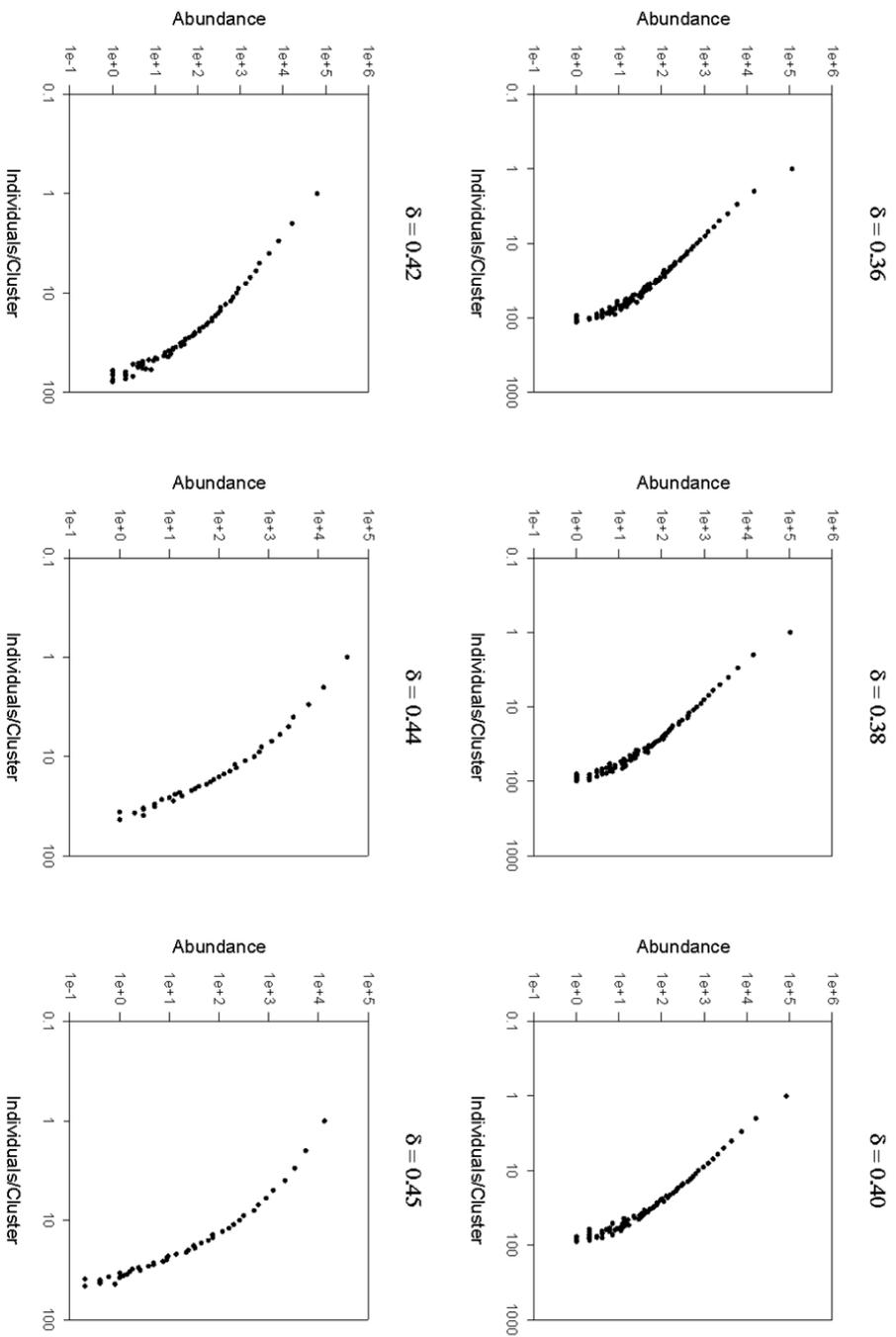


Figure 16 Distributions of number of individuals per cluster for the fission model at $\mu = 0.90$. Values of δ read from left to right with δ_c in bottom row, middle column.

appears to be a trend toward increased linearity on the log-log plots of the abundance curves as $\delta \rightarrow \delta_c$. In contrast to Figures 9 and 10, these results illustrate a trend toward power law behavior on the approach to δ_c from the regime of survival (δ approaching δ_c from above), instead of from the absorbing state (δ approaching δ_c from below).

