

The effect of the dispersal distribution on isolation-by-distance in a continuous population

Tara N. Furstenau* and Reed A. Cartwright

Center for Evolutionary Medicine and Informatics, The Biodesign Institute,
and School of Life Sciences, Arizona State University, Tempe, AZ

85287-5301 USA

*tfursten@asu.edu



Abstract

Under models of isolation by distance, population structure is determined by the probability of identity-by-descent between pairs of genes according to the geographic distance between them. Wright's neighborhood size represents a standardized measure of dispersal which partially quantifies the relationship between identity-by-descent and isolation-by-distance in continuous populations. Here I model dispersal of haploid individuals on a two-dimensional torus using various dispersal distributions. The distributions are parametrized to produce the same neighborhood size. The populations are evaluated to determine if similar patterns of isolation-by-distance are observed when neighborhood size is held constant.

Introduction

- Wright's (1946) neighborhood size ($N_b = 4\pi\sigma^2D$) is commonly used as a measure of isolation-by-distance and local genetic differentiation where the second moment of dispersal distance, $2\sigma^2$, is the only factor of the dispersal distribution that is taken into account.
- Rousset (1997) points out that dispersal distributions and isolation-by-distance have a more complicated relationship that involves the shape of the dispersal distribution. However, he points out that, for many classes of dispersal functions, σ^2 is likely the only parameter that matters.
- Here we simulate local dispersal in a continuous population to determine how patterns of isolation-by-distance and spatial structure vary based on the shape of four different dispersal distributions that all have the same second moment of dispersal distance.
- If patterns of isolation-by-distance are equivalent across dispersal models, then it will be possible to speed up individually-based, spatially-explicit simulations by using simpler models without the an reduction in quality.

Simulation

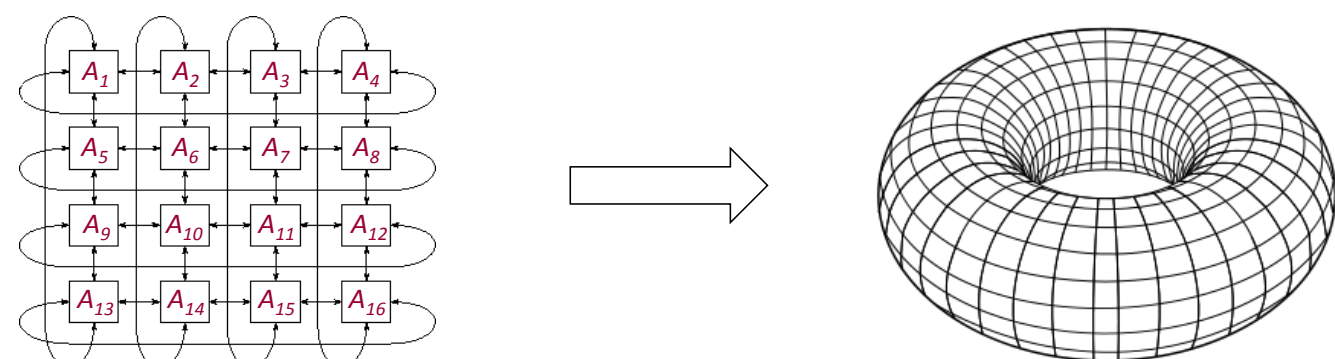


Figure 1: Haploid individuals occupy an $N \times N$ torus. A population exists on an $N \times N$ rectangular lattice with periodic boundaries (a torus). Each cell is occupied by one haploid individual with a single locus. The population is initialized so that each individual carries a unique allele. Image adapted from <http://www.mcs.anl.gov/itf/dbpp/text/node33.html> and http://commons.wikimedia.org/wiki/File:Simple_Torus.svg.

