RICE UNIVERSITY

Quantitative Genetics and Evolutionary Consequences of Variation in Social Group Structure

By

Eric Wesley Wice

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Julia Saltz
Chair
Associate Professor of BioSciences

Scott Egan
Associate Professor of BioSciences

Volker Rudolf
Professor of BioSciences

Frederick Oswald
Professor of Psychology
Herbert S. Autrey Chair in Social Sciences

HOUSTON, TEXAS
August 2021
TITLE
Quantitative Genetics and Evolutionary Consequences of Variation in Social Group Structure

AUTHOR
Eric Wesley Wice

ABSTRACT
The structure of social groups - i.e., the patterning of interactions amongst individuals - is highly variable, and can be important for nearly every aspect of an individual’s behavior and fitness. How individuals are nested within social groups can be modeled using network analysis, which quantifies how direct and indirect social interactions shape the structure of social groups. While an individual’s position within a network - i.e., network position - is a frequent target of selection, we understand little about when social structure can respond to selection and evolve. For social structure to evolve, it must have a genetic basis. Yet estimating the genetic basis of network structure is not straightforward, because an individual’s position within a network is inherently dependent on interactions between multiple individuals. To understand the genetic basis of network structure, we need to know how individuals’ own genotypes and the genotypes of interacting partners shape social structure. Testing these genetic components of variation is empirically challenging, as it requires controlling and replicating the genotypes of all social group members. Despite these challenges, studying the genetic basis of social group structure is necessary for furthering our understanding of the quantitative genetic causes and evolutionary consequences of variation in social group structure. *Drosophila melanogaster* flies provide a powerful system to address questions about the quantitative
genetics and evolutionary consequences of social group structure, as we can control and replicate the genotypes of individuals in social groups and measure fitness effects across the lifespans of individuals. Using a combination of motion tracking software to quantify social interactions, and network analysis to model social group structure, I address how direct and indirect genetic effects shape variation in social structure. I also address how selection acts on genetic variation in network positions, across variable social and environmental contexts. My research program has found that both direct and indirect genetic effects shape the network structure of flies, and selection on social structure varies depending on the environmental context social groups experience. These findings suggest that genetic variation in social structure is pervasive, and can be adaptively maintained due to context-dependent selection operating on social group structure.
ACKNOWLEDGEMENTS

I thank my advisor and chair of my committee, Dr. Julia Saltz; as well as my other committee members, Drs. Scott Egan, Fred Oswald, and Volker Rudolf. I thank my lab mates: Madeline Burns, Marina Hutchins, Nick Keiser, Tracey Douglas, Charlotte Hovland, Raleigh Anderson, and Adam Geiger. An especially big thanks to all of the undergraduate and high school students who have worked with me over the years: Nicole Koonce, Ellie Dullea, Simone Maddox, Quynhmai Tran, Sophie Streeter, Jared Beshai, Claudia Xian, Shreyas Satewar, Tracey Dibbs, Anna Girardeau, Lilly Snellman, and Tyler Sakaweeny. Seeing you all grow as scientists has easily been the most rewarding part of graduate school. I would like to thank Rice University, the Weiss School of Natural Sciences, the Department of Biosciences, and the Rice Counseling Center. And lastly, I would like to thank the following institutions for financial support: Rice University, the Society for the Study of Evolution, the Animal Behavior Society, the National Science Foundation, the Houston Livestock Show and Rodeo, the Rice University Department of BioSciences, and the Rice University Institute of Biosciences and Bioengineering.
# TABLE OF CONTENTS

Title Page ........................................... i

Abstract ............................................. ii - iii

Acknowledgments ...................................... iv

Table of Contents .................................... v

Chapter 1: Selection on heritable social network positions is context-dependent in *Drosophila melanogaster* ........................................... 1 - 43

Chapter 2: Indirect genetic effects for social network structure in *Drosophila melanogaster* ........................................... 44 - 79

Chapter 3: Frequency-dependent selection and selection on social network structure also depends on indirect genetic effects in *Drosophila melanogaster* ........................................... 80 - 107
Chapter 1: Selection on heritable social network positions is context-dependent in *Drosophila melanogaster*

Eric Wesley Wice* and Julia Barbara Saltz

1 Department of Biosciences, Rice University, Houston, TX 77005

*Corresponding author (eric.wesley.wice@gmail.com)

**ABSTRACT**

Social group structure is highly variable and can be important for nearly every aspect of behavior and its fitness consequences. Group structure can be modeled using social network analysis, but we know little about the evolutionary factors shaping and maintaining variation in how individuals are embedded within their networks (i.e. network position). While network position is a pervasive target of selection, it remains unclear whether network position is heritable and can respond to selection. Furthermore, it is unclear how environmental factors interact with genotypic effects on network positions, or how environmental factors shape selection on heritable network structure. Here we show multiple measures of social network position are heritable, using replicate genotypes and replicate social groups of *Drosophila melanogaster* flies. Our results indicate genotypic differences in network position are largely robust to changes in the environment flies experience, though some measures of network position do vary across environments. We also show selection on multiple network position metrics depends on the environmental context they are expressed in, laying the groundwork for better understanding how spatio-temporal variation in selection contributes to the evolution of variable social group structure.
INTRODUCTION

From classic examples of eusocial animals, to species that are commonly thought of as solitary, the diversity of social structures evident in nature fascinates biologists and naturalists. Characteristics of social group structure emerge from the patterning and organization of social interactions occurring between members of a social group, which are commonly modeled using social network analysis\(^1\)–\(^3\). How individuals are embedded within their social groups, termed “social network position”, has been demonstrated to influence many critical aspects of an individual’s life (mating opportunities\(^4\),\(^5\), risk of disease\(^1\),\(^6\), access to social information\(^7\),\(^8\), foraging potential\(^8\),\(^9\), etc.). Despite this importance and pervasive evidence that social network positions are under selection, we still know relatively little about the underlying causes of variation in social network positions and group structure. Such an understanding is necessary for identifying ways that social structures may form and evolve, provoking recent calls for studies investigating the mechanisms underlying social network structure\(^10\)–\(^12\). Potential mechanisms that could give rise to variation in social network structure include the genotypes of individuals comprising a social group, as well as the environmental conditions in which social interactions occur\(^13\),\(^14\). However, very few studies have investigated whether social network positions of individuals within a group are heritable\(^10\),\(^11\),\(^15\)–\(^20\), or how the structure of a social network varies across environments\(^12\),\(^15\),\(^21\). Furthermore, we know even less about how selection shapes heritable variation in social network structure over generations\(^22\).

One of the challenges of studying the genetic basis of and how selection acts on social network positions is that network position is not inherent to the individual, and is
instead an extended phenotype dependent on the direct and indirect interactions of conspecifics in an individual’s social group\textsuperscript{23,24}. Thus, to measure the genetic basis of social network position, one has to measure focal individuals of known genotype within a replicate social group comprised of the same conspecifics\textsuperscript{25}. Repeatedly measuring the same social group in the wild presents additional challenges, as prior environmental and social experiences can induce plasticity and variation in social network positions\textsuperscript{6,26,27}, and/or reinforce individual differences in social network positions\textsuperscript{20,26,28–31}. These empirical challenges leave gaps in our knowledge about whether social network positions are heritable traits that can respond to selection, as no prior study has estimated the heritability of individuals’ network positions while also controlling for the genotypes of their social partners and prior experience. Furthermore, no prior study has investigated the heritability of network positions in conjunction with environmental variation, and no prior study has addressed how selection acts on network positions in tandem with known genotypic and environmental influences. Studying the causes of variation in social network structure, including how genotypic and environmental influences align with patterns of selection on social network positions, is crucial for understanding how diverse patterns of social group structures form and evolve\textsuperscript{15,22}.

In the present study, we identify how genetic, environmental, and genotype-by-environment interaction effects contribute to variation and plasticity in commonly studied social network positions. We also address how selection acts on social network positions, across environmental contexts, to shape and adaptively maintain genetic variation in network structure\textsuperscript{32–36}. 
Drosophila melanogaster flies are an ideal study system for investigating the genetic and environmental causes of social network positions, in tandem with the fitness consequences of social network positions. Flies form non-random social networks\textsuperscript{7,13} and spend the majority of their lives living and socially interacting on nutritionally-variable rotting fruit environments\textsuperscript{37}. We can control and replicate the genotypes and environments of all individuals within a social group, allowing us to create replicate social groups with genetically identical composition and measure their social structure across variable environments\textsuperscript{38}. We created 98 replicate social groups of flies. Each social group consisted of 20 unrelated, heterozygous genotypes bred from the Drosophila Genetic Reference Panel (DGRP) (10 males and 10 females per replicate group)\textsuperscript{38}. Each replicate social group was placed on one of five nutritional environments varying in either protein-to-carbohydrate ratio or caloric concentration of the food, both of which have been shown to affect various components of fly behavior and fitness\textsuperscript{39–41}. Social groups were video recorded twice over the course of two days, and social networks were generated using automated motion tracking software, resulting in 56 weighted and directed networks comprising over 600,000 seconds of social interactions\textsuperscript{42}. We analyzed the five most commonly studied social network positions for each individual within a social group\textsuperscript{3}: Instrength and outstrength - the amount of time other individuals spend socially engaging with a focal individual, and the amount of time a focal individual spends socially engaging with other individuals, respectively; Clustering coefficient - how interconnected a focal individual’s direct social partners are to one another (i.e. cliquishness); Betweenness centrality - the number of shortest paths between any two individuals that transverse a focal individual; and Eigenvector centrality - how critical a
focal individual is to the overall structure of the group based on the strength of its direct
social connections, the strength of its partners’ connections, the strength of its partners’
partners’ connections, etc.\textsuperscript{1,2}. We also measured multiple metrics of fitness for each
individual, allowing us to estimate the strength and direction of selection on social
network positions. These fitness measures include the total number of observed matings
and the latency to mate for males, and the lifetime reproductive success and lifespan of
females (Supplementary Fig. 1).

In this work, we find that genotype significantly affects all five measures of social
network position, with low to moderate estimates of broad-sense heritability. We also
find effects of sex, genotype-by-environment interactions, and sex-by-environment
interactions on many measures of network position. Selection on network positions is
limited to male flies, and the strength and direction of selection on measures of network
position varies depending on the nutritional environment. Spatio-temporal variation in
selection, combined with heritable variation for traits under selection, shapes and
maintains genetic variation over generations\textsuperscript{32,36}. These findings not only suggest that the
structure of social groups can respond to selection and evolve, they also shed light on
how variation in social structure may evolve under variable environmental conditions.