Results from the Scott et al. investigation showed minimal survival of the population for the values of μ presented here for the random mating scheme. Furthermore, no phase transition existed with respect to the control parameter μ ; instead, the population size followed a smooth, Gaussian-like curve as μ was varied. In strong contrast to the assortative and fission models, clustering in the random model only consisted of 'one giant component' – which was to be expected due to the similarity the random mating scheme bears to random graph theory. For the random mating scheme presented here, minimal survival has also been observed for $\mu \leq 0.90$, suggesting that this model will show similar behavior to the results of Scott et al.. Investigation further into the random mating scheme with respect to δ will be the focus of future simulations.

Discussion

Nonequilibrium continuous phase transition behavior has been demonstrated for both order parameters of population size and numbers of clusters and for both the asexual fission and assortative mating models. A transition to an active ordered state of survival occurs for $\delta > \delta_c$, while for $\delta_c < \delta$ the absorbing state of extinction is one from which the system can never escape – thus the reason this system is classified as 'nonequilibrium'. The approximate values of δ_c for both mating schemes and all values

of μ have been identified by the sharp peak in the standard deviations (Figures 5b – 8b) of the measures serving as the order parameters (population size and number of clusters). These values of δ_c are estimates, since the standard deviation plots were obtained over five simulations only; a larger number of simulations, and also a finer spacing of values of δ , would yield a more accurate determination of these values. Nevertheless, the existence of the fluctuating ordered state at δ_c demonstrates that this is a continuous phase transition, for there is no discontinuous jump in the order parameters. The existence of power law behavior in the distributions of lifetimes (Figures 9 and 10) and possibly in the distribution of cluster sizes (Figures 11 through 16) in the neighborhood of δ_c is further evidence of the continuous nature of the transition. Unique to power law behavior and continuous phase transitions is the ability to classify a system into a particular universality class. The control condition of random mating still showed similar behavior to the results of Scott et al., thus indicating that the *type* of mating has an effect on collective behaviour. Further investigation will determine the universality class, and examine more critically the effect random mating has on the present model.

An increased robustness of the system is presented here by the fact that, as μ increased, δ_c also increased for both the assortative mating and fission schemes. This indicates that populations are able to survive in less hospitable environments (or harsher death conditions) if they are able to mutate further from their parents. The simulations presented here showed that populations could still survive with $\delta_c = 0.44$ at $\mu = 0.90$ for the fission model and $\delta_c = 0.43$ for assortative mating. Experimental

(Sniegowski et al., 1997) and computational (Taddei et al., 1997) studies have demonstrated that *Escherichia coli* can increase its mutation rate in order to maintain survival in inhospitable conditions. Sniegowski et al. (1997) demonstrated a 'rise in mutators in populations of *E. coli* undergoing long-term adaption to a new environment,' and ascertained that 'mutator alleles must have arisen by mutation;' while, Taddei et al. (1997) demonstrated an increased mutation rate depending on the number of mutator alleles present. Generally, an allele 'is an alternate version of a gene that produces distinguishable phenotypical effects' (Campbell, 2005). Similarly, since the model presented here is representative of phenospace where independent x,y coordinates represent organisms' phenotypes, rather than explicit genotypes, these *simulated* organisms also demonstrate an increased ability to survive based on decreasing phenotypic similarity between offspring and parent.

There is also evidence suggesting that aggregation behavior is determined by μ . Previous investigation by Scott et al. (submitted) showed a phase transition curve for the number of clusters as a function of μ , which is similar to that observed here as a function of δ (Figures 7 and 8). In both cases, the number of clusters exhibited a sharp peak for values of the control parameter beyond the critical range (note that this corresponds to $\mu > \mu_c$ for the transition as μ is varied, and for $\delta < \delta_c$ for the transition shown in Figures 7 and 8). Using the Clark and Evans (1954) nearest neighbor index, R , Scott et al. showed that, for values of μ below this sharp peak in number of clusters, the organisms form aggregated, clumped clusters, and for values of μ above this peak the organisms form 'more uniformly spaced clusters'. Preliminary investigation (data not

shown) has shown similar aggregation behavior at the value of $\mu = 0.30$ when δ is varied for both assortative mating and fission schemes. Furthermore, Figures 7 and 8 illustrate the erosion of the sharp peak for $\mu = 0.60, 0.90$, and thus, for these values of μ , the qualitative change in clustering does not appear to occur. This suggests that the characteristics of 'aggregation' in the model may be heavily dependent on μ . These qualitative changes in clustering might be characterized better through percolation theory, which deals primarily with the permeation of clusters through space. Below, I will sketch out possible directions for future studies investigating the percolation properties of the system, and then discuss how percolation will help to determine the universality class of the system.