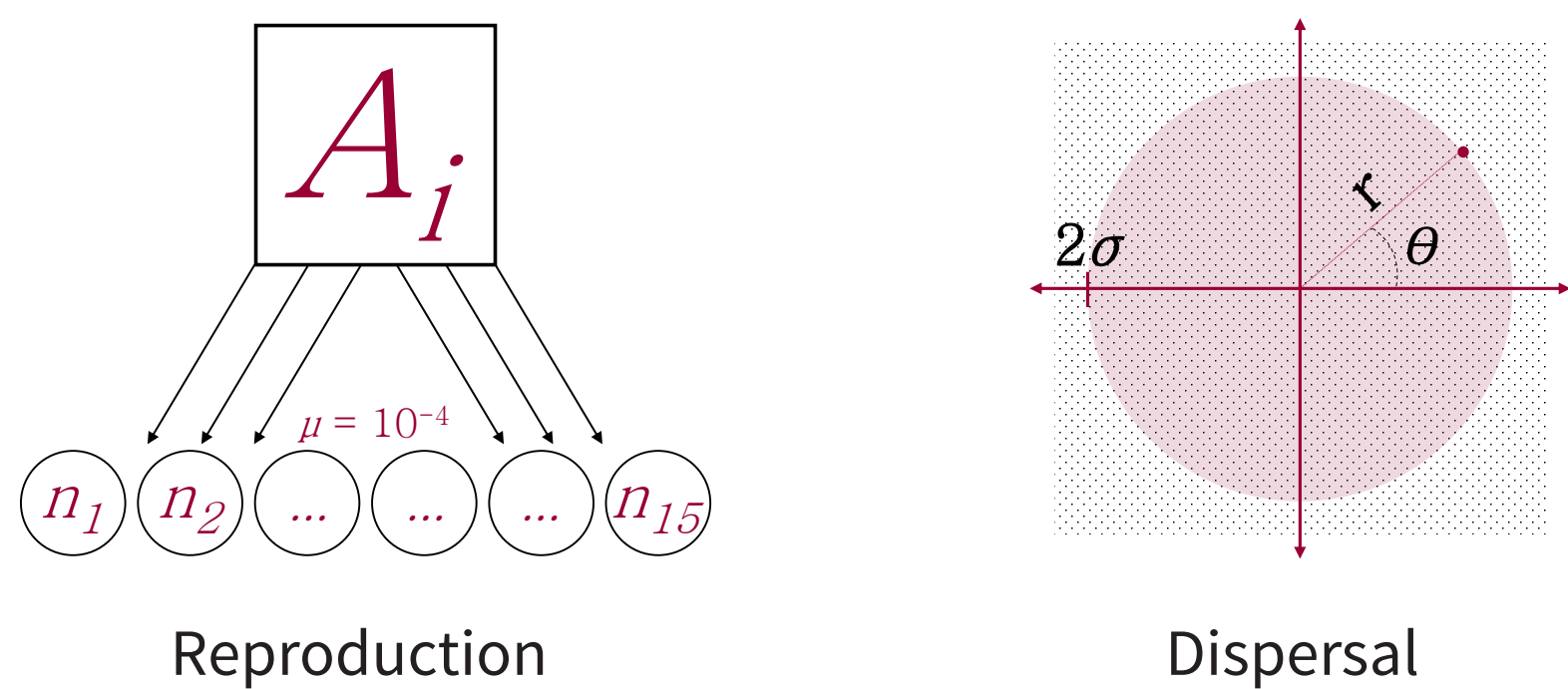


Figure 2: Generations are discrete. During each cycle, every individual reproduces by generating 15 clonal offspring which experience mutations according to the infinite alleles model at rate $\mu = 10^{-4}$. The offspring are then dispersed according to the dispersal distribution that is set for the simulation. After dispersal, only one uniformly selected offspring per cell will survive to become the parent of the next generation. The simulations were run for a total of 2 million generations after a burn-in period of 10,000 generations. Data was collected every 1,000 generations by sampling individuals from a transect (a single row across the landscape).

Modeling Dispersal

The dispersal function returns polar coordinates where θ is a uniform direction, and r is the distance determined by the given probability distribution. The σ parameter for each distribution is adjusted so that they all yield the same second moment of Euclidean distance.

Table 1: Dispersal Functions

Distribution	Dispersal Function	Parameter
Rayleigh	$f(r, \theta \sigma) = \frac{1}{2\pi} \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}$	$\forall x \geq 0 \quad \sigma = \sigma$
Exponential	$f(r, \theta \lambda) = \frac{1}{2\pi} \lambda e^{-\lambda x}$	$\forall x \geq 0 \quad \lambda = \frac{1}{\sigma}$
Half-normal	$f(r, \theta s) = \frac{1}{2\pi s \sqrt{\pi}} e^{-\frac{r^2}{2s^2}}$	$\forall x \geq 0 \quad s = \sigma \sqrt{2}$
Triangular	$f(r, \theta c) = \frac{1}{2\pi} \frac{2x}{c^2}$	$\forall 0 \leq x \leq c \quad c = 2\sigma$

