

Forum

Using Networks
to Connect
Individual-Level
Reproductive Behavior
to Population Patterns

Rebecca J. Safran,^{1,6,*}
Iris I. Levin,^{1,2,6}
Bailey K. Fosdick,^{3,6,@}
Molly T. McDermott,¹
Georgy A. Semenov,¹
Amanda K. Hund,^{1,4,@}
Elizabeth S.C. Scordato,^{1,5} and
Sheela P. Turbek¹

We propose an application of network analysis to determine which traits and behaviors predict fertilizations within and between populations. This approach quantifies how reproductive behavior between individuals shapes patterns of selection and gene flow, filling an important gap in our understanding of the connection between evolutionary processes and emergent patterns.

Analyzing Fertilizations Provides Insights into Mechanisms Underlying Population Patterns

Mate selection decisions drive gene flow within and between populations of sexually reproducing organisms and are thus an important aspect of the speciation process [1–3]. Across animal taxa, mating systems vary in terms of the presence or absence of stable social pair bonds, as well as the extent to which fertilizations occur through within- and extra-pair matings. Accounting for fertilizations between all possible male–female pairs is essential for estimating patterns of sexual selection within populations that are not genetically monogamous. Further, analyses of fertilizations in **hybrid zones** (see [Glossary](#))

provide the opportunity to examine the relative importance of hybridization as a mechanism of gene flow. Yet, analyses of fertilizations alone do not take into account the landscape of mating opportunities experienced by both members of a mating pair. Missing from most studies are the details of phenotypic and behavioral variation associated with individuals who are and are not connected through shared fertilizations, as these provide information about the context in which fertilizations occur.

Recently, **network analyses** have been applied to studies of mate choice to understand **reproductive interactions** and their consequences for sexual selection within populations (e.g., [4,5]), including the influence of the social environment on extra-pair mating decisions [6]. Here, we propose the use of network methods to link fertilizations to emergent properties of populations, including phenotypic variation and the degree to which populations hybridize.

Most studies of reproductive behavior analyze patterns of mate and fertilization choice from the perspective of one of the sexes (e.g., hypotheses related to female choice). However, mating patterns are often influenced by properties of **dyads** (i.e., similarity/differences of potential pairs) as well as properties of the population (i.e., distribution of phenotypes, frequency of fertilizations). Network approaches can examine the influence of properties at all of these levels (individual, dyadic, and population level) on reproductive decisions to answer questions about whether and how **attributes** of individuals or dyads predict fertilizations. A network approach can thus incorporate details of individual and pair bond variation that are left out in traditional population summaries.

Studying reproductive behavior with network methods can answer a broader

Glossary

Assortative: nonrandom association of mating pairs by ancestry or phenotype. Positive assortative mating can contribute to reproductive barriers between hybridizing populations or result in disruptive selection within populations.

Attribute: features of an individual's phenotype or genotype (e.g., mate signaling traits, body size, age, social interactivity with other individuals, ancestry).

Dyad: two nodes connected by an edge in a network, such as a male and female that share fertilizations.

Edge weight: the degree to which nodes are connected in a network. Edge weight could represent the number of genetic offspring between a male and a female in a fertilization network or the number of social interactions between individuals in a social network.

Hybrid zone: a geographic region where two genetically distinct populations or species are found in spatial contact and produce offspring of mixed ancestry.

Network analysis: using statistical tools to make population inferences from observed network structure, such as how connectivity between nodes is associated with node attributes. Applied to natural populations, network analysis can be used to estimate which aspects of phenotype influence reproductive transactions.

Node attribute: characteristic of an individual (e.g., ancestry coefficient or phenotypic trait value).

Reproductive interactions: mating interactions, such as courtship behavior or copulation, which may or may not result in fertilization events (i.e., reproductive transactions).

Reproductive isolation: barriers to reproduction between individuals, resulting in barriers to gene flow between populations.

Reproductive skew: occurs when only a subset of individuals in a population gains reproductive transactions.