Percolation theory (or ordinary percolation) is the description of how individual components group together in space in a given generation and is not concerned with the change from generation to generation. Of particular interest is the formation of a cluster that spans from end to end of the space – when this happens it is said that percolation is achieved. This point at which percolation is first achieved is called the percolation threshold p_c , which is the probability (or fraction of space occupied by organisms) for which the emergence of an 'infinite' cluster – one that spreads from end to end of the landscape, but theorized to reach infinity if the landscape was infinite – occurs. For example, if the the landscape has N individuals, then $p = 1$ corresponds to space being completely filled, $p = 0$ to no individuals in the space at all, and $p_c N$ indicates the fraction of individuals needed for percolation to be achieved. This is important because this threshold defines another nonequilibrium continuous phase

transition, in which the system moves from a state where, before the threshold ($p < p_c$), only clumped, aggregated clusters form, to one where, after the threshold ($p > p_c$), only uniformly distributed clusters are formed. Therefore, p_c defines the critical point of a phase transition between the probability of connected components where before p_c the system will never fully connect (or reach across the landscape), and after p_c the system will always reach across the landscape (often times forming a 'giant' component). For example, if the density of coffee grains is too high, then water will never percolate through the space – thus p_c defines the fraction of grains necessary for water to percolate across the space.

At p_c , the system is said to have *scale free* behavior in the number of steps (analogous to the number of individuals per cluster) it takes for a cluster to form and in the path length (the measured distance between each organism of a cluster starting with the first cluster 'seed' organism and ending with the last) of cluster formation (Stauffer & Aharony, 1994). Note that, at least for the case of this model, while a cluster is forming, the shortest route from end to end of the landscape is not taken, since an organism chooses its mate based on proximity. For example, consider the assortative mating scheme and the algorithm of how organisms choose the most 'similar' mates (i.e., the shortest Euclidian distance) in order to form a cluster. The first individual that starts the algorithm is not directed in any particular direction, for it 'searches' within a 360° radius of itself and then chooses the closest individual as its mate, then that mate searches for the next closest to itself within 360° , and the next mate, and next mate, and so on... until a closed set is formed. If the above implementation of the mating

algorithm is thought of in terms of bonds that form in space and time with each 'mate step' taken considered as a time step, it would appear that 'mating' exhibits characteristics of Brownian motion since these individual step lengths of mating are random in direction and restricted to be 'near' each other. Thus, this type of mating behavior can lead to highly connected (or lengthy) clusters, which is why the within-cluster path lengths form a power law distribution at criticality. The same reasoning applies to the number of individuals (or steps) in cluster formation.

Keeping in mind the previous rationale of the clustering algorithm, the percolation behavior of clusters above and below p_c , and the Clark and Evans nearest neighbor index (which indicates that the clustering behavior shifts from aggregated to uniform as the plot of the number of clusters reaches its peak), I hypothesize that the peak of the clusters curve at $\mu = 0.30$ should occur at the value of δ for which percolation is achieved (call it δ_p). δ_p gives the probability or the percentage of organisms removed at p_c , thus since p_c indicates the fraction of individuals on the space, then $p_c = 1 - \delta_p$ is the fraction of opened space. To test the hypothesis that percolation occurs at δ_p , future simulations, particularly at δ_p , will reveal whether critical behavior of the formation of clusters (Note that Figures 11-16 provide preliminary evidence of linear log-log behavior of cluster sizes for values of $\delta < \delta_c$) at the value of δ_p exists— thus indicating whether $p_c = 1 - \delta_p$. Since percolation depends on the spread of a cluster in a given generation, the initial population size would need to be set as indicated by the number given by the population curve in Figures 5 and 6 at the hypothesized value of δ_p (starting the population at that value should eliminate transient generations when the