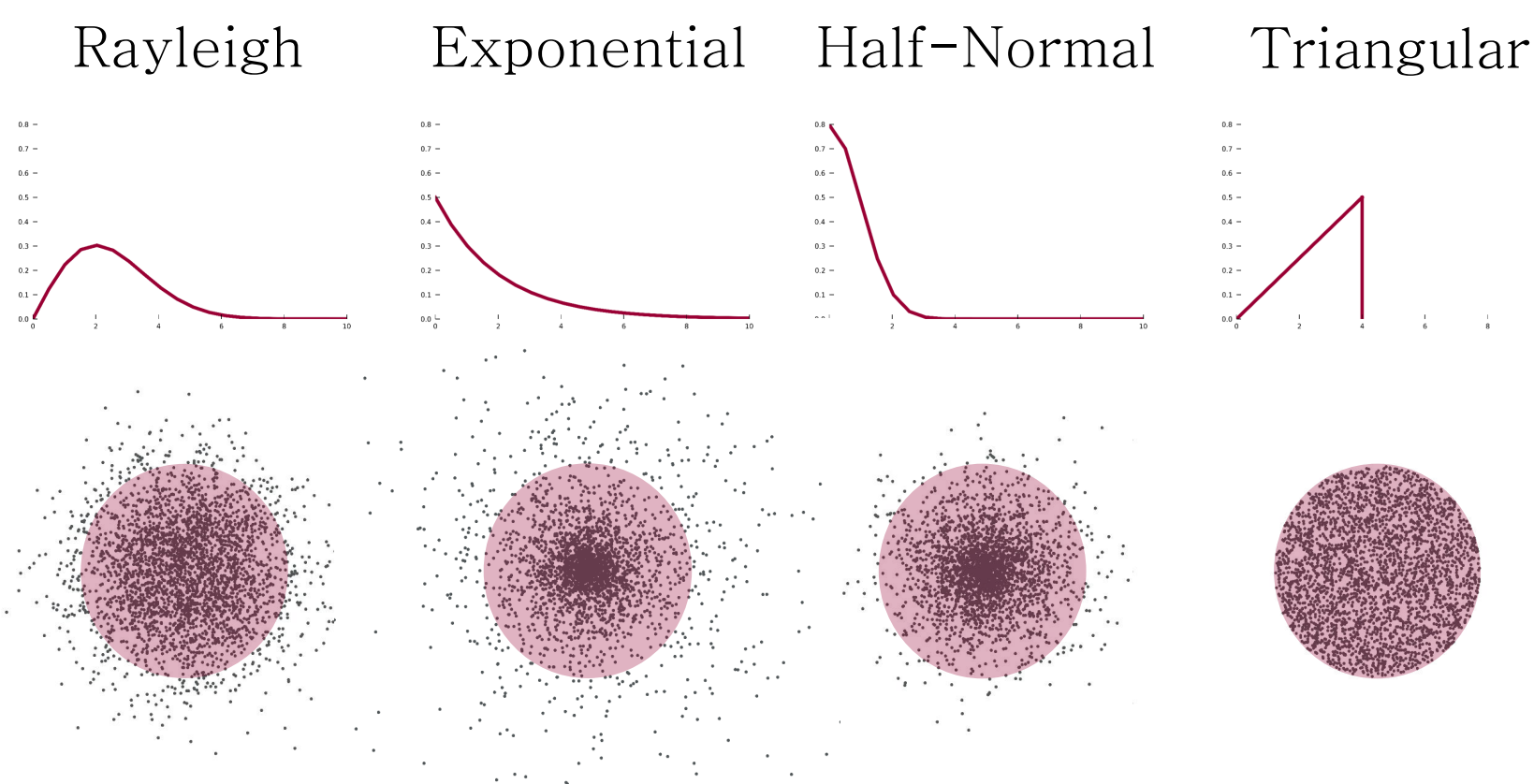


Figure 3: Example of the shape of the dispersal distributions and the result of dispersal in 2 dimensions. Here $\sigma = 2$ and the parameter for each function has been adjusted according to table 1. The red circle represents the Neighborhood area with a radius of $2\sigma = 4$.

Postdoc Positions Available in Computational Evolutionary Genomics

a conspiracy of
μinions
cartwright.ht/lab/

Email cartwright@asu.edu if interested.