RESULTS

Social Network Positions

We discovered that an individual’s genotype was a significant predictor for all five
network positions analyzed (likelihood ratio tests: instrength, $LRT = 22.713, P_R < 0.001$;
outstrength, $LRT = 24.13, P_R < 0.001$; clustering coefficient, $LRT = 54.726, P_R < 0.001$;
betweenness centrality, $LRT = 1086, P_R = 0.022$; and eigenvector centrality, $LRT =$
24.348, \( P_R < 0.001 \); Fig. 1). Estimates of broad sense heritability (\( H^2 \)), defined as the extent to which genotypic differences explain variation in a phenotype\(^{43} \), ranged from 2.4 - 16.6% for the five network positions analyzed (instrength, \( H^2 = 0.024 \); outstrength, \( H^2 = 0.025 \); clustering coefficient, \( H^2 = 0.050 \); betweenness centrality, \( H^2 = 0.166 \); and eigenvector centrality, \( H^2 = 0.042 \)). We also found significant differences between the sexes in four of the five network positions analyzed (type III Wald \( \chi^2 \) tests: instrength, \( \chi^2 = 8.072, P_R = 0.002 \); outstrength, \( \chi^2 = 10.108, P_R = 0.001 \); clustering coefficient, \( \chi^2 = 106.832, P_R < 0.001 \); and eigenvector centrality, \( \chi^2 = 9.248, P_R = 0.006 \); Fig. 1), but not for betweenness centrality (\( \chi^2 = 6.208, P_R = 0.683 \); Fig. 1). As expected, females received more social interactions (instrength) than males, and males initiated more social interactions (outstrength) than females. Males also tended to be more central to the overall structure of the social group (eigenvector centrality), while females tended to be more cliquish (clustering coefficient) compared to males (Fig. 1).
We saw no effects of genotype-by-environment interactions on the social network positions of instrength, outstrength, betweenness centrality, and eigenvector centrality; meaning that genotypic differences in these measures of network position remained constant across the nutritional contexts (all $P_R > 0.05$; Supplementary Table 1). However, we did find evidence of genotype-by-environment interactions influencing how cliquish (clustering coefficient) individuals were ($LRT = 5.000, P_R = 0.007$) across nutritional environments that varied in both protein:carbohydrate ratio ($LRT = 1.843, P_R = 0.001$) and caloric concentration ($LRT = 1.047, P_R = 0.028$). When selection acts on traits that

\[ H^2 = 0.024 \]
\[ H^2 = 0.025 \]
\[ H^2 = 0.050 \]
\[ H^2 = 0.166 \]
\[ H^2 = 0.042 \]
are influenced by genotype-by-environment interactions, the genotype(s) with the highest relative fitness depends on the environmental context in which the trait is expressed. Thus, genotype-by-environment interactions are one of the primary mechanisms invoked to explain the persistence of genetic variation in traits that are under selection. Our findings of genotype-by-environment interactions for how cliquish individuals are (clustering coefficient) provides a potential mechanism for explaining the maintenance of genetic diversity in this trait. We also find sex-by-environment interactions for the network positions of outstrength and clustering coefficient (outstrength, $\chi^2 = 13.691, P_R = 0.002$; clustering coefficient, $\chi^2 = 38.130, P_R < 0.001$; Fig. 2). Specifically, males and females initiated similar amounts of social interactions (outstrength) when on nutritional environments that were low-calorie or protein-rich, but males initiated more social interactions than females on nutritional environments that were high-calorie or carbohydrate-rich (sex-by-caloric concentration interaction, $\chi^2 = 4.591, P_R = 0.033$; sex-by-P:C ratio interaction, $\chi^2 = 9.599, P_R < 0.001$; Fig. 2). Females increased how cliquish they were (clustering coefficient) in high-calorie and carbohydrate-rich environments, while males were less cliquish in low-calorie environments and maintained similar levels of cliquishness across environments of varying protein-to-carbohydrate ratio (sex-by-caloric concentration interaction, $\chi^2 = 33.250, P_R < 0.001$; sex-by-P:C ratio interaction, $\chi^2 = 25.890, P_R < 0.001$; Fig. 2).
The nutritional environment also played a significant role on its own in affecting instrength, outstrength, clustering coefficient, and variation in eigenvector centrality. Individuals interacted less overall (lower instrength and outstrength) on nutritional environments that were more carbohydrate-rich compared to protein-rich environments (instrength, $\chi^2 = 5.175, P = 0.023$; outstrength, $\chi^2 = 8.122, P = 0.004$), and individuals tended to be more cliquish (clustering coefficient) overall on high-calorie ($\chi^2 = 11.618, P$...
< 0.001) and carbohydrate-rich environments ($\chi^2 = 7.057, P = 0.008$). In high-calorie environments, we detected increased variation in how central individuals were to the structure of their social groups (Fligner-Killeen test: eigenvector centrality, $\chi^2 = 8.097, P = 0.017$). Notably, these patterns of genotypic, sex, genotype-by-environment, and sex-by-environment effects on network position could not be attributed to differences in the overall locomotor activity of each fly (Supplementary Table 2). For a sub-sample of social groups in which network structure was measured across two days, we found that individuals displayed significant consistency across days in four of the five network positions analyzed (Kendall’s rank correlations: instrength, $\tau = 0.167, z = 4.017, P < 0.001$; outstrength, $\tau = 0.222, z = 5.337, P < 0.001$; clustering coefficient, $\tau = 0.458, z = 11.007, P < 0.001$; and eigenvector centrality, $\tau = 0.091, z = 2.179, P = 0.029$) but not for betweenness centrality ($\tau = 0.024, z = 0.524, P = 0.6$), further supporting our inference that genetic variation contributed to individual differences in network positions.

For a trait to respond to selection and evolve, it must have an additive genetic basis\textsuperscript{44}. Our findings that genotypes significantly vary in all measures of social network position indicate network structure can likely evolve in response to selection. Prior studies estimating the heritability of social network positions have reached variable conclusions about which measures of network position are heritable\textsuperscript{16–18}. In wild populations, estimates of heritability may be confounded by prior experience\textsuperscript{11,34,39,45,46}, social group composition\textsuperscript{11,14,22,25,45,46}, and carryover effects (e.g. maternal effects or social inheritance)\textsuperscript{47,48}. Here, we demonstrate that social network positions are heritable under controlled laboratory conditions, and with direct replication of social groups of
natural genotypes. Interestingly, our broad-sense heritability estimates were generally comparable to prior heritability estimates of social network positions in the wild\textsuperscript{16-18}.

**Fitness and Selection on Social Network Positions**

To further probe the evolutionary potential of social networks, we quantified selection gradients for social network position, and how these gradients varied across nutritional environments. For males, we observed directional selection acting on network position measures of instrength and eigenvector centrality. For males, receiving more interactions (instrength) and being more central to group structure (eigenvector centrality) were associated with lower mating success overall (type III Wald $\chi^2$ tests: instrength, $\chi^2 = 8.795$, $P = 0.003$; and eigenvector centrality, $\chi^2 = 9.938$, $P = 0.002$; Supplementary Table 3). Additionally, males who received more interactions (instrength) had longer latencies to mate ($\chi^2 = 9.687$, $P = 0.002$). We also found that the strength and direction of selection, measured as the relationships between male fitness measures and instrength, outstrength, and eigenvector centrality, varied across nutritional environments (Fig. 3, Supplementary Table 3). These findings are particularly exciting, as context-dependent selection can maintain genetic variation when the strength and direction of selection on a phenotype changes depending on the environmental context in which that phenotype is expressed\textsuperscript{32,36}. In high-calorie environments, males who received more social interactions (instrength) or were more central to the structure of their social groups (eigenvector centrality) also showed both an increased observed mating success (instrength-by-caloric concentration, $\chi^2 = 7.098$, $P = 0.008$; and eigenvector centrality-by-caloric concentration, $\chi^2 = 9.859$, $P = 0.002$; Fig. 3) and shorter latencies to mate (instrength-by-caloric concentration, $\chi^2 = 10.367$, $P = 0.001$; and eigenvector centrality-by-caloric concentration, $\chi^2 = 9.938$, $P = 0.002$)
concentration, $\chi^2 = 10.628, P = 0.001$). This pattern was reversed in low-calorie environments, where receiving more interactions and being central to group structure was associated with lower mating success and longer latencies to mate. Males who initiated more social interactions (outstrength) had marginally shorter latencies to mate in high-calorie environments, but tended to have longer latencies to mate in low-calorie environments (outstrength-by-caloric concentration, $\chi^2 = 6.471, P = 0.011$; corrected significance threshold = 0.01). Similarly, males who were more central to the structure of their social group had higher mating success in carbohydrate-rich environments, but had lower mating success in protein-rich environments (eigenvector centrality-by-P:C ratio, $\chi^2 = 6.678, P < 0.010$; Fig. 3); and males who engaged in more social interactions (instrength and outstrength) had shorter latencies to mate in carbohydrate-rich environments, but had longer latencies to mate in protein-rich environments (instrength-by-P:C ratio, $\chi^2 = 7.317, P = 0.007$; outstrength-by-P:C ratio, $\chi^2 = 8.183, P = 0.004$). We also detected marginally significant effects of increased social interactions (instrength and outstrength) associated with greater male mating success in carbohydrate-rich environments, with the opposite trend occurring in protein-rich environments (instrength-by-P:C ratio, $\chi^2 = 6.478, P = 0.011$; and outstrength-by-P:C ratio, $\chi^2 = 5.733, P = 0.017$; corrected significance threshold = 0.01; Fig. 3).
Male genotypes also significantly differed in their mating success (likelihood ratio test: LRT = 99.76, $P < 0.001$) and latency to mate (LRT = 87.104, $P < 0.001$). Our findings indicate that higher rates of social interactions and centrality to group structure are beneficial in high-calorie and carbohydrate-rich nutritional environments, but costly in low-calorie and protein-rich environments, suggesting that social groups of flies in high-calorie and carbohydrate-rich environments would evolve to be more cohesive and interactive compared to groups in low-calorie and protein-rich environments.

**Figure 3 | Selection on network position is context-dependent.** Selection gradients, quantified as the correlations between male mating success and three network position metrics measured the first day after social groups were established, varied across nutritional environments that varied in protein:carbohydrate ratio (a-c) and caloric concentration (d-f). Lines show best fit linear relationships with standard error (shaded areas) (1:4-1:1, 4x-1x; $n = 40, 60, 40, 70, 50$, respectively). The colors of each line correspond to the given protein:carbohydrate ratio or caloric concentration. Source data are provided as a Source Data file.
While we found context-dependent and directional selection operating on social network positions in males, we found no effects of network positions on components of female fitness (all $P > 0.05$; Supplementary Table 3). We also observed no differences among female genotypes in lifespan (LRT = 0.16, $P = 0.689$) or lifetime offspring production (LRT = 0.04, $P = 0.842$). However, the nutritional environment social groups interacted on did influence female lifespan and offspring production, with females living longer from low-calorie and protein-rich environments (caloric concentration, $\chi^2 = 10.815, P = 0.001$; P:C ratio, $\chi^2 = 7.443, P = 0.006$) and producing more offspring after interacting on protein-rich environments (P:C ratio, $\chi^2 = 4.860, P = 0.028$). These results are largely consistent with past work showing caloric restriction and short-term stress increases lifespan, and elevated dietary protein increases reproduction\textsuperscript{40,49}. However, our findings are particularly noteworthy, as flies were only on their treatment nutritional environments for less than 3 days out of their total lifespan (average female lifespan 60 days $\pm 27$ standard deviation), and were otherwise reared and housed on standard lab food\textsuperscript{39}.

**DISCUSSION**

Our findings of context-dependent and directional selection operating on heritable network positions were evident only for males, and not for females. One potential hypothesis explaining this pattern is that females may be better able to adjust their behavior to adaptively match their environment, compared to males\textsuperscript{9,36,46,50,51}. This hypothesis is consistent with our findings of sex-by-environment interactions for network positions. An alternative hypothesis for our sex-by-environment interaction results could be that males and females perform different social functions depending on the
environment, irrespective of the fitness consequences of engaging in different social behaviors. This would align with previous work demonstrating sex-specific patterns of gene expression change in response to social and nutritional cues from the environment. While our study utilized a quantitative genetic approach to addressing how genotype influences measures of social network position, this genotype-to-phenotype relationship is likely complex and involves many intermediary links. Understanding the links between genotype and social network position is an active area of study, with recent work exploring everything from how genes for sensory processing affect network structure, to how individual behavioral differences inform network structure (e.g. exploratory differences, group size preferences, and sociability). Future work should continue to explore the genetic and behavioral underpinnings of network structure, as well as how these effects may vary across environmental contexts.

Across a diverse range of taxa, evidence that social network positions are under selection is mounting (mating success, disease risk, effects on foraging, review) in tandem with evidence for remarkable intra-population variation in these same traits. How can heritable variation in social network positions be maintained in the face of selection and drift? While we found some support for the hypothesis that genotype-by-environment interactions influence measures of social network position (namely how cliquish individuals were, i.e. clustering coefficient), we did not find evidence of selection acting on variation in cliquishness. However, we did find support for the hypothesis that context-independent genetic variation in network position, coupled with context-dependent selection, can maintain heritable genetic variation in measures of social network positions. Studies estimating how selection varies across space or time are
severely underrepresented for behavioral traits\textsuperscript{57}, and very few have obtained multiple measures of selection on social network positions\textsuperscript{58,59}. The strength of our methodology rests in our ability to create replicate social groups of identical genotypic composition. This allowed us to investigate processes that may be obscured by genetic variation in social group membership in natural populations. Our results, in conjunction with prior estimates of the heritability of social network positions\textsuperscript{16–18}, indicate that the evolutionary potential of social network structure may be a more widespread phenomenon than currently known. It is important to note that measures of social network position are all derived from the same underlying network of social interactions. Thus, if any single measure of network position is heritable, the entire underlying social network is likely to change in response to selection. Future studies investigating the heritability of social network positions across multiple species, populations, and social groups will hopefully bring us closer to understanding the evolutionary potential of social group structure. Continued work should also seek to address how genetic and environmental factors shape the evolution of complex social behaviors, to develop a better understanding of the behavioral, ecological, and evolutionary factors shaping diversity in social group structure. Our findings suggest that studying how both social and ecological factors interact to shape variation in and selection on social phenotypes is necessary for understanding the evolution of social dynamics.