population size is too small to reach across the landscape). Examination of how the clusters fill the space by tracking the clustering algorithm will be performed as follows: since each assignment of a mate is considered a 'time' step, then the number of time steps can be recorded for each cluster. Also, since each position of each organism in the phenospace is recorded, the path length of the clusters can be measured as well. If scale free behavior is found in the length of time (number of mating steps) required for cluster formation or in the path length, then δ_p corresponds to the percolation threshold $p_c = 1 - \delta_p$. Determining how clusters fill the space and the value of δ that gives the value of p_c will help to determine the universality class of system because, if δ_p is not identical to δ_c , then the percolation of clusters through the landscape does not correlate with the phase transition in the order parameter of the number of clusters on the landscape.

Clarification the system's universality class will begin by determining what the value of p_c is. According to Ódor (2002):

If the critical point of the order parameter does not coincide with the percolation threshold, then at the percolation transition the order parameter coherence length is finite and does not influence percolation properties. We observe random percolation in that case. In contrast if the critical order parameter and percolation threshold occur at the same critical point percolation is influenced by the order parameter behavior and we find different, **correlated percolation universality** whose exponents coincide with that of the order parameter.

Therefore, if $\delta_c \neq \delta_p$, then the percolation transition occurs at random (and no information is provided about the universality class of the system), but if $\delta_c = \delta_p$, then it can be concluded that the system exhibits correlated percolation universality, and the critical exponents can be obtained from percolation theory. It is possible that we might observe, since percolation is related to clustering, and the peak of the number of

clusters depends heavily on μ (i.e., the peak starts to vanish for larger μ), that $\mu_c = \mu_p$ and $\delta_c \neq \delta_p$, so that the phase transition as one control parameter is varied belongs in a different universality class from the transition as the other control parameter is varied. Preliminary results from the Scott et al. paper (data not shown), show a possible percolation transition occurring at μ_c . It could be that since μ imposes a local restriction on percolation (i.e., the next generation of offspring are confined to a certain space allotted by μ), and δ effects individuals randomly and globally (death can happen anywhere on landscape), a different universality class should be expected for each control parameter of the system. Thus, if the transition in relation to the control μ belongs to correlated percolation universality, and the transition in relation to δ does not, then what universality class does the transition in relation to δ belong to?

A very general universality class that describes many nonequilibrium systems is the directed percolation universality class. Characteristic of directed percolation (as implied by its name) is the directing of agent-based processes such that direction can either occur in space or time – or in both. Because it is a percolating process, directed percolation describes nonequilibrium processes, since it is characteristic of a percolating system to reach an absorbing state; thus, directed percolation is a simple way to describe critical phenomena and many mean field models have been developed from it (Hinrichsen, 2000). The directed percolation conjecture was constructed by Grassberger and Janssen, and presented in Henkel et al. (2008):

According to this conjecture, it is thought that a given model should generically belong to the DP universality class if 1. The model displays a continuous phase transition from a *fluctuating* active phase into a unique absorbing state, 2. The transition is

characterised by a non-negative one-component order parameter, 3. The dynamic rules are short ranged, 4. The system has no special attributes such as unconventional symmetries, conservation laws, or quenched randomness.

Regarding point 1, I have demonstrated continuous phase transitions from a state of fluctuating survival to an absorbing state of extinction as the parameter δ is varied. Secondly, each phase transition exists for a positive one-component order parameter (number of clusters or population size). Third, the dynamic rules of the mating systems are short ranged, i.e., μ restricts how far offspring are generated from parents and assortative mating restricts individuals to mate with the most phenotypically similar individuals. Lastly, the system has no symmetries, conserved quantities, or any quenching of any kind. The system is, in fact, asymmetric with respect to the birth and death processes, for births occur locally, near the parent(s), while deaths may occur anywhere on the landscape (globally). This last point could also be a fundamental reason the organisms cluster for an individual-based model, for previous work has shown clustering of asexual organisms and credited the clustering to the asymmetric birth and death process as well (Meyer et al., 1996, Young et al. 2001). Thus, based on the directed percolation conjecture, and how well the current evolution model fits each point, it is probable this model will fall into the directed percolation universality class. Note the phase transition as μ is varied (Scott et al., submitted) also satisfies this conjecture, but that there is also evidence suggesting this phase transition occurs at p_c , which suggests that it could belong to the correlated percolation universality class instead. More work will need to be done to parse out this seemingly paradoxical relationship between the space and time behavior of μ .