Results

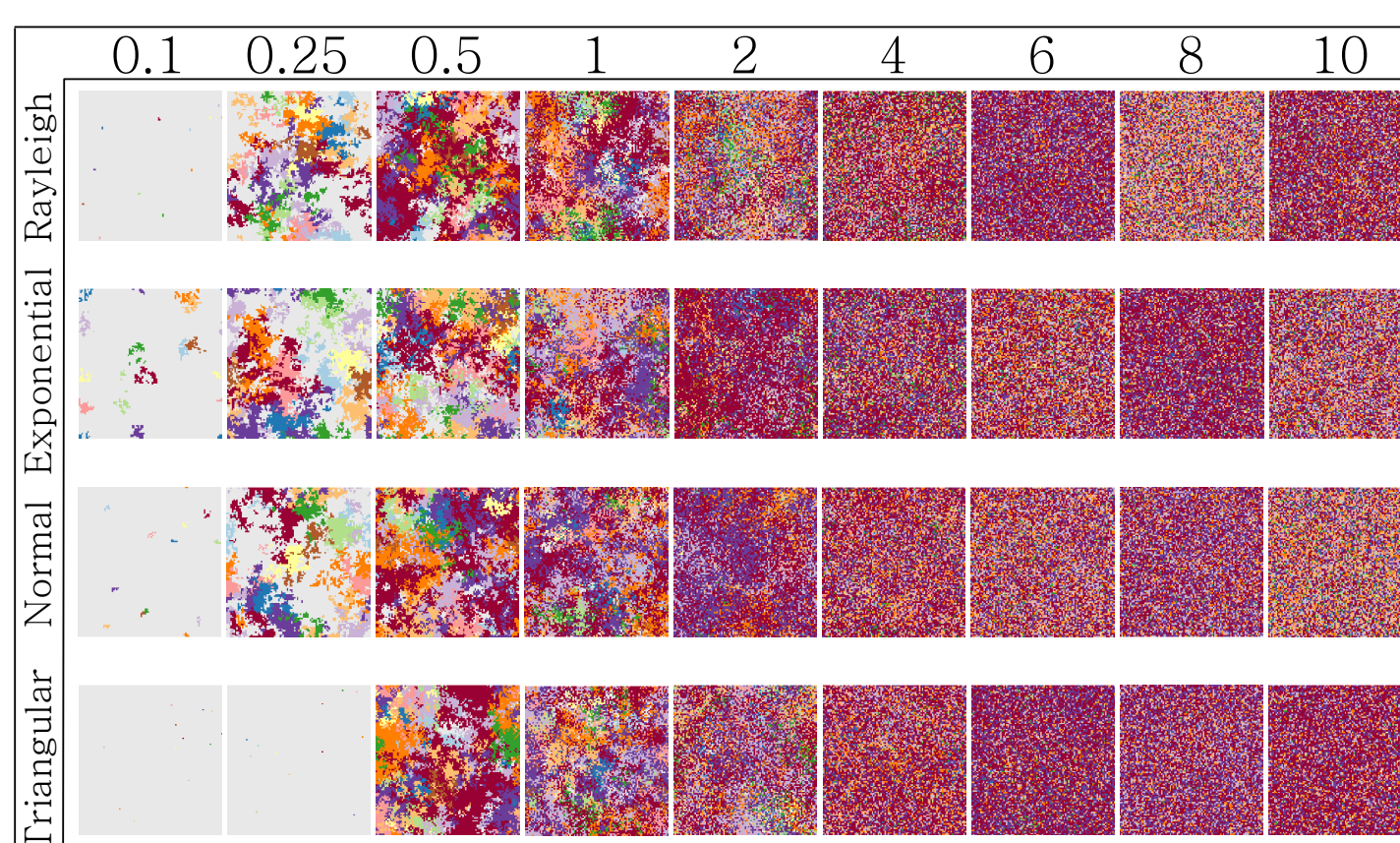


Figure 4: The different dispersal functions show similar patterns of spatial structure in a 100x100 landscape. Each box depicts a different simulation at 50,000 generations after a 10,000 generation burn-in period. The top 12 high frequency alleles are represented by different colored pixels and the rest of the alleles are grey.