**METHODS**

**Study System and Social Groups**

To create replicate social groups, 40 inbred lines of flies were randomly chosen from the *Drosophila* Genetic Reference Panel (DGRP), a panel of dozens of inbred homozygous
genotypes derived from a wild population\textsuperscript{38}. These 40 inbred lines were randomly divided between maternal and paternal lines, and 20 mating crosses were randomly generated to produce 20 replicate heterozygous genotypes (maternal lines 136, 208, 237, 306, 315, 335, 360, 387, 786, 852, 153, 235, 307, 378, 57, 59, 637, 732, 801, and 855; crossed with paternal lines 427, 313, 714, 774, 229, 365, 380, 324, 716, 721, 517, 712, 176, 765, 862, 820, 318, 40, 513, and 391; respectively). Note that parental line numbers are arbitrary and unindicative of relationships between genotypic lines. Using replicate heterozygous genotypes allows us to alleviate the potentially deleterious effects of homozygous recessive alleles in inbred lines, and generate individuals representative of a subset of naturally segregating genetic variation found in the wild\textsuperscript{60,61}. This approach thus balanced our need to create replicated social groups against the challenges of adequately representing natural genetic variation with inbred lines, which may or may not be perfectly representative of quantitative genetic parameters in wild populations. Ten males and ten virgin females of each inbred line were paired and placed in vials of standard fly food to control for maternal and larval density, and heterozygous offspring were allowed to develop for 2-3 weeks. Newly emerged virgin females and males were collected under light CO\textsubscript{2} anesthesia from the 20 mating crosses (10 replicate genotypes per sex). Each fly from each sex was randomly marked with a unique color identifier on its mesothoracic segment to visually identify each individual. Flies were aged in same-sex groups for three days to allow for development to sexual maturity and recovery from CO\textsubscript{2} anesthesia\textsuperscript{62,63}. All flies were reared and aged on a 12:12 light:dark cycle, at 24°C and 50% relative humidity, and on standard fly food. On the evening of the third day, the 10
males and 10 females were anesthetized via chilling and combined into a social group. Care and treatment of all flies complied with all relevant ethical regulations.

**Nutritional Environments**

Each social group was placed in a 10cm petri dish filled with one of five nutritional environments that varied in two dimensions of nutrient composition: protein-to-carbohydrate ratio (P:C ratio) and caloric concentration\textsuperscript{64,65}. Three nutritional environments were constant in the total amount of calories they contained, but varied in whether the calories were derived from protein or carbohydrates (1:1, 1:2, and 1:4 P:C ratio). To vary caloric concentration, three nutritional environments contained a constant ratio of P:C (1:4), but varied in caloric concentration by 4x, 2x, and 1x. Note that the 1:4 P:C ratio and the 4x caloric concentration environments are the same. All recipes contained a base of 27g agar, 11.1mL tegocept acid mix (70g tegocept/270 mL H\textsubscript{2}O), and 3mL propionic acid; per 1L H\textsubscript{2}O. Caloric concentration and P:C ratio were manipulated by adjusting the amounts and ratio of nutritional yeast and malt sugar added to the foods (1:1 P:C, 4x Concentration = 146.9g yeast and 45.6g sugar; 1:2 P:C, 4x Concentration = 97.9g yeast and 94.6g sugar; 1:4 P:C, 4x Concentration = 58.8g yeast and 133.7g sugar; 1:4 P:C, 2x Concentration = 29.4g yeast and 66.9g sugar; and 1:4 P:C, 1x Concentration = 14.7g yeast and 33.4g sugar)\textsuperscript{40,41}.

**Fly Behavior and Social Network Analysis**

Social groups were given a period of overnight acclimation, then video-recorded for 20 minutes during the hour immediately following lights-on (when flies are most active) over the course of two days\textsuperscript{66}. Videos were recorded with Nikon D3300 cameras at 30 fps and 9.97±0.50 pixels/mm. Each video was processed with the motion-tracking software
Caltech FlyTracker 1.0.5, which outputs the position and orientation of every fly in every frame of a video\textsuperscript{42}. We manually verified the tracking output for all individuals to ensure all tracking identities were consistent and accurate. We also manually validated the tracking output of the software by comparing hand-annotated fly positions to FlyTracker’s output in a stratified random subsample of 700 frames (correlation > 0.999).

Using the tracking output from FlyTracker, we calculated the weighted and directed social interactions occurring between every pairwise combination of flies to build a social interaction matrix for each video\textsuperscript{67}. A focal fly was considered interacting with another fly if three criteria were met: (1) the distance between the two flies was < 2.5 average fly body lengths, (2) the interacting fly was within a 320° field-of-view of the focal fly, and (3) if the aforementioned criteria were met for a minimum duration of 0.6 seconds\textsuperscript{68}. Note that our first criterion ensured that the radius in which flies were considered to be interacting was fixed for all individuals. For example, larger flies (i.e. females) were not assigned a larger space in which interactions were counted, which could bias our results. Our second criterion allows for one fly to be in another’s field-of-view, without reciprocally having the other fly in its field-of-view; thus, we can distinguish between the directedness of social interactions occurring. While proximity between individuals does not necessarily imply a social interaction occurred, these interaction criteria filter out random interactions, such as two flies walking quickly past one another without socially engaging\textsuperscript{68}. Edges in our interaction matrices were defined as the total duration of time all pairwise combinations of flies spent interacting throughout the duration of a video. Since all individuals were observed at all times throughout a video, the interaction networks we built allowed us to directly proceed with
network analysis without having to compute association indices to account for sampling error or bias\textsuperscript{69,70}. We chose to analyze the five most commonly studied individual-level social network positions using the R package igraph 1.2.4.2\textsuperscript{71}: instrength and outstrength - the amount of time other individuals spend socially engaging with a focal individual, and the amount of time a focal individual spends socially engaging with other individuals, respectively; weighted and directed clustering coefficient - how interconnected a focal individual’s direct social partners are to one another (i.e. cliquishness); weighted and directed betweenness centrality - the number of shortest paths between any two individuals that transverse a focal individual; and weighted and directed eigenvector centrality - how critical a focal individual is to the overall structure of the group based on the strength of its direct social connections, the strength of its partners’ connections, the strength of its partners’ partners’ connections, etc.\textsuperscript{1–3}.

**Measures of Fitness**

We measured two components of fitness for males and two components of fitness for females, to understand how selection acts on social network phenotypes across different environmental contexts. In the two hours immediately after males and females were combined into a social group, the identities of all mating pairs and their latency to mate were recorded. The total number of matings and latency to first copulation constitute the two measured components of males’ fitness. Females were rarely observed to remate (2/507 observed copulations). On the morning of the third day after social groups were established, females were removed from their groups using cold anesthesia and isolated in individual vials containing standard fly food. Females were transferred to new vials with fresh food medium every week until death, and lifespan was recorded as our first
component of female fitness. After a female was removed from a vial, we allowed eggs and larvae to develop and counted all eclosed adult offspring up to 21 days after a female was first introduced to a vial. The minimum duration from egg to adult in *D. melanogaster* is approximately 11 days \(^2\), so we were able to ensure all counted offspring were the female’s progeny and not F2 individuals. By transferring females to a new vial every week before any offspring eclosed, we were able to ensure that all offspring produced were derived from matings that occurred when the female was in the social group. The total lifetime offspring production of each female constituted our second component of female fitness.

**Replication**

Ninety-eight replicate social groups were created and divided amongst the five nutritional environment treatments (1:1 P:C/ 4x Concentration, \(n = 24\); 1:2 P:C/ 4x Concentration, \(n = 18\); 1:4 P:C/ 4x Concentration, \(n = 14\); 1:4 P:C/ 2x Concentration, \(n = 22\); 1:4 P:C/ 1x Concentration, \(n = 20\)). If any flies died or escaped, the social group they were a part of was excluded from being videoed, resulting in the exclusion of almost half of our replicates from network analysis. This was necessary however, as uncontrolled dynamics within the group (e.g. presence of a dead individual affecting living social group members) likely contribute variation to measures of social group structure and cause groups to no longer be truly replicated. For intact social groups, up to two videos were taken (one/day over two days). Forty-three independent social groups had fully tracked videos, 13 of which were videoed on both days, giving us 56 fully tracked videos of social groups to be used for network analysis (1:1 P:C/ 4x Concentration, \(n = 16\); 1:2 P:C/ 4x Concentration, \(n = 12\); 1:4 P:C/ 4x Concentration, \(n = 13\); 1:4 P:C/ 2x Concentration,
All flies from a group were excluded from fitness analyses if any flies in their group died or escaped before we began to quantify their fitness metrics. For analyses of total number of matings and latency to first copulation, 640 males were used, including males that were never observed to mate (number of matings = 0, latency to mate was right-censored at 120 minutes). Of these, 340 males also had measured network positions. For analyses of female lifespan and lifetime offspring production, 354 females were used, excluding females that either died during cold anesthesia while being removed from their social groups or escaped after they were removed from their social groups and into individual vials. Of these, 247 females also had measured network positions.

**Analyses of Social Network Positions**

All analyses were conducted in R version 3.6.2\textsuperscript{73}. For social network positions of instrength, outstrength, and betweenness centrality, we ran Poisson-distributed generalized linear mixed models (GLMMs), as these network positions are counts. For clustering coefficient and eigenvector centrality, we ran linear mixed models (LMMs)\textsuperscript{74–76}. All LMMs and GLMMs were constructed using R package lme4 1.1\textsuperscript{77}. In models for instrength, outstrength, betweenness centrality, and clustering coefficient, fixed effects of sex, nutritional environment, and sex-by-environment interactions; and random effects of genotype and social group ID were included. For models of eigenvector centrality, we included fixed effects of sex and sex-by-environment interactions; and random effects of genotype and individual ID to account for non-independence of multiple measures of the same individuals. Group-level properties (e.g. nutritional environment and social group identity) are not able to meaningfully influence mean measures of eigenvector centrality,
as this is inherently a relative measure. Model fits were assessed using R package DHARMa 0.2.7. Accommodations for zero-inflation were applied to models of betweenness centrality using R package glmmTMB 1.0.0\cite{78}, and accommodations for overdispersion were applied to models for instrength and outstrength using an observation-level random effect\cite{76}. Fixed effect interactions were assessed using Type III Wald $\chi^2$ tests, and non-significant interactions were removed from our models\cite{74}. The resulting models for social network positions constitute our base models, from which further investigations are based. To test for effects of genotype and genotype-by-environment interactions on social network positions, the base models were compared to models excluding genotype and including a genotype-by-environment random slopes interaction, respectively, using likelihood ratio tests\cite{74}. To test for fixed effects of nutritional environment and sex on social network positions, Type III Wald $\chi^2$ tests were employed\cite{74}. Since group-level properties such as nutritional environment cannot influence mean levels of eigenvector centrality, but can influence overall variation, Fligner-Killeen tests were used to test homogeneity of variance in eigenvector centrality across the nutritional environments\cite{79}. Since network data is non-independent, significance of within-group effects (sex, genotype, sex-by-environment interactions, and genotype-by-environment interactions) on social network positions was tested by comparing the observed test statistics to test statistics from 1000 null networks using a one-tailed $t$ test (significance values reported as $P_R$ in text; Supplementary Table 1)\cite{80,81}. Null networks were generated by permuting sex within social groups, for assessing the significance of sex and sex-by-environment interactions; and permuting genotype within each sex within social groups, for assessing the significance of genotype and genotype-
by-environment interactions. To better understand the behavioral processes underlying variation in social network positions, we tested for effects of sex, genotype, nutritional environment, sex-by-environment interactions, and genotype-by-environment interactions on fly locomotor activity (measured as the total distance a fly moved in a video) using LMMs with fixed effects of nutritional environment, sex, and their interaction; and random effects of genotype and social group ID. Upon finding significant effects of sex ($\chi^2 = 15.679, P_R = 0.001$), sex-by-environment interactions ($\chi^2 = 12.615, P_R = 0.003$), and genotype (likelihood ratio = 118.87, $P_R < 0.001$) on activity, we added an activity covariate to our base models for social network positions to clarify that observed patterns are due to social processes and not sex and genotypic differences in activity (Supplementary Table 2). In cases where groups had measures of network structure taken both one day and two days after social groups were established, we tested for individual consistency in network positions using Kendall’s rank correlations. Kendall’s rank correlations were used since our measures of network positions are non-normally distributed. As our five nutritional environment treatments differed in two dimensions of nutrition (protein-to-carbohydrate ratio and caloric concentration), we sought to further understand how variation in nutritional environments affects social network positions. We analyzed the effects of P:C ratio and caloric concentration in separate models using the previously described base models for individual-level network positions, only with P:C ratio and caloric concentration levels substituted for the effect of the nutritional environment. Broad-sense heritability ($H^2$) estimates for social network positions were acquired by estimating the proportion of total variance explained by the random effect of genotype in our base models using R package MuMIn 1.43.1782-84. Current methods of
estimating the proportion of variance explained by random effects are unequipped to handle zero-inflated models, so our base model for betweenness centrality was amended with an observation-level random effect to account for the excess variation attributed to zero-inflation.