If the above hypothesis is correct, such that ordinary percolation determines $p_c = 1 - \delta_p$ (and $\delta_c \neq \delta_p$) then the critical behavior of *how* organisms cluster on the space will not be correlated with δ_c . It is likely δ_c is the critical point for a transition in the DP universality class. If the model is found to belong to the directed percolation universality class, this would mean that (at least) one universality class could describe the system in relation to time (the transitions driven by the parameters δ and μ); while another describes its relation to space (the percolation transition). These results will open fertile ground for speculation on the biological implications of the model. If it is shown that how the clusters fill the space is not correlated with the number of clusters, this may have a suggestive implication for the different structures of various types of biological diversity (between species vs. within species). The demonstrated increase in the robustness of the system as μ increases could have relevance for the broad biological question of whether 'evolvability' itself can be selected for; simulations involving competition between organisms with different values of μ and δ may be helpful in this regard. In the broadest sense, the phase transition approach to modeling speciation may ultimately contribute to the "hard problem" of multiple levels of evolution / group selection. Further studies, including more a biologically realistic version of the present model – such as the inclusion of an explicit genetics – will undoubtedly be necessary in order to achieve that goal.

Works Cited

Buchanan, M. (2000). *Ubiquity: Why Catastrophies Happen*. New York: Three Rivers Press.

- Barabási, A. (2003). *Linked*. New York: Penguin Group.
- Buss, L. W. (1987). *The Evolution Of Individuality*. Princeton, New Jersey: Princeton University Press.
- Campbell, N., & Reese, J. (2005). *Biology* (7 ed.). New York: Benjamin Cummings.
- Cipra, B. A. (1987). An introduction to the Ising model. *The American Mathematical Monthly*, 94 (10), 937-959.
- Clark, P., & Evans, F. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology*, 35 (4), 445-453.
- Dawkins, R. (2006). *The Selfish Gene* (30th Anniversary ed.). New York: Oxford University Press.
- Dees, N., & Bahar, S. (2010). Mutation size optimizes speciation in an evolutionary model. *PLoS ONE*, 5 (8).
- Henkel, M., Hinrichsen, H., & Lubeck, S. (2008). *Non-Equilibrium Phase Transitions: Volume 1: Absorbing Phase Transition*. The Netherlands: Springer Science and Business Media B.V.
- Hinrichsen, H. (2000). Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances In Physics*, 49 (7), 815-958.
- Hubbell, S. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton, New Jersey: Princeton University Press.
- Johnson, N. (2007). *Simply Complexity: a Clear Guide to Complexity Theory*. Oxford, England: Oneworld.
- Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature*, 217, 624-626.
- Kimura, M. (1983). *Neutral Theory of Molecular Evolution*. Cambridge UK: Cambridge University Press.
- Meyer, M., Havlin, S., & Bunde, A. (1996). Clustering of independently diffusing individuals by birth and death processes. *Phys. Rev, E*, 54 (5), 5567-5570.
- Mikhailov, A., & Calenbuhr, V. (2002). *From Cells to Societies: Models of Complex Coherent Action*. Berlin: Springer.
- Ódor, G. (2008). *Universality in Nonequilibrium Lattice Systems: Theoretical Foundations*. New Jersey: World Scientific.
- Sniegowski, P., Gerrish, P., & Lenski, R. (1997). Evolution of high mutation rates in experimental populations of E. Coli. *Nature*, 387, 703-706.
- Solé, R. (2011). *Phase Transitions*. Princeton, New Jersey: Princeton University Press.

Stauffer, D., & Aharony, A. (1994). *Introduction to Percolation Theory*. (2, Ed.) New York: CRC Press.

Taddei, F., Radnam, M., Maynard-Smith, J., Toupance, B. G., & Godelle, B. (1997). Role of mutator alleles in adaptive evolution. *Nature*, *387*, 700-702.

Yeomans, J. (1992). *Statistical Mechanics of Phase Transitions*. New York: Oxford University Press.

Young, W., Roberts, A., & Stune, G. (2001). Reproductive pair correlations and the clusterings of organisms. *Nature*, *412*, 328-331.