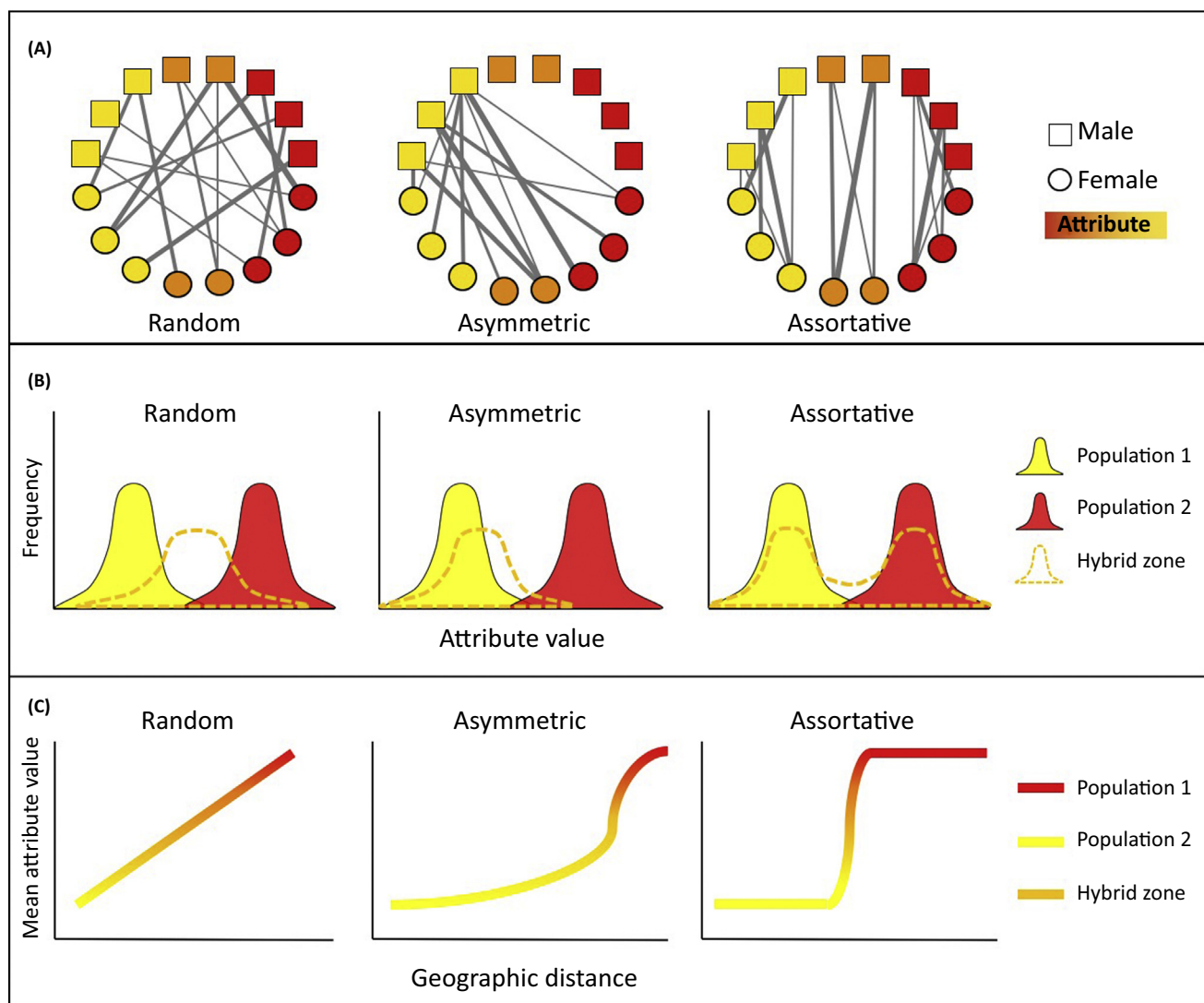
range of questions, including how phenotypic traits predict the presence and, just as critically, the absence of fertilizations between males and females. In addition, networks are excellent visualization tools for analyzing how individual behaviors aggregate into population patterns. When possible, including data on individual interactions through the construction of a social network can add important insights into how the social environment influences fertilizations both within and

outside of the pair bond [6]. Thus, the construction of fertilization networks provides the opportunity to test hypotheses about how behavior and phenotype shape, and are shaped by, fertilizations within and between closely related populations.