**Fitness Analyses**

For number of matings, lifespan, and lifetime offspring production, we ran Poisson-distributed GLMMs, as these fitness components were either counts or best represented by a Poisson distribution\(^74,76\). For latency to first copulation, we ran mixed effect Cox proportional hazards models using the R package coxme 2.2, as our latency data was right-censored\(^85\); nearly half of males (331/640, 48.5\%) did not mate in our initial two hours of observations, meaning their latency to mate was an unknown duration of >2 hours. We confirmed that our data met the assumptions of Cox proportional hazards models by fitting a survival function including a variable indicating whether each measurement was censored or not (all global \(P > 0.05\)). Initial models for each fitness component contained a fixed effect of nutritional environment, and random effects of genotype and social group ID. Model fit for GLMMs was assessed using R package DHARMa 0.2.7. Accommodations for zero-inflation were applied to models for number of matings, and accommodations for overdispersion were applied to models for lifetime offspring production and lifespan by specifying a negative binomial distribution using R package glmmadmb 0.8.3.3\(^86\). These models constitute our base fitness component models from which further analyses are based. We used likelihood ratio tests to test the effect of genotype, and Type III Wald \(\chi^2\) tests to test the effect of nutritional environment in each of our base fitness component models. For models testing the effects of each
social network position and its interaction with the nutritional environment on our fitness components, we only considered individuals that had both a fitness response and measured network phenotypes. Each individual fly had only one possible measurement of fitness per fitness component, but individuals could have two measures of each social network position (in cases where videos of the social group were taken both one and two days after social groups were formed). Because individuals displayed significant consistency in network positions across days, we used network data collected on the first day after social group formation in our fitness analyses, as the first day of network data had a more robust sample size (Supplementary Table 4). Social network positions are often highly correlated, which can create problems when multiple collinear variables are included in multiple regression models. We tested for Kendall’s rank correlations between all pairwise combinations of our five social network positions, and found all to be significantly correlated (Supplementary Fig. 2). To test if the observed multicollinearity would pose problems in multiple regression models, we measured the variance inflation factors of the social network positions by adding all network positions to the base models for each fitness component. The observed multicollinearity was far above amounts generally considered acceptable for multiple regression analyses. As such, we analyzed the effect of each of the five social network positions on each fitness component in a model on its own, and applied a Bonferroni correction for multiple testing to $P$ values from these models. Because our measures of network position are highly multicollinear and non-independent, significance tests will likely be conservative with corrections for multiple testing. Each of the social network positions and its interaction with the nutritional environment were added to the base models for each
fitness component. If the two-way network position-by-environment interaction effect was found to not affect fitness, the interaction was removed from the model. The significance of each social network position and network position-by-environment interaction was assessed using Type III Wald $\chi^2$ tests\textsuperscript{74}. To better understand how variation in nutritional environments affect the strength and direction of selection on social network positions, we analyzed the effects of our two dimensions of nutrition separately by substituting the effects of P:C ratio and caloric concentration for nutritional environment in the models described above (Supplementary Table 3).

**DATA AVAILABILITY**

Source data for all figures and tables are provided with this paper. All other data is publicly available on Zenodo repository (https://doi.org/10.5281/zenodo.4642991)\textsuperscript{89}.

**CODE AVAILABILITY**

Source code for this work is publicly available on Zenodo repository (https://doi.org/10.5281/zenodo.4642991)\textsuperscript{89}.

**REFERENCES**


22. Farine, D. R., Montiglio, P. O. & Spiegel, O. From Individuals to Groups and


correlates of social network variation in plurally breeding degus (Octodon degus).


69. James, R., Croft, D. P. & Krause, J. Potential banana skins in animal social


87. James, G., Witten, D., Hastie, T. & Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*. Springer (2013). doi:10.1016/j.peva.2007.06.006


89. Wice, E. W. & Saltz, J. B. Data from: Selection on heritable social network

ACKNOWLEDGEMENTS

We thank V. H. W. Rudolf, S. P. Egan, F. Oswald and the Saltz laboratory for discussions; E. Dullea, S. Maddox, N. Koonce, Q. Tran, J. Beshai, A. Geiger, and R. Anderson, for data collection assistance; N. Pinter-Wollman, T. W. Wey, and D. N. Fisher for assistance with network and statistical analysis; and SJ Art and Frames for donating supplies. This research was supported by the National Science Foundation (IOS-1856577 to J. B. S.), a Society for the Study of Evolution Rosemary Grant award to E. W. W., and a Houston Livestock Show and Rodeo award to E. W. W.

AUTHOR CONTRIBUTIONS

E. W. W. designed, performed, and analyzed most of the research. J. B. S. assisted with experimental design and analysis, and supervised the research. E. W. W. wrote the manuscript with assistance from J. B. S.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary tables and figures are available for this paper. Correspondence and requests for materials should be addressed to E. W. W. or J. B. S.

SUPPLEMENTARY FIGURES AND TABLES
Supplementary Figure 1 | Experimental schematic. Solid-colored flies represent the 40 DGRP inbred lines that were uniquely crossed to create 20 heterozygous genotypes (bi-colored flies). The 20 heterozygous genotypes were combined into a social group, which was replicated 98 times. Each social group was placed on one of five nutritional environments. The motion-tracking and social network data presented are representative of data gleaned from a single social group.
**Supplementary Figure 2 | Correlations between network position variables.** Diagonal shows the distribution of each of the five network variables. Lower-left scatterplots show plotted relationships between all pairwise combinations of network variables. Upper-right quadrants show Kendall’s correlation estimates for each pairwise combination of network variables.
### Supplementary Table 1 | Model results for effects on network positions

Both observed p-values (P), and p-values gleaned from network permutation tests (PR) are presented. Bolded PR values indicate significant effects (* indicates marginal significance).
### Supplementary Table 2 | Model results for effects on network positions with activity covariate.

Both observed p-values ($P$), and p-values gleaned from network permutation tests ($P_R$) are presented. Bolded $P_R$ values indicate significant effects (* indicates marginal significance).

<table>
<thead>
<tr>
<th>Social Network Positions</th>
<th>Sex</th>
<th>Genotype</th>
<th>Nutritional Environment</th>
<th>Sex-by-Environment</th>
<th>Genotype-by-Environment</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\chi^2$</td>
<td>Df</td>
<td>$P$</td>
<td>$P_R$</td>
<td>LRT</td>
<td>Df</td>
</tr>
<tr>
<td>In-strength</td>
<td>4.816</td>
<td>1</td>
<td>0.028</td>
<td><strong>0.015</strong></td>
<td>25.339</td>
<td>1</td>
</tr>
<tr>
<td>Out-strength</td>
<td>11.948</td>
<td>1</td>
<td>&lt;0.001</td>
<td><strong>0.001</strong></td>
<td>26.814</td>
<td>1</td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>114.125</td>
<td>1</td>
<td>&lt;0.001</td>
<td><strong>&lt;0.001</strong></td>
<td>57.175</td>
<td>1</td>
</tr>
<tr>
<td>Betweenness Centrality</td>
<td>0.854</td>
<td>1</td>
<td>0.419</td>
<td><strong>0.888</strong></td>
<td>984.26</td>
<td>1</td>
</tr>
<tr>
<td>Eigenvector Centrality</td>
<td>8.796</td>
<td>1</td>
<td>0.003</td>
<td><strong>0.002</strong></td>
<td>24.236</td>
<td>1</td>
</tr>
<tr>
<td>Fitness Effects</td>
<td>Number of Matings</td>
<td>Latency to Mating</td>
<td>Lifetime Offspring Production</td>
<td>Lifespan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-------------------------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Test Statistic</td>
<td>Df</td>
<td>P</td>
<td>Test Statistic</td>
<td>Df</td>
<td>P</td>
</tr>
<tr>
<td>Genotype</td>
<td>99.76 1 &lt;0.001</td>
<td>87.104 1 &lt;0.001</td>
<td>0.04 1 0.842</td>
<td>0.16 1 0.689</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nutritional Environment</td>
<td>1.021 4 0.907</td>
<td>2.416 4 0.659</td>
<td>6.84 4 0.145</td>
<td>27.277 4 &lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC Ratio</td>
<td>0.273 1 0.901</td>
<td>1.190 1 0.294</td>
<td>4.860 1 0.028</td>
<td>7.413 1 0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caloric Concentration</td>
<td>0.113 1 0.737</td>
<td>0.020 1 0.886</td>
<td>&lt;0.001 1 0.989</td>
<td>10.815 1 0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insstrength</td>
<td>8.795 1 0.003</td>
<td>9.687 1 0.002</td>
<td>0.857 1 0.355</td>
<td>0.607 1 0.936</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outstrength</td>
<td>4.075 1 0.044*</td>
<td>3.070 1 0.080*</td>
<td>0.570 1 0.456</td>
<td>0.615 1 0.903</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coefficient</td>
<td>2.410 1 0.121</td>
<td>3.814 1 0.051*</td>
<td>0.717 1 0.397</td>
<td>0.000 1 0.996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Betweenness Centrality</td>
<td>2.060 1 0.149</td>
<td>0.555 1 0.476</td>
<td>1.593 1 0.207</td>
<td>2.138 1 0.144</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eigenvector Centrality</td>
<td>9.908 1 0.002</td>
<td>0.009 1 0.925</td>
<td>1.406 1 0.439</td>
<td>1.150 1 0.284</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insstrength-by-Environment</td>
<td>52.014 4 &lt;0.001</td>
<td>13.720 4 0.008</td>
<td>7.074 4 0.132</td>
<td>1.017 4 0.907</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insstrength-by-PC Ratio</td>
<td>6.478 1 0.011**</td>
<td>7.317 1 0.007</td>
<td>4.130 1 0.042*</td>
<td>0.816 1 0.366</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outstrength-by-PC Ratio</td>
<td>7.058 1 0.008</td>
<td>10.367 1 0.001</td>
<td>2.002 1 0.157</td>
<td>0.345 1 0.557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outstrength-by-Caloric Concentration</td>
<td>20.440 4 &lt;0.001</td>
<td>10.602 4 0.031*</td>
<td>4.130 4 0.389</td>
<td>2.184 4 0.702</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outstrength-by-PC Ratio</td>
<td>5.733 1 0.017**</td>
<td>8.183 1 0.004</td>
<td>2.341 1 0.126</td>
<td>1.540 1 0.215</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outstrength-by-Caloric Concentration</td>
<td>2.659 1 0.100</td>
<td>6.471 1 0.011**</td>
<td>1.311 1 0.252</td>
<td>0.953 1 0.356</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coefficient-by-Environment</td>
<td>5.975 4 0.299</td>
<td>4.129 4 0.399</td>
<td>0.983 4 0.907</td>
<td>1.220 4 0.876</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering Coefficient-by-PC Ratio</td>
<td>0.019 1 0.899</td>
<td>0.057 1 0.811</td>
<td>0.502 1 0.479</td>
<td>0.402 1 0.526</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Betweenness Centrality-by-PC Ratio</td>
<td>1.676 1 0.196</td>
<td>2.293 1 0.130</td>
<td>&lt;0.001 1 0.993</td>
<td>0.813 1 0.367</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Betweenness Centrality-by-Caloric Concentration</td>
<td>7.473 4 0.112</td>
<td>5.857 4 0.210</td>
<td>2.943 4 0.567</td>
<td>5.371 4 0.251</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Betweenness Centrality-by-PC Ratio</td>
<td>4.177 1 0.146</td>
<td>3.509 1 0.530</td>
<td>1.359 1 0.244</td>
<td>3.471 1 0.063*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Betweenness Centrality-by-Caloric Concentration</td>
<td>5.016 4 0.043</td>
<td>2.910 4 0.068*</td>
<td>0.799 1 0.371</td>
<td>0.920 1 0.338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eigenvector Centrality-by-Environment</td>
<td>49.699 4 &lt;0.001</td>
<td>12.803 4 0.012**</td>
<td>2.269 4 0.687</td>
<td>0.394 4 0.983</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eigenvector Centrality-by-PC Ratio</td>
<td>6.678 1 &lt;0.010</td>
<td>4.330 1 0.037*</td>
<td>0.732 1 0.392</td>
<td>0.023 1 0.880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eigenvector Centrality-by-Caloric Concentration</td>
<td>9.679 1 0.002</td>
<td>10.928 1 0.001</td>
<td>1.052 1 0.305</td>
<td>0.198 1 0.664</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Table 3** | Model results for effects on fitness components. The significance threshold for models including network data was Bonferroni adjusted to account for multiple testing of the five network position variables. Uncorrected p-values are reported. Bolded p-values indicate significant effects, and p-values marked with * indicate marginally significant effects, after adjusting for multiple testing if applicable. P-values marked with ^ or + indicate significant and marginally significant effects, respectively, without multiple testing correction.
<table>
<thead>
<tr>
<th>Nutritional Environments</th>
<th>Number of Social Groups with Networks from Only Day 1</th>
<th>Number of Social Groups with Networks from Only Day 2</th>
<th>Number of Social Groups with Networks from Both Days 1 &amp; 2</th>
<th>Sum by Nutritional Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1 P:C / 4x Concentration</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>1:2 P:C / 4x Concentration</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>1:4 P:C / 4x Concentration</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>1:4 P:C / 2x Concentration</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>1:4 P:C / 1x Concentration</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Sum by Day</td>
<td>22</td>
<td>8</td>
<td>13</td>
<td>Total Networks 56</td>
</tr>
</tbody>
</table>