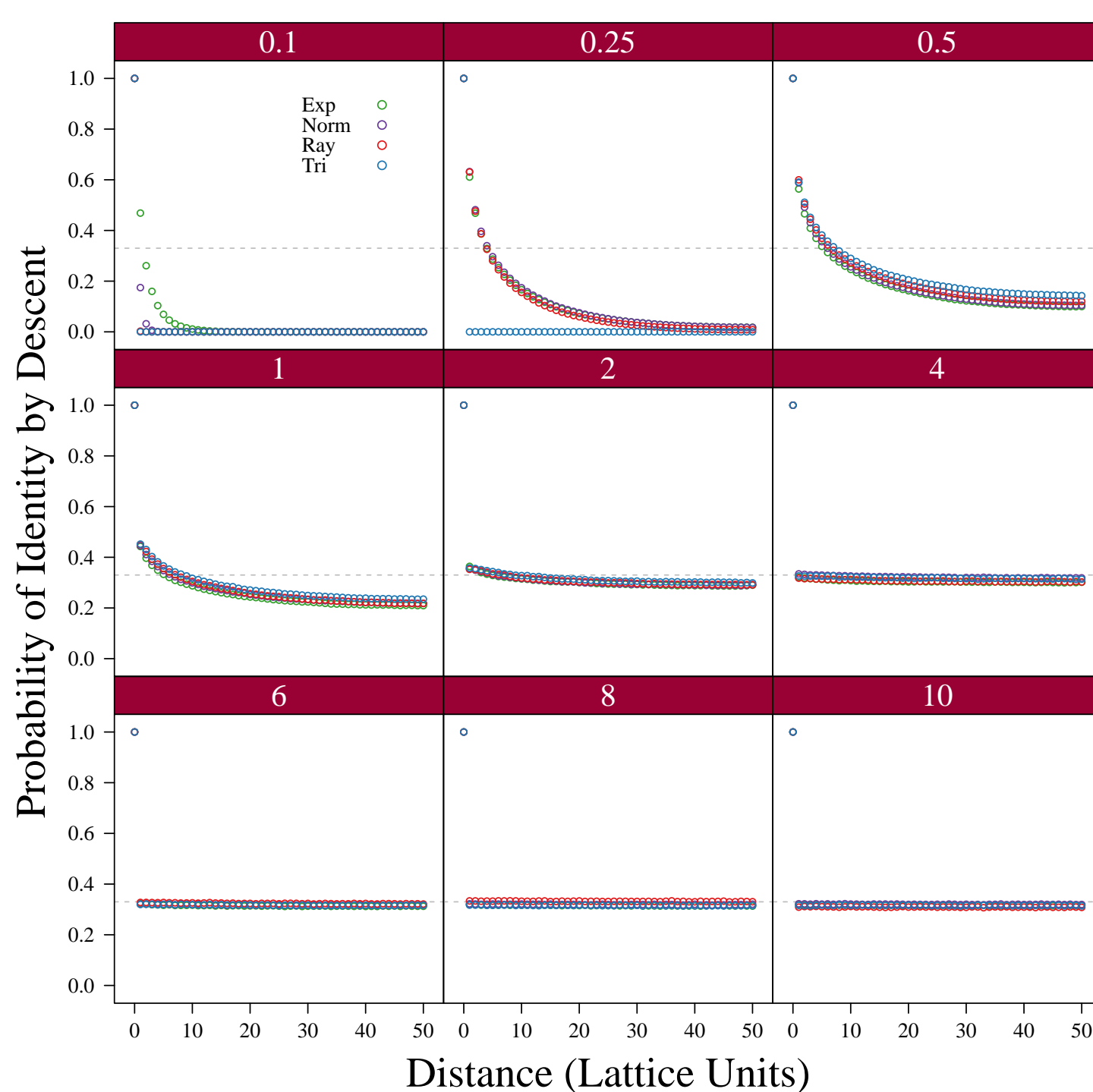


Figure 5: The different dispersal functions show similar patterns of isolation by distance in a 100x100 landscape. Every possible pairing of individuals in the sampled transect were placed in classes according to the minimal distance between them. The probability of identity-by-descent among each distance class was averaged over 2,000 sampled generations for the four dispersal distributions over a range of σ parameters. The dashed line represents a panmictic population simulation.