Process to Pattern in Hybrid Zones

Hybridization is a potential outcome when diverging populations come into secondary contact. Although much is known about patterns of mate selection and gene flow within populations, we know

relatively little about the processes that affect reproductive barriers between closely related populations. An explicit analysis of fertilizations in a network context is key for identifying whether and how mate choice contributes to **reproductive isolation**.



Trends in Ecology & Evolution

Figure 1. Fertilization Networks Illustrate the Processes That Underlie Different Population Patterns, Including the Degree to Which Populations Hybridize. (A) Portrays three different fertilization networks to illustrate cases of random, asymmetric, and positive assortative mating. (B) Shows population-level patterns (e.g., mean attribute value of phenotype or genotype in a hybrid zone center) associated with each fertilization network. (C) Portrays phenotypic or genetic variation across a geographic cline, another population pattern associated with each fertilization network shown in panel A. A table of predictions is indicated below panels A, B, and C to directly indicate process-pattern connections (Table 1). For example, a fertilization network showing random mating will result in a unimodal distribution of phenotypes and genotypes and, in a hybrid zone context, unrestricted gene flow (i.e., broad geographic cline that follows a pattern of isolation by distance).

Table 1. Connecting Process to Pattern in Studies of Gene Flow within and between Populations

Network structure	Pattern	Within population	Between populations	
		Selection	Gene flow	Reproductive isolation
Similarity of node attributes does not predict edge weights	Random mating	Absence of sexual selection	High	Low
Nodes(s) on one end of the attribute spectrum have higher edge weights than expected	Asymmetric; reproductive skew	Directional	Directional, asymmetric	Variable
Nodes with similar attributes have higher edge weights than expected	Positive assortative mating	Disruptive/divergent	Low	High

Typically, hybrid zones are described by patterns of phenotypic and genomic differences that can be visualized as either frequency distributions of phenotypes and genotypes in a given location [7] or gradients (clines) across space (Figure 1). Figure 1, Table 1 and Box 1 show how we suggest connecting pattern to process using networks (e.g., associating asymmetry in mating decisions with asymmetry in introgression). Panel A shows cartoon fertilization networks that underlie different mechanisms of mate selection: random, asymmetric (**reproductive skew**), and **assortative** ('like with like'). Whether fertilizations are uniformly or unevenly distributed among individuals in the population can dramatically influence the phenotypic and genetic variation of the next generation (Figure 1B), as well as associated evolutionary processes at work (selection, and degree and direction of gene flow; Figure 1C).

Suggested Workflow

We propose a layered network approach [8] that combines direct measures of fertilizations and attributes of pairs of individuals. Networks are most informative when they include all potentially interacting individuals. For fertilization networks, this means careful tracking of reproductive behavior and thorough genetic and phenotype sampling of all adults and offspring. When

possible, collecting data on social interactions provides important information on the social context in which fertilizations take place. Further, including samples from nonsurviving offspring are helpful for estimating postreproductive incompatibilities.

An example fertilization network is shown in Figure 1. Males and females are represented by 'nodes', and successful fertilizations between a male/female pair are portrayed as 'edges' (i.e., connections) between nodes. **Edge weight** (thickness) indicates the number of genetic offspring between a given male and female. The mathematical representation of a network is a matrix, in which rows are indexed by females, columns by males, and each entry represents the number of fertilizations between the corresponding pair (see Figure 1A in Box 1).