**Supplementary Table 4 | Network sample size summary.** This table summarizes the number of networks analyzed by each nutritional environment (rows), as well as whether the networks were generated from videos of social groups taken 1, 2, or both days after social groups were established (columns). Forty-three independent social groups had their network structure analyzed, 13 of which had replicate measures of network structure taken across both days of videoing.
Chapter 2: Indirect genetic effects for social network structure in Drosophila melanogaster

Eric Wesley Wice\textsuperscript{1} and Julia Barbara Saltz\textsuperscript{1}

\textsuperscript{1}Department of Biosciences, Rice University, Houston, TX 77005

*Corresponding author (eric.wesley.wice@gmail.com)

ABSTRACT

The position an individual holds in a social network is dependent on both its direct and indirect social interactions. Because social network position is dependent on the actions and interactions of conspecifics, it is likely that the genotypic composition of individuals within a social group impacts individuals’ network positions. However, we know very little about whether social network positions have a genetic basis, and even less about how the genotypic makeup of a social group impacts network positions and structure. With ample evidence indicating network positions influence various fitness metrics, studying how direct and indirect genetic effects shape network positions is crucial for furthering our understanding of how the social environment can respond to selection and evolve. Using replicate genotypes of Drosophila melanogaster flies, inbred from a natural population, we created social groups that varied in their genotypic makeup. Social groups were videoed, and networks were generated using motion-tracking software. We found that both an individual’s own genotype and the genotypes of conspecifics in a social group affect its position within a social network. These findings provide an early example of how indirect genetic effects and social network theory can be linked, and shed new light on how quantitative genetic variation shapes the structure of social groups.
INTRODUCTION

The social interactions individuals engage in, and the emergent structure of the social groups they are a part of, is one of the most consequential components of individuals’ lives (Krause et al. 2007; Croft et al. 2008). One of the many factors contributing to social interactions and emergent social group structure is the genotypes of individuals comprising a group (Fisher and McAdam 2017; Montiglio et al. 2018; Bentzur et al. 2020). From an evolutionary perspective, the genetic underpinnings of social interactions are particularly interesting: when the genotype of one individual affects the phenotype of an interacting partner (def. indirect genetic effect, or IGE), this genetic component of the social environment allows the social environment itself to respond to selective pressures and evolve (Wolf et al. 1998, 1999; Bijma et al. 2007).

Theoretical and empirical work incorporating IGEs has seeded a better understanding how variation in social interactions and dynamics takes shape (Bleakley and Brodie III 2009; Chenoweth et al. 2010; Dingemanse and Araya-Ajoy 2015; Bailey et al. 2018; Araya-Ajoy et al. 2020). However, models of IGEs have struggled to move beyond incorporating dyadic social interactions, to encompassing the genetic effects of multiple direct and indirect interacting phenotypes simultaneously (Saltz 2013; Schneider et al. 2016; Fisher and McAdam 2017; Montiglio et al. 2018; Araya-Ajoy et al. 2020). This limitation is problematic for applying IGE theory to data from wild study systems, because individuals frequently engage in social interactions with multiple conspecifics.

One of the ways IGE theory has been extended to apply to social groups is that the effect of multiple individuals on a focal individual is averaged across all social groupmates (Montiglio et al. 2018). An unintended consequence of this approach is that, in a social
group of infinite size, no single individual can have an effect on anyone else (Montiglio et al. 2018). This is problematic because we know that how individuals affect others is not simply an average. Previous studies have demonstrated how singular individuals can have an outsize influence on their social groups (see ‘keystone individuals’, Modlmeier et al. 2014) and networks (e.g. Flack et al. 2006). Furthermore, how interacting individuals affect others can depend on the specific individuals they are paired with (Wade 1998; Wolf and Brodie III 1998; Wolf 2000; Kraft et al. 2016; Signor et al. 2017; Baud et al. 2018; Jaffe et al. 2020; Lane et al. 2020). If the effect of interacting individuals depends on both their own genotype and the genotype of the individuals they are interacting with, then these effects can be considered a type of genotype-by-environment interaction, where the ‘environment’ is the indirect genetic effects imposed by the social environment (Saltz 2011; Kraft et al. 2016; Signor et al. 2017; Baud et al. 2018; Jaffe et al. 2020; Rooke et al. 2020). However, we know very little about how variability in IGEs and genotype-by-IGE interactions are predicted to shape variation in group structure, as there is currently no systematic way of predicting how genotypes will vary with given environments, especially social environments (D’Aguillo et al. 2019).

In contrast, social network analysis has emerged as a widely used tool for describing how individuals’ direct and indirect social interactions are nested within the emergent social structure of their group (Proulx et al. 2005; Krause et al. 2007; Croft et al. 2008; Wey et al. 2008; Sih et al. 2009; Webber and Vander Wal 2019). Network analysis provides a unique opportunity to move beyond studying IGEs in dyadic contexts, as it allows us to address how the genetic composition of groups and the effects of social group members cascade through networks of direct and indirect interactions (Fisher and
Extending our knowledge of how the genetic effects of social group members affect the structure and dynamics of networks of social interactions - i.e., IGEs within social networks - can allow us to better understand how variation in social structures form and evolve (Fisher and McAdam 2017; Montiglio et al. 2018).

Numerous studies have addressed how behavioral variation shapes the positions individuals hold in their networks of social interactions: movement and exploratory behavior (Aplin et al. 2013; Jezovit et al. 2020), aggression and policing behaviors (Flack et al. 2006; Lea et al. 2010; Kilgour et al. 2020), and group size preference and experience (Firth et al. 2017; Bentzur et al. 2020; Rooke et al. 2020). However, it remains unclear whether these behavioral effects on network structure have an underlying quantitative genetic basis. Understanding how an individual’s genotype affects not only its own network position, but the network positions of conspecifics in the social group, is key to understanding how genetic variation manifests into variation in network structure (Fisher and McAdam 2017; Montiglio et al. 2018; Bentzur et al. 2020; Radersma 2020).

A handful of studies have addressed how the genotypic makeup of a social group impacts various aspects of its structure and dynamics: social niche construction (Saltz 2011, 2017; Saltz and Foley 2011; Geiger and Saltz 2020), exploratory behavior (Jaffe et al. 2020), collective foraging (Walsh et al. 2020), antipredator behavior (Bleakley and Brodie III 2009), and aggression (Saltz 2013; Kilgour et al. 2020). Yet it remains unclear how the genotypes of multiple individuals within social groups impacts the structure of their networks of social interactions.
Because the position an individual holds in a network is inherently dependent on the direct and indirect social interactions occurring amongst all social group members, the effects of individuals’ own genotypes on their own network position are difficult to parse from the indirect genetic effects of other social group partners (Radersma 2020). Perhaps as a consequence of these empirical challenges, few studies have addressed how an individual’s own genotype influences its position within a social network (Fowler et al. 2009; Lea et al. 2010; Brent et al. 2013), and we know even less about how the genotypic composition of social groups and indirect genetic effects of interacting group members shape variation in network phenotypes (Montiglio et al. 2018; Radersma 2020). Merging studies of the quantitative genetic basis of social traits, IGEs, and social networks can open new windows into understanding the genetic basis and evolutionary potential of social group structures.

In this study, we started to bridge this gap in knowledge by examining how multiple individuals’ genotypes and the genotypes of their social partners affected multiple measures of social network positions using replicate genotypes of *Drosophila melanogaster* flies. Flies are a great study system to address questions about the quantitative genetic basis of social traits, as we can replicate the genotypes of individuals engaged in social interactions using multiple inbred lines (Mackay et al. 2012). Additionally, flies have been shown to vary in social group preference (Saltz 2011, 2017), actively choose the social groups they are apart of (Saltz and Foley 2011; Geiger and Saltz 2020), and have non-random social networks of interactions (Schneider et al. 2012; Pasquaretta et al. 2016). In general, we expect an individual’s own genotype and the genotypes of its social groupmates to affect its position within a social network. By
manipulating the genotypic composition of social groups, our experimental design allowed us to quantify the role of direct genetic effects (DGEs) and IGEs on social network structure.