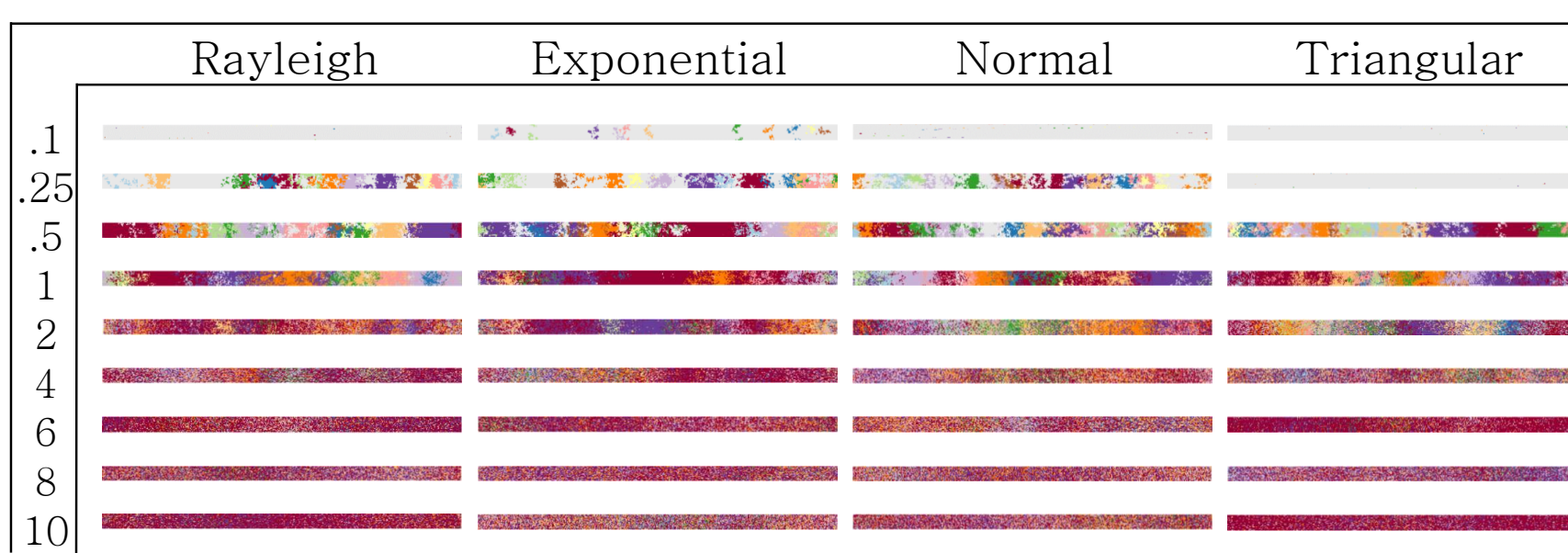


Figure 6: The different dispersal functions show similar patterns of spatial structure in a nearly-linear 20x500 landscape. Each box depicts a different simulation at 50,000 generations after a 10,000 generation burn-in period. The top 12 high frequency alleles are represented by different colored pixels and the rest of the alleles are grey.