To study factors affecting fertilizations, the fertilization matrix can be compared with a set of similarity (or distance) matrices constructed from different **node attributes**. Entries in these matrices represent pairwise similarity (or dissimilarity) with respect to a variety of traits, behaviors, or genetic relationships (ancestry). Statistical analyses tailored to networks can then reveal the degree to which each attribute is predictive of fertilizations (see Figure 1B in Box 1). In collection, these analyses can be used to determine which

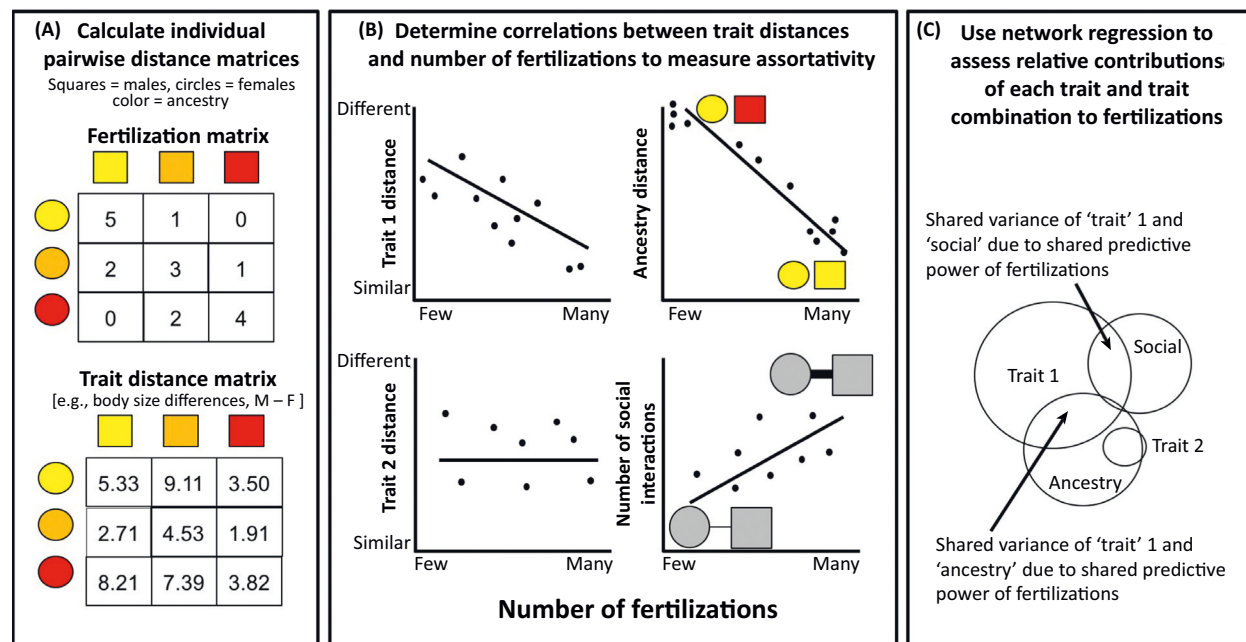
axes of phenotypic differentiation promote or impede hybridization, as well as directly examine their relative contribution to patterns of gene flow (see Figure 1C in Box 1).

The key statistical challenge when comparing fertilization and distance matrices is that entries associated with the same individual are not independent. This issue can be accounted for statistically through permutation tests. Here, we propose a novel application of the permutation-based quadratic assignment procedure to a bipartite network [9] that can be used to assess relationships between the behavior and phenotype matrices, and the fertilization matrix (see Figure 1B in Box 1) (see the supplemental information online). To assess statistical significance, the observed correlation between the behavior or attribute and the fertilization network is compared with correlations obtained when the rows and columns of one of the matrices are permuted. If none of the tests of association for the behavior or phenotype attributes are significant, this suggests random mating, whereas significant correlations suggest nonrandom mating.

To quantify the relative importance of behavior and phenotype attributes to fertilizations, the fertilization matrix and behavior (pairwise social interactivity)

Box 1. Workflow for Connecting Process to Pattern Using Fertilization Networks

Once a fertilization network has been constructed using genetic data from all potentially interacting individuals in a population and their offspring (Figure 1A), we suggest a workflow for analyzing relationships between fertilization networks and individual attributes (e.g., traits, ancestry, and social interactivity). Figure 1B shows several correlations derived from trait distance and fertilization matrices. The cartoon example in Figure 1C shows how the number of fertilizations among individuals within a hybrid zone is explained by a combination of trait 1, social, and ancestry variables, controlling for shared variation among these traits. Due to its correlation with ancestry, trait 2 has little effect *per se*. Size of circle = amount of variation in fertilizations explained by matrix regression analyses.