METHODS

Genotypes

Heterozygous genotypes were created by performing mating crosses of inbred homozygous genotypes of *Drosophila melanogaster* flies derived from the *Drosophila* Genetic Reference Panel (DGRP), a collection of dozens of inbred homozygous genotypes derived from a natural population in North Carolina (Mackay et al. 2012; Figure 1). Each genotype was generated by establishing a mating cross of 10 virgin females of homozygous genotypes 208, 315, 786, or 637 with 10 males of homozygous genotypes 313, 229, 716, or 318, respectively. Virgin females were collected from mating crosses 208x313, 315x229, and 786x716; and virgin males were collected from mating cross 637x318. Genotype numbers refer to arbitrary labeling within the DGRP and are unindicative of similarities or differences between genotypes (Mackay et al. 2012). Seeding the vials with a standard number of individuals per vial (10 females and 10 males), allowed us to minimize variation in larval density. We chose to use heterozygous genotypes, as opposed to homozygous flies from inbred lines, to create individuals that are more representative of genetic variation found in the wild. Additionally, creating heterozygous genotypes allays concerns about potentially deleterious effects of homozygous recessive alleles in inbred homozygous lines (Wahlsten 2001; Brakefield 2003).
The three heterozygous female genotypes were chosen using *a priori* information about genotypic differences in network position from Chapter 1. Specifically, we chose three genotypes of female flies that were observed to differ in the network position of eigenvector centrality (i.e. how critical an individual is to the structure of the network based on the strength of its direct and indirect social connections) (Croft et al. 2008; Wey et al. 2008; Webber and Vander Wal 2019). We chose female genotypes known to vary in eigenvector centrality for two reasons: First, this measure of network position incorporates information about both an individual’s direct and indirect social interactions, which is a key part of addressing our question of how an individual’s own genotype, and the genotypes of its direct and indirect social partners, affects its network position. Second, eigenvector centrality is an inherently relevant measure, meaning how central an individual is to the structure of its group inherently depends on its social groupmates (i.e. if one or more individuals are more central to group structure, this means other individuals in the group have to be less central to group structure). We chose to manipulate the genotypic composition and ratios of only the females in the group, as our results from Chapter 1 indicate that females receive more social interactions and are more central to the structure of their social groups. More importantly though, manipulating both sexes could introduce confounding variation about which specific genotypic and sex combinations affect group structure (Saltz and Foley 2011; Saltz 2013); and this was not the goal of our current study.

**Social Groups**

Social groups varied in both their genotypic composition, as well as the ratio of genotypes within them. Each social group contained two different heterozygous female
genotypes, in a full-factorial design, in a 1:9 ratio (Figure 1). Each combination of females was paired with 10 males of a standardized heterozygous genotype (Figure 1). Thus, each social group contained 20 individuals with a 50/50 sex ratio.

After 2-3 weeks of development, virgin flies were collected under light CO$_2$ anesthesia. Female flies were split between their specific treatment combinations (1:9 genotypic ratios of all pairwise combinations of two different female genotypes; Figure 1), and flies from both sexes were randomly marked with a unique paint color on their mesothoracic segment to visually distinguish individuals. Each group of 10 treatment females and 10 standardized males were then aged in same-sex groups for three days to allow for recovery from CO$_2$ anesthesia and development to sexual maturation (Barron 2000; Bartholomew et al. 2015). After three days, the treatment ratios and compositions of female genotypes were combined with the standardized males into a 10 cm petri dish layered with fly food (58.8g nutritional yeast, 133.7g malt sugar, 27g agar, 11.1mL tegocept acid mix [70g tegocept/270mL H$_2$O], and 3mL propionic acid; per 1L H$_2$O). All flies were reared, aged, and housed on a 12:12 light:dark cycle, at 24°C and 50% relative humidity, and on standard fly food unless otherwise noted.
Figure 1: Overview of the experimental design indicating how heterozygous genotypes were derived from homozygous inbred lines of flies, and how social groups of variable composition and ratios of female genotypes were established. Genotype numbers refer to arbitrary indicators from the DGRP (Mackay et al. 2012). Sample sizes indicate the number of networks analyzed for each social group treatment.
Fly Behavior and Social Network Analysis

See Chapter 1 for details of how social interactions and network structure were quantified. Also see Supplementary Figure 1 for definitions and depictions of the social network positions analyzed.

Replication

One hundred and twelve social groups were created, split amongst the six treatments of two of three female genotypes in a 1:9 ratio in a full-factorial design: 1:9 ratio of genotypes 208x313 and 786x716, respectively (n = 19); 1:9 ratio of 208x313 and 315x229 (n = 19); 1:9 ratio of 786x716 and 315x229 (n = 19); 9:1 ratio of 208x313 and 786x716 (n = 19); 9:1 ratio of 208x313 and 315x229 (n = 19); 9:1 ratio of 786x716 and 315x229 (n = 18). Social groups were excluded from analyses if any flies died or escaped before they were videoed. This excluded over half of our social groups from analyses, as keeping all 20 flies in a group alive throughout the duration of the experiment was challenging. However, doing so was necessary, because variation introduced by the presence of a dead individual and subsequent changes in group size cause groups to no longer be replicated within each genotype ratio/composition treatment. For intact social groups, up to two videos were taken (one/day over the course of two days). Forty-four independent social groups had fully-tracked videos, 21 of which were videoed on both days, resulting in 65 tracked videos of social groups used for network analysis (1:9 ratio of genotypes 208x313 and 786x716, respectively (n = 7); 1:9 ratio of 208x313 and 315x229 (n = 14); 1:9 ratio of 786x716 and 315x229 (n = 14); 9:1 ratio of 208x313 and 786x716 (n = 15); 9:1 ratio of 208x313 and 315x229 (n = 4); 9:1 ratio of 786x716 and 315x229 (n = 11; Figure 1; Supplementary Table 1).
**Analyses of DGEs and IGEs on Network Structure**

We subsetted our data prior to analyses in two ways: First, because all males had the same standardized genotype, we could not address how a male individual’s genotype (DGE) affected their own network position. Thus, males were analyzed separately from females, in which we could measure DGEs on network position. Secondly, because a female individual’s own genotype covaried with which group of female genotypes the social group contained in 90% of the data (in cases where an individual was one of 9 genotypes in the whole group) and 10% of the remaining data (in cases where an individual was a singular focal genotype in the group), this presented issues with convergence in many of our models. Thus, we analyzed focal females (the unique female individual in each group) separately from the nine female genotypes that were consistent within a social group (hereafter referred to as the “stimulus genotype”). This also allowed us to address not only how the genotype(s) an individual’s social group partners affect its network position, but how the genotype of a single focal individual could affect the structure of the entire group.

**Model Specifications**

All analyses were conducted in R version 3.6.2 (Team 2019). For focal and stimulus genotype females, we analyzed how an individual’s own genotype (DGEs) affected their network positions. For focal and stimulus genotype females, and males, we analyzed how the genotype of their social partner(s) (IGEs) affected network positions. Fixed factors included in models for females included an interaction between an individual’s own genotype and the genotype of the female(s) they were paired with, as well as a random effect of the social group individuals were a part of. Fixed factors included in models for
males included an interaction between the identity of the focal female genotype and stimulus female genotype, and a random effect of the social group males were a part of. Genotype was treated as a fixed effect because genotypes were chosen non-randomly based on prior information (see ‘Genotypes’ section), and because there were only three levels of genotype for females.

One of the behavioral factors that can underly an individual’s network position, is how much it physically moves within its social group. To clarify that our results for how DGEs and IGEs affect network position were due to differences in social interactions and not simply physical movement, we analyzed how DGEs and IGEs affect fly activity (measured as the total distance a fly moved throughout a video) using the same analysis structure as our models investigating network positions. Upon finding evidence of DGEs for activity in females and IGEs for activity in males, we incorporated an activity covariate in our previous models investigating the effects on network positions, and we report results from these models.

**Model Fitting**

Network positions of instrength, outstrength, and betweenness centrality were analyzed using Poisson-distributed generalized linear mixed models (GLMMs), as these measures of network position are counts; and network positions of clustering coefficient and eigenvector centrality were analyzed using linear mixed models (LMMs) using R package lme4 1.1 (Bates et al. 2015). Model fits were assessed using the package DHARMa 0.2.7 (Hartig 2020). Accommodations for overdispersion were applied as needed using an observation-level random effect in models for instrength and outstrength,
and for zero-inflation by specifying a negative binomial distribution as needed in models for betweenness centrality (Harrison et al. 2018).

**Inference**

The significance of higher-order interactions among fixed effects in our models was assessed using Type III Wald $\chi^2$ tests, and non-significant interactions were broken down into their subsequent main effects (Bolker et al. 2009). The significance of fixed effects were assessed using Type III Wald $\chi^2$ tests.

Measures of network position are non-independent, meaning one individual’s network position inherently informs other individuals’ network position (Croft et al. 2011; Farine 2017). Because of this, permutation tests are often required to generate an appropriate null distribution to compare observed results to, where the identities of individuals within each group are randomized (Farine 2017; Weiss et al. 2021). However, because we subsetted our analyses of females by whether females were focal individuals (in cases where an individual was one of 9 genotypes in the whole group) or stimulus individuals (one of 9 social groupmates of the same genotype), these null model randomizations are unneeded, because the genotypic identities of all individuals within a group within any given analysis are identical.

**RESULTS**

**DGEs and Activity Effects on Network Positions**

Female individuals’ own genotype significantly correlated with three measures of network position that describe aspects of both an individual’s direct and indirect social interactions: clustering coefficient (type III Wald $\chi^2$ tests: stimulus genotypes, $\chi^2 = 12.525, P = 0.002$), betweenness centrality (stimulus genotypes, $\chi^2 = 12.828, P = 0.002$;
focal genotypes, $\chi^2 = 9.481, P = 0.009$), and eigenvector centrality (stimulus genotypes, $\chi^2 = 13.652, P = 0.001$; focal genotypes, $\chi^2 = 6.914, P = 0.032$; Figure 2). While an individual’s own genotype was found to affect measures of network position encompassing both direct and indirect social interactions, we found no evidence that an individual’s own genotype influenced measures of network position that only encompass direct social interactions (instrength and outstrength; all $P > 0.05$; Figure 2).
Figure 2: The effects of a female’s own genotype on its network position, broken down by whether the female was a focal individual in its group (A,C,E,G,I; its genotype was unique amongst the social group) or the female was one of nine individuals of the same stimulus genotype (B,D,F,H,J).

Genotype labels indicate the maternal parent genotype crossed with ("x") the paternal genotype to create replicate heterozygous genotypes (Mackay et al. 2012). Note that genotype numbers are arbitrary and do not describe similarities or differences between genotypes. Boxplots indicate the median, interquartile range (IQR), values within ±1.5x IQR, and outliers. Significance of genotypic differences are indicated by asterisks (* <0.05, ** <0.01, *** <0.001).
In addition to genotypic differences in network positions, we also detected
genotypic differences in activity (focal genotypes, $\chi^2 = 11.640, P = 0.003$). Further, we
found an effect of activity on all measures of network position. More specifically, focal
females who were more active engaged in fewer social interactions (instrength, $\chi^2 =
3233.944, P < 0.001$; and outstrength, $\chi^2 = 4169.524, P < 0.001$) and were less of a
bridge between indirectly connected individuals (betweenness centrality, $\chi^2 = 11.139, P <
0.001$). For stimulus females, being more active was also negatively associated with
being a bridge between indirectly connected individuals (betweenness centrality, $\chi^2 =
115.078, P < 0.001$). Additionally, more active stimulus females were less cliquish
(clustering coefficient, $\chi^2 = 5.565, P = 0.018$) and more central to group structure
(eigenvector centrality, $\chi^2 = 17.488, P < 0.001$). Males exhibited a similar pattern, where
being more active was associated with being less of a bridge between indirectly
connected individuals (betweenness centrality, $\chi^2 = 51.971, P < 0.001$). However, how
activity affected the number of social interactions males initiated (outstrength) differed
from focal females, where being more active was associated with higher outstrength ($\chi^2 =
4.664, P = 0.031$).

**IGEs on Activity and Network Positions**

How the genotype of female’s social partner(s) affected their own network position
depended on whether the female was a focal or stimulus individual. We found that the
genotype of a single focal female genotype had no influence on the network positions of
the stimulus female genotypes in a group (all $P > 0.05$; Figure 3). However, we did find
that the stimulus female genotype that focal females and males were paired with could
influence their network positions. The stimulus genotype influenced how cliquish focal
females were (clustering coefficient, $\chi^2 = 7.028, P = 0.030$) and how focal females served as a bridge between indirectly connected individuals (betweenness centrality, $\chi^2 = 7.107, P = 0.029$; Figure 3). Additionally, the stimulus genotype of females also influenced how central males were to the structure of their social group (eigenvector centrality, $\chi^2 = 7.862, P = 0.020$). We did not find evidence to support the hypothesis that the stimulus genotype of females influenced our other tested measures of network position for males and focal females (all $P > 0.05$; Figure 3). We also found only marginal evidence for interactions between a focal female’s own genotype and the genotype of its social groupmates on how many interactions a focal female engaged in (instrength, $\chi^2 = 3.210, P = 0.073$; outstrength, $\chi^2 = 3.762, P = 0.052$) and how central focal females were to the structure of their social group (eigenvector centrality, $\chi^2 = 3.047, P = 0.081$). And we found no evidence for interactions between a stimulus female’s own genotype and the genotype of the focal female genotype they were paired with (all $P > 0.05$). Similarly, we found no evidence of an interaction between the genotypes of the focal and stimulus females that males were paired with (all $P > 0.05$).