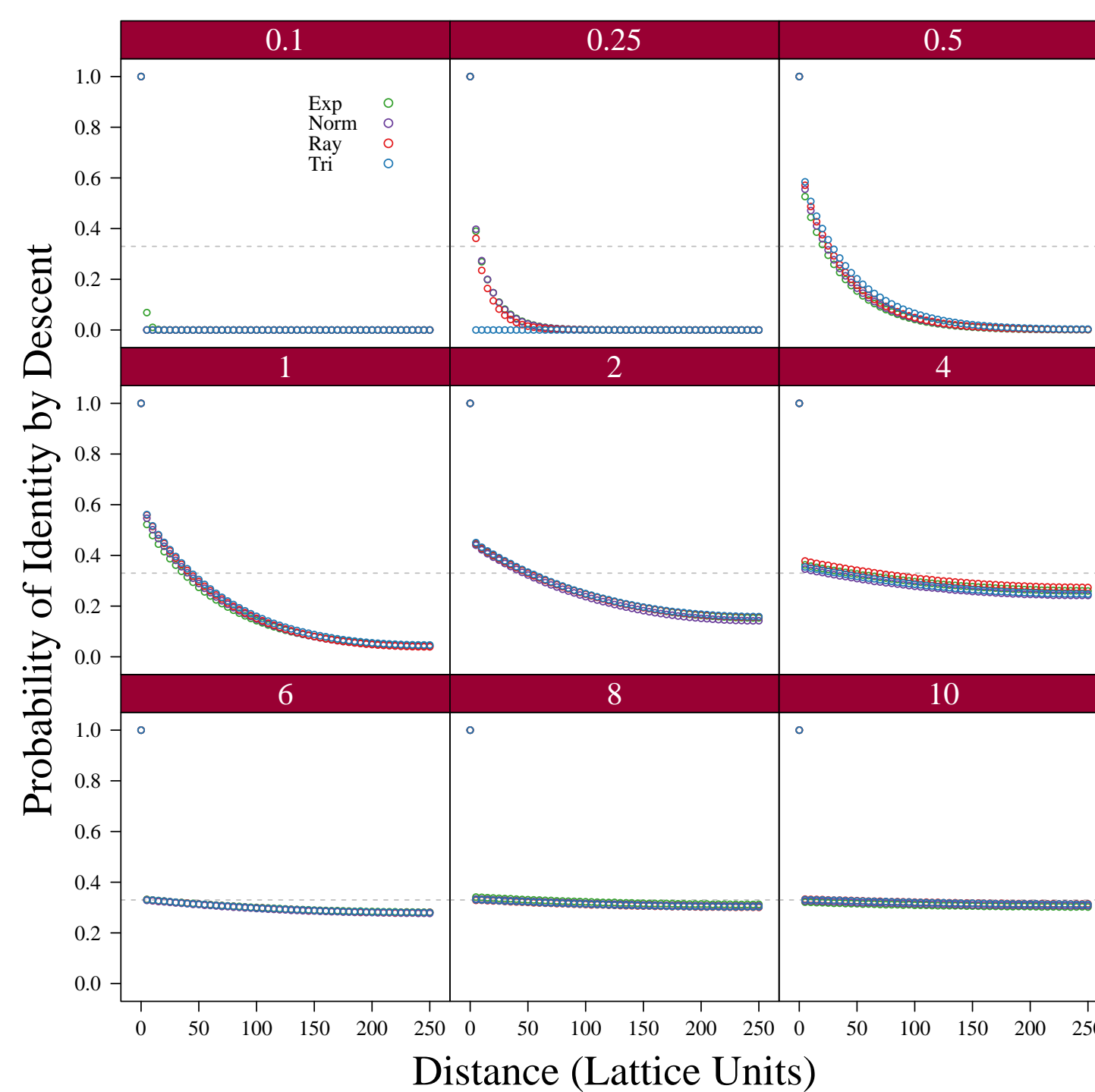


Figure 7: The different dispersal functions show similar patterns of isolation by distance in a 20x500 landscape. The probability of identity-by-descent among each distance class was averaged over 2,000 generations for the four dispersal distributions over a range of σ parameters. The dashed line represents a panmictic population simulation.