Trends in Ecology & Evolution

Figure 1. Cartoon Workflow for Analyzing Fertilization Data Using Network Models. (A) The image shows how network data are transformed into matrices to analyze all fertilizations (including none) between all males and females in the population, as well as additional matrices to describe pair-wise distances for attributes of interest (e.g., behavior, morphology, ancestry). (B) An example relationships between fertilizations and network attributes where slopes that differ from zero indicate patterns of assortativity with respect to that trait. Finally, (C) Cartoons to portray overlapping and independent predictors of fertilization patterns. In this case, when controlling for various attributes, including trait 1, trait 2, and information about an individual's social interactions and genomic ancestry, the model reveals that trait 1, ancestry, and social interactions explain variation in fertilization patterns, whereas variation in trait 2 is completely overlapping with ancestry.

and phenotype matrices are analyzed simultaneously using multiple matrix regression (see Figure 1C in Box 1). These methods estimate the importance of each behavior and phenotype predictor for fertilizations, given all other predictors in the model. These methods carefully account for correlation among observations involving the same individual to provide accurate confidence intervals for the relative association between each attribute and the fertilization network, as well as *P* values for testing these effects [10].

Software is available in the R package 'netregR' [11].

Concluding Remarks

Network analyses provide a comprehensive method for examining mate selection in the context of all possible reproductive interactions. Here, we suggest a workflow to study interactions between social behavior, phenotype variation, and fertilizations within populations, as well as link reproductive behavior to the degree of gene flow in hybrid zones.

Supplemental Information

Supplemental information associated with this article can be found in the online version at <https://doi.org/10.1016/j.tree.2019.03.009>.

¹Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309, USA

²Department of Biology, Agnes Scott College, Decatur, GA 30030, USA

³Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA

⁴Department of Ecology, Evolution, and Behavior, St. Paul, MN 55108, USA

⁵Department of Biological Sciences, California State Polytechnic University, Pomona, CA 91768, USA

⁶Co-first authors

®Twitter: @baileyfosdick, @AmandaHund

*Correspondence:

rebecca.safran@colorado.edu (R.J. Safran).

<https://doi.org/10.1016/j.tree.2019.03.009>

© 2019 Elsevier Ltd. All rights reserved.



References

1. Safran, R.J. *et al.* (2013) Contributions of natural and sexual selection to the evolution of premating reproductive isolation: a research agenda. *Trends Ecol. Evol.* 28, 643–650
2. Culumber, Z.W. *et al.* (2014) Assortative mating and the maintenance of population structure in a natural hybrid zone. *Am. Nat.* 184, 225–232
3. Kopp, M. *et al.* (2018) Mechanisms of assortative mating in speciation with gene flow: connecting theory and empirical research. *Am. Nat.* 191, 1–20
4. Moore, J.A. *et al.* (2015) Social network analysis of mating patterns in black bears (*Ursus americanus*). *Mol. Ecol.* 24, 4010–4022
5. McDonald, G.C. and Pizzari, T. (2016) Why patterns of assortative mating are key to study sexual selection and how to measure them. *Behav. Ecol. Sociobiol.* 70, 209–220
6. Maldonado-Chaparro, A.A. *et al.* (2018) Linking fine-scale social environment to mating decisions: a future direction for the study of extra-pair paternity. *Biol. Rev. Camb. Philos. Soc.* 93, 1558–1577
7. Jiggins, C.D. and Mallet, J. (2000) Bimodal hybrid zones and speciation. *Trends Ecol. Evol.* 15, 250–255
8. Silk, M.J. *et al.* (2018) Can multilayer networks advance animal behavior research? *Trends Ecol. Evol.* 33, 376–378
9. Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220
10. Dekker, D. *et al.* (2007) Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika* 72, 563–581
11. Marrs, F.W. *et al.* (2018) Standard errors for regression on relational data with exchangeable errors. *arXiv* 1701.05530