We found only one measure of males’ network positions was affected by the female genotype(s) they were paired with. Specifically, how central males were to the structure of their social group (eigenvector centrality) varied depending on the stimulus female genotype they were paired with ($\chi^2 = 7.862, P = 0.020$). In contrast, we did find evidence that males’ overall activity was significantly affected by both the stimulus female genotype ($\chi^2 = 8.057, P = 0.018$) and the focal female genotype ($\chi^2 = 7.383, P = 0.025$). Neither focal nor stimulus females’ activities were affected by the stimulus or focal female genotype, respectively, they were paired with (all $P > 0.05$).
Figure 3: The effects of females' social group partners on their network positions, broken down by the effects of singular focal genotypes on nine stimulus genotypes (A,C,E,G,I) and the effects of nine stimulus genotypes on focal individuals (B,D,F,H,J). Genotype labels indicate the maternal parent genotype crossed with ("x") the paternal genotype to create replicate heterozygous genotypes (Mackay et al. 2012). Note that genotype numbers are arbitrary and do not describe similarities or differences between genotypes. Boxplots indicate the median, interquartile range (IQR), values within ±1.5x IQR, and outliers. Significance of genotypic differences are indicated by asterisks (*<0.05, **<0.01, ***<0.001).
DISCUSSION

Taken together, our findings show that an individual’s position within a social network is dictated not only by its own genotype, but also by indirect genetic effects of other individuals within the network. Such results indicate that studies of the genetic basis of network position should extend to include the genotypes of all social group members. To not incorporate information about indirect genetic effects in studies of network structure risks missing crucial components of quantitative genetic variation in the structure of social groups.

The idea that an individual’s position within a social network is dependent not only on its own genotype, but by the genotypes of other individuals within a network, is perhaps intuitive, as an individual’s position within a network is dependent on both its own direct social interactions and indirect interactions amongst groupmates as well (Krause et al. 2007; Croft et al. 2008; Wey et al. 2008; Sih et al. 2009). However, studying how DGEs and IGEs affect network structure has remained challenging (Fisher and McAdam 2017; Montiglio et al. 2018; Bentzur et al. 2020; Radersma 2020). Previous studies investigating the heritability of measures of network position have estimated the direct genetic contributions to network position phenotypes, but have done so using individuals of known genotype or relatedness nested within social networks of individuals of unknown genotypes or relatedness (Fowler et al. 2009; Lea et al. 2010; Brent et al. 2013). Because of the challenges of controlling for the genetic identities of all individuals within a social group, it has remained unclear to what degree an individual’s own genotype and the genotypes of their social groupmates affect each individual’s position within a social network. For example, even if focal individuals of known
relatedness or genotype were to behave and socially interact identically across different social groups, how their social partners interact with indirectly connected individuals can have profound effects on those focal individuals’ network positions. In other words, apparent DGEs may be driven wholly by IGEs. This is perhaps one of many reason why prior estimates of the genetic contributions to measures of network position have had variable results, whereby not controlling for the genetic effects of all social group members could have increased variation in measures of network position and obscured results (Fowler et al. 2009; Lea et al. 2010; Brent et al. 2013; Radersma 2020). Because of this, it will be necessary to investigate IGEs in wild social groups within a broader range of species in order to come to generalizable conclusions about the quantitative genetic basis of social group structure.

It is interesting to note that the measures of network position we found to be affected by DGEs and IGEs (eigenvector centrality, betweenness centrality, and clustering coefficient) incorporate information about an individual’s indirect interactions in addition to direct interactions, while measures of network position that only encompass an individual’s direct social interactions (instrength and outstrength) were not affected by DGEs or IGEs. These findings are perhaps surprising, as we might expect an individual to have more control over its direct social interactions compared to indirect interactions, and for DGEs to have a stronger effect on these measures of direct social interactions. One potential factor influencing this finding is that all females were paired with males of a standardized genotype that did not vary. Male flies initiated more interactions than female flies (outstrength, $\chi^2 = 3.854$, $P < 0.050$), though it’s impossible to disentangle the effect of sex from genotype in this study, as all males had the same genotype. However,
males initiated similar amounts of interactions regardless of which female genotypes they were paired with (all $P > 0.050$). With males initiating more social interactions, and not discriminating amongst the genotypes of female flies they were paired with, this could potentially explain why we found no evidence of DGEs for measures of direct social interactions for females.

Additionally, prior studies have similarly found that social interactions that are received, in addition to measures of network position that incorporate indirect interactions, can have stronger direct genotypic effects than measures of initiated interactions (Fowler et al. 2009; Lea et al. 2010; Brent et al. 2013). Similarly, studies of gene-environment correlations (def. when an individual’s genotype correlates with the environment it experiences) have shown that an individual’s own genotype can evoke a heritable response from social partners (e.g. parenting behaviors in studies of twins reared apart) (Plomin et al. 1977; Rutter et al. 2006; Jaffee and Price 2007; Avinun and Knafo 2014). In the context of social groups, gene-environment correlations can also dictate the social environment individuals choose to experience (Scarr and Mccartney 1983; Stamps and Groothuis 2010; Saltz 2011). In social settings in which all individuals are attempting to join a group that matches their preferred social environment, individuals’ preferences can come into conflict with each other, resulting in some individuals being unable to realize their preferred social environment (Saltz and Foley 2011; Saltz 2017). This variation can further obscure how DGEs and IGEs can manifest into differences in how individuals are nested within their broader social environment.

One of the limitations of our study is that we were only able to manipulate three genotypes in creating groups of variable genotypic composition. While social groups
often encompass individuals of many more variable genotypes, we note that this biological reality is a serious impediment to our understanding of how DGEs and IGEs affect variation in social structure; manipulating multiple genotypes in a factorial design with sufficient replication is exceedingly difficult, even in organisms amenable to laboratory manipulation such as flies. Additionally, many of the individuals in our social groups (all males, and 9/10 females in each group) were genotypically identical, which is not typically representative of biological reality. Exceptions to this include haplodiploid social hymenoptera (e.g. ants, bees, and wasps), which alternatively, are classic examples of how relatedness affects social group dynamics (Crozier and Pamilo 1996). While *Drosophila melanogaster* flies may not experience genetically identical individuals in nature, the aim of our study was less about creating social groups that are representative of variation found in the wild, and more about taking an initial step toward integrating IGEs and social network theory.

An additional limitation of our study is that our approach of studying social network positions does not account for the potential functionality of network position, meaning it remains unclear what consequences different network positions have in this specific system. However, our findings of DGEs for measures of network position does translate well to findings of individual differences in other complex social behaviors (i.e. social roles, social responsiveness, social complexity, and social niche specialization; Bergmüller and Taborsky 2010; Montiglio et al. 2013; Dingemanse and Araya-Ajoy 2015; Kappeler et al. 2019). Individuals have been shown to consistently vary in network position across both time (Aplin et al. 2015; Formica et al. 2017; Blaszczyk 2018; Kulahci et al. 2018; Smith et al. 2018) and contexts (Wilson et al. 2013, 2015; Jacoby et
al. 2014; Krause et al. 2017; Muller et al. 2018), with individuals’ behaviors often being tied to their positions within a social group (Flack et al. 2006; Pike et al. 2008; Aplin et al. 2013; Firth et al. 2017; Jolles et al. 2017; Morimoto et al. 2017; Jezovit et al. 2020; Rooke et al. 2020). Correspondingly, individuals have also been shown to plastically alter their network positions in response to prior experience (Pike et al. 2008; Aplin et al. 2013; Jolles et al. 2017; Gaffney and Webster 2018), behavioral changes within their social group (Muller et al. 2018), and perturbations such as demographic changes (reviewed in Shizuka and Johnson 2019) and disease prevalence (Stroeymeyt et al. 2018). While we were able to control for prior experience and the genotypic composition of social groups, the specific behaviors individuals engaged in remains unknown. We were able to observe how the behavior of locomotor activity was influenced by IGEs for males, though this did not transfer to IGEs for network position in males. Similarly, while we found evidence of IGEs for network position in females, we found no evidence of IGEs for activity in females. These findings support previous studies in flies showing individual’s activity is influenced by the sex (Fujii et al. 2007) and genotype (Lone and Sharma 2011) of interacting individuals. Yet it still remains unclear what behavioral factors shape the genotype-to-network position phenotype relationship, and what behavioral traits allow social partners to shape individuals’ network positions. Future work should continue to link genotypic differences in behavior to DGEs and IGEs for network structure.

Social network analysis has emerged as a well-suited tool for resolving how direct and indirect interactions affect how group structure evolves, as it provides a consistent framework in which social interactions and the genotypes of social group members can
be studied simultaneously. However, studies employing both network analysis and IGEs remain scarce, despite the potential to help resolve unanswered questions about how variation in social group structure arises and evolves. The work presented here is a first step toward integrating IGEs and network theory. Future work should continue to leverage this integration, to resolve questions about how variation in DGEs and IGEs shapes the structure and evolution of social groups.

REFERENCES


in Ecology and Evolution 8:1309–1320.


Kraft, B., E. Williams, V. A. Lemakos, J. Travis, and K. A. Hughes. 2016. Genetic Color


Scarr, S., and K. Mccartney. 1983. How People Make Their Own Environments: A


**SUPPLEMENTARY FIGURES AND TABLES**

*Instrength:* The number of received social interactions an individual has, measured as the amount of time other individuals spent interacting with it.

*Outstrength:* The number of initiated social interactions an individual has, measured as the amount of time it spends interacting with other individuals.

*Eigenvector Centrality:* How central and important an individual is to social group structure, measured by the strength of its direct connections, the strength of its partners’ connections, the strength of its partners’ partners’ connections, etc.

*Betweenness Centrality:* How much an individual serves as a bridge between individuals in the group who do not socially interact, measured as the number of shortest paths between all pairwise combinations of individuals that transmit through a single individual.

*Clustering Coefficient:* How cliqueish an individual is, measured as the proportion of social interactions that are shared amongst an individual’s own direct social partners.

**Supplementary Figure 1:** Depictions and definitions of the five most commonly studied metrics of social network position (Webber and Vander Wal 2019, Krause et al. 2007, Croft et al. 2008, Wey et al. 2008, Sih et al. 2009, Proulx et al. 2005). The red individual and the black lines in each network denote how each measure of network position manifests for the red individual.
<table>
<thead>
<tr>
<th>Genotype Ratio and Composition Treatments</th>
<th>Number of Social Groups with Networks from Only Day 1</th>
<th>Number of Social Groups with Networks from Only Day 2</th>
<th>Number of Social Groups with Networks from Both Days 1 &amp; 2</th>
<th>Sum by Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:9, 208x313 and 786x716</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>1:9, 208x313 and 315x229</td>
<td>6</td>
<td>0</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>1:9, 786x716 and 315x229</td>
<td>2</td>
<td>0</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>9:1, 208x313 and 786x716</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>9:1, 208x313 and 315x229</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>9:1, 786x716 and 315x229</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Sum by Day</td>
<td>19</td>
<td>4</td>
<td>21</td>
<td>Total Networks 65</td>
</tr>
</tbody>
</table>
Chapter 3: Frequency-dependent selection and selection on social network structure also depends on indirect genetic effects in *Drosophila melanogaster*

Eric Wesley Wice¹* and Julia Barbara Saltz¹

¹Department of Biosciences, Rice University, Houston, TX 77005

*Corresponding author (eric.wesley.wice@gmail.com)

**ABSTRACT**

The structure of individuals’ social groups can be an important driver of variation in components of individuals’ fitness. Yet processes that drive variation in social group structure and its impact on fitness -i.e., social selection- are not well-described. The genotypic composition of social groups and how it affects individuals’ fitness -i.e., indirect genetic effects (IGEs)- is thought to be an important driver of variation in social selection; subsequently, studying the quantitative genetic basis of social group structure and its effects on fitness is necessary for understanding how groups both generate and respond to social selection. Two ways that the genotypic composition of social groups and IGEs can impact social selection is via frequency-dependent selection, and selection on individuals’ direct and indirect social interactions (i.e., social network position).