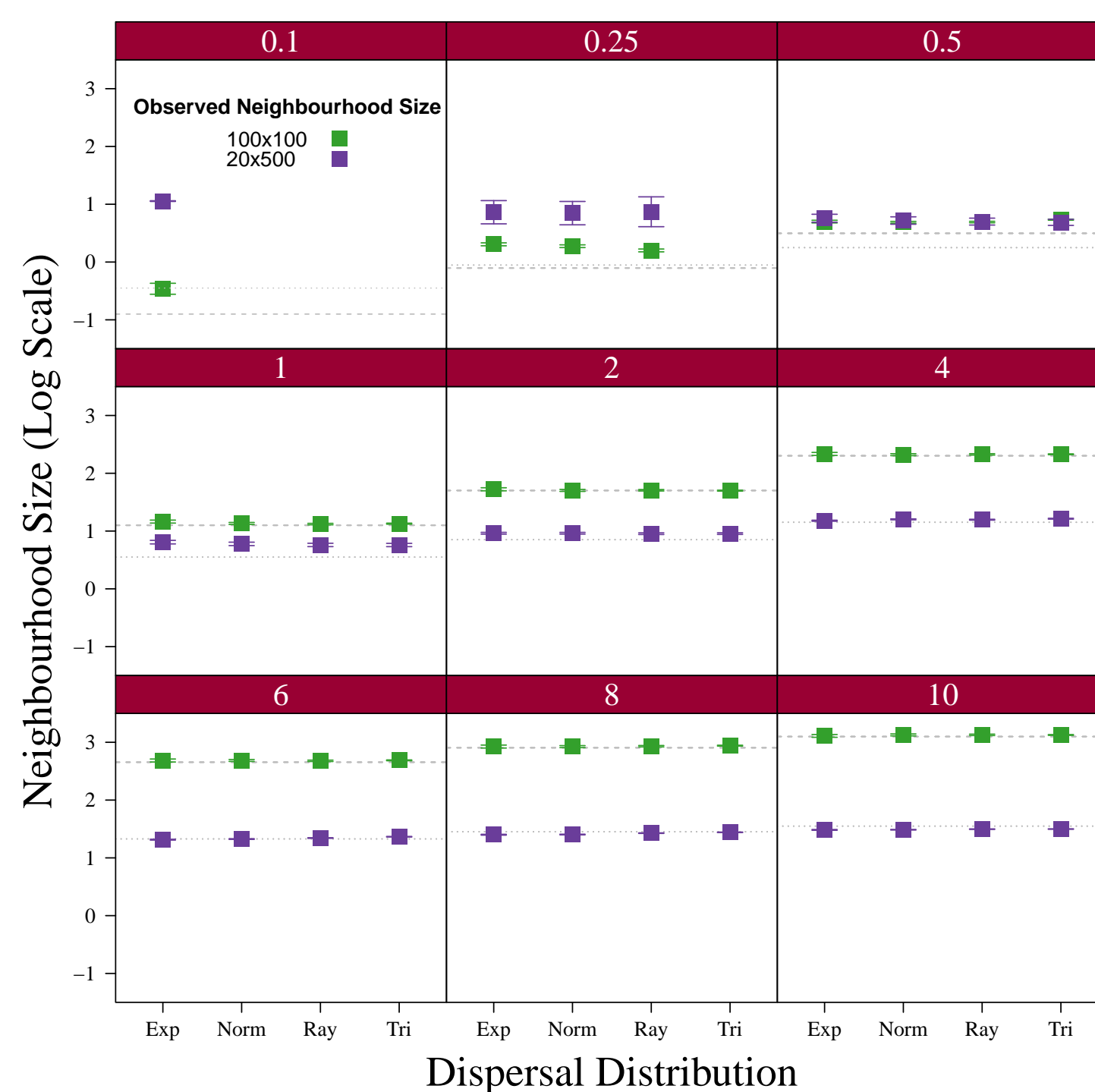


Figure 8: Observed and expected neighborhood size for 100x100 and 20x500 populations. The observed σ^2 was calculated as $1/2$ the mean squared distance of offspring from their parent. The effective population size N_e was calculated from observed allelic diversity via $\theta_k = 2N_e\mu$ (Ewens, 2004, eq. 9.32). The observed Neighborhood size is $N_b = 2\sqrt{\pi}\sigma N_e/A$ in one-dimension and $N_b = 4\pi\sigma^2 N_e/A$ in two-dimensions where $A = 10,000$ and error bars are SE over 2,000 samples. The grey lines are the expected neighborhood size for a one (dots) and two-dimensional (dashed) landscape.

Conclusions

- Each of the dispersal distributions show similar patterns of isolation-by-distance and spatial structure when σ is between 0.5 and 10.
- When sigma is less than 0.5, individuals dispersing under the bounded triangular distribution are unable to disperse outside of their original cell. When $\sigma = 0.1$, only the normal and exponential distributions had dispersal occur outside of the original cell.
- As expected, the decrease in identity-by-descent falls more rapidly with distance in the two-dimensional populations than in the nearly one-dimensional populations (Malécot, 1969).
- When simulating dispersal, the triangular distribution may be a good substitute for more computationally demanding distributions with exponential tails when σ is greater than 0.5.
- When dispersal follows a triangular distribution, it is uniform within an area of size $4\pi\sigma^2$. This suggests that the popular description of N_b as measuring a local panmictic population is valid for some families of dispersal.

Acknowledgments

T. N. Furstenau was supported by startup funds provided to R. A. Cartwright.

References

- Ewens, W. J. 2004. *Mathematical Population Genetics I Theoretical Introduction*. Springer Science, 2 edn..
- Malécot, G. 1969. *The Mathematics of Heredity*. W.H. Freeman and Company.
- Rousset, F. 1997. Genetic Differentiation and Estimation of Gene Flow from *F*-Statistics Under Isolation by Distance. *Genetics* 145:1219–1228.
- Wright, S. 1946. Isolation by distance under diverse systems of mating. *Genetics* 31:39–59.