Using replicate genotypes of *Drosophila melanogaster* flies, we manipulated the genotypic composition of social groups, and measured social network positions and multiple components of fitness for individuals within these groups. We found positive frequency-dependent selection operating on measures of females’ fitness, and this frequency-dependent selection varied with indirect genetic effects of social groupmates. Additionally, we found selection operating on measures of individuals’
social network positions, and this selection also varied depending on indirect genetic effects. These findings suggest variation in social selection can be driven by variation in the social context individuals experience. Because social context had a quantitative genetic basis, these findings also suggest social context-dependent social selection provides a mechanism for the adaptive maintenance of social group structure.

INTRODUCTION

Social selection, the process by which the traits of interacting partners affects individuals’ fitness, is a key component of the evolution of sociality (Wolf et al. 1998, 1999; Bijma et al. 2007). Social selection, like natural selection, is likely to vary depending on the temporal or spatial context individuals find themselves in; and in the case of social selection, the specific social context individuals experience (Caruso et al. 2017; Radersma 2020). When the genotypes of socially interacting partners influence individuals’ fitness—a phenomenon termed indirect genetic effects (IGEs)—the social environment itself can respond to selection and evolve, producing unusual and important evolutionary dynamics (Wolf et al. 1998, 1999; Bijma et al. 2007; McGlothlin et al. 2010). A classic example of this process is frequency-dependent selection: the situation in which a genotype’s fitness depends on whether it is common or rare in its population or social group (Ayala and Campbell 1974; Brisson 2018). Frequency-dependent selection can be a vitally important process for shaping phenotypic and genetic variation in populations and groups (Ayala and Campbell 1974; Brisson 2018). The strong evidence for frequency-dependent selection—both positive and negative—on behaviors
across taxa illustrates that an individual’s fitness can depend on the genotype(s) of interacting partners (Ayala and Campbell 1974).

At the same time, recent social network analysis theory and data have highlighted how social interactions play a critical role in shaping fitness and its variation (Krause et al. 2007; Croft et al. 2008; Wey et al. 2008; Sih et al. 2009). Social network analysis provides a quantitative approach for describing the patterning of individuals’ direct and indirect social interactions (Proulx et al. 2005; Krause et al. 2007; Croft et al. 2008; Wey et al. 2008; Sih et al. 2009; Webber and Vander Wal 2019). A key emerging insight from literature utilizing social network analysis is that the interactions occurring amongst social group partners are not always homogenous: some social partners can have an outsize effect on a particular focal individual, compared to that individual’s other social partners (Flack et al. 2006; Pike et al. 2008; Aplin et al. 2013; Firth et al. 2017; Jolles et al. 2017; Jezovit et al. 2020; Rooke et al. 2020). For example, the network structure of pigtailed macaque social groups changes when only a single individual, known to police aggressive interactions within the group, was experimentally removed (Flack et al. 2006; also see "keystone individual concept", Modlmeier et al. 2014). Additionally, the positions that individuals hold in their social network have been shown to impact many aspects of individuals’ fitness: offspring production (Wey et al. 2013; Cheney et al. 2016), mating success (Formica et al. 2012; Farine and Sheldon 2015; Wice and Saltz 2021), and lifespan (Stanton and Mann 2012; Lehmann et al. 2016; Ellis et al. 2017; Blumstein et al. 2018; Thompson and Cords 2018).

To understand how variation in social context impacts social selection, the next step is to integrate these perspectives by studying how both the genotypic composition of
social groups and the patterning of social interactions within groups (measured via social network structure) affects variation in social selection. Currently, it remains unclear how frequency-dependent selection varies depending on the social context (i.e. group genotypic compositions) individuals experience. An individual being rare, versus being common, inherently depends on the traits or genotypes of the social partners that an individual is paired with. Yet, frequency-dependent selection commonly looks at variation from the perspective of an individual’s own traits/genotype, while the traits of an individual’s social context can be just as important in affecting said individual’s fitness (Ayala and Campbell 1974; Wolf et al. 1998, 1999; Westneat 2012; Brisson 2018). This information is critical to understanding how social selection may produce evolutionary change. For example, if negative frequency-dependent selection is context-dependent, then its role in adaptively maintaining genetic variation may be weaker than is commonly suspected. Similarly, it remains unclear how selection on network position might vary across different social contexts, i.e, group genotypic compositions. If selection for high levels of social connectivity is context-independent, this process should generate strong directional selection on the behaviors that produce social connections. Thus, understanding how diverse social relationships among individuals of diverse genotypes together produce individual fitness and its variation will be critical to understanding social selection and its role in behavioral evolution.

Quantifying context-dependent social selection is challenging, in part because it requires measuring selection multiple times (Caruso et al. 2017; Araya-Ajoy et al. 2020). Comparisons of selection in different social groups can be difficult to interpret if the genetic compositions of the groups cannot be measured and/or replicated (Radersma
Similarly, recurrent estimates of selection in a single social group can be difficult to interpret, as social groups are constantly changing and evolving (McGlothlin et al. 2010; Montiglio et al. 2018; Araya-Ajoy et al. 2020). Because of these challenges, it remains unclear what processes generate variation in patterns of social selection.

In this study, we addressed how variation in the genotypic composition of social groups impacted frequency-dependent selection, and selection on social network structure, using *Drosophila melanogaster* flies. Flies are an ideal system for addressing questions about how the quantitative genetic basis of social context impacts selection, as we can generate replicate genotypes of flies in social groups of controlled and variable genotypic composition (Mackay et al. 2012). Additionally, flies have emerged as a model system for studying the quantitative-genetic basis for the structure of social groups, as flies exhibit genetic variation in preference and choice of different social contexts (Saltz 2011, 2017; Saltz and Foley 2011; Geiger and Saltz 2020), and non-random social network structure (Schneider et al. 2012; Pasquaretta et al. 2016; Wice and Saltz 2021). Social context has also been demonstrated to impact selection in flies (Billeter et al. 2012; Saltz 2013), and frequency-dependent selection has been observed to shape behaviors in flies (Fitzpatrick et al. 2007; Kilgour et al. 2018).

Using this system, we created social groups that varied in both their genotypic frequency and genotypic composition. We measured how variation in genotypic composition influenced individual’s social network positions (Chapter 2). Here, we further consider how variation in the genotypic composition and network structure of social groups results in variation in social selection. Our approach allowed us to address
how individuals’ genotypic frequency and network position in a group affected multiple metrics of individuals’ fitness, and whether these processes varied depending on interacting genotypes to produce IGEs. By studying frequency-dependent selection and selection on network structure in conjunction with one another, we can generate a more comprehensive understanding of how variation in social group composition shapes variation in social selection.

METHODS

Genotypes and Social Group Composition
See Chapter 2 for details of how genotypes and social groups were chosen and created.

Fly Behavior and Social Network Analysis
See Chapter 1 for details of how social interactions and network structure were quantified. Also see Supplementary Figure 1 for definitions and depictions of the social network positions analyzed.

Measures of Fitness
See Chapter 1 for details of how fitness metrics were generated. Note that we have only analyzed one measure of fitness for males (number of matings), and have not yet finalized analyses for the effects on males’ latency to mate. Also note that only 6/820 females were observed to mate more than once.

Replication
See Chapter 2 for details on social group and network sample sizes. If any fly from a social group died or escaped before measures of fitness were taken, all flies from the group were excluded from analyses. For females, we measured lifespan and total lifetime offspring production for 298 individuals, excluding females who either died during cold
anesthesia or escaped after they had been removed from their social groups: 1:9 ratio of genotypes 208x313 and 786x716, respectively (n = 66); 1:9 ratio of 208x313 and 315x229 (n = 48); 1:9 ratio of 786x716 and 315x229 (n = 39); 9:1 ratio of 208x313 and 786x716 (n = 69); 9:1 ratio of 208x313 and 315x229 (n = 49); 9:1 ratio of 786x716 and 315x229 (n = 27; Supplementary Table 1). Of these, 193 individuals also had measured network positions: 1:9 ratio of genotypes 208x313 and 786x716, respectively (n = 18); 1:9 ratio of 208x313 and 315x229 (n = 38); 1:9 ratio of 786x716 and 315x229 (n = 39); 9:1 ratio of 208x313 and 786x716 (n = 49); 9:1 ratio of 208x313 and 315x229 (n = 22); 9:1 ratio of 786x716 and 315x229 (n = 27; Supplementary Table 1). For males, we measured the number of matings for 820 individuals: 1:9 ratio of genotypes 208x313 and 786x716, respectively (n = 140); 1:9 ratio of 208x313 and 315x229 (n = 170); 1:9 ratio of 786x716 and 315x229 (n = 110); 9:1 ratio of 208x313 and 786x716 (n = 160); 9:1 ratio of 208x313 and 315x229 (n = 130); 9:1 ratio of 786x716 and 315x229 (n = 110; Supplementary Table 1). Of these, 400 individuals also had measured network positions: 1:9 ratio of genotypes 208x313 and 786x716, respectively (n = 40); 1:9 ratio of 208x313 and 315x229 (n = 100); 1:9 ratio of 786x716 and 315x229 (n = 80); 9:1 ratio of 208x313 and 786x716 (n = 70); 9:1 ratio of 208x313 and 315x229 (n = 30); 9:1 ratio of 786x716 and 315x229 (n = 80; Supplementary Table 1).

**Analyses of Female Lifetime Offspring Production and Lifespan**

We modeled how females’ lifetime offspring production and lifespan were associated with each individual’s own genotype (DGE), the genotype of their social group partner(s) (IGE), and an individual’s genotypic frequency within their social group (1/20 for focal females, and 9/20 for stimulus female genotypes) using linear mixed models (LMMs)
Each females’ own genotype was modeled as a fixed effect, as opposed to a random effect, because genotypes were non-randomly chosen (see Chapter 2), and had only three levels. We also included a random effect for the identity of each group, reflecting the fact that each individual was treated as a sub-sample of its group, and that groups are the unit of replication for this study. Model fits were assessed using R package DHARMa 0.2.7 (Hartig 2020). Females’ lifespan was found to be best modeled using a Poisson-distributed generalized linear mixed model (GLMM). All LMMs and GLMMs were constructed using R package lme4 1.1 (Bates et al. 2015).

We tested for a three-way interaction between a DGE, IGE, and an individual’s genotypic frequency in a group, as well as all pairwise two-way interactions between these effects, using Type III Wald $\chi^2$ tests (Bolker et al. 2009). Non-significant interactions were broken down into their main effects, and the significance of all subsequent interactions and main effects were assessed using Type III Wald $\chi^2$ tests.

To further understand our initial results, we tested how measures of each female’s position within their social network influenced their lifespan and lifetime offspring production, as well as how this effect varied depending on which other genotype was present (IGEs). Measures of network positions are often colinear, which presents problems when analyzing multiple colinear variables in multiple regression models (James et al. 2013; Wice and Saltz 2021). As such, we analyzed the effect of each of our five measures of network position on female fitness traits in separate models. These models included fixed factors of DGEs, individuals’ genotypic frequency within their group, and the interaction between IGEs and each measure of network position. This approach directly tests the hypothesis that the fitness effect of occupying a specific
network position depends on the social context in which these traits are expressed. Non-significant interactions were broken down into their main effect components, and significance of effects within the resulting models was assessed using Type III Wald $\chi^2$ tests. All analyses were performed in R version 3.6.2 (Team 2019).

**Analyses of Male Mating Success**

To investigate the role of social group composition on male mating success (number of mates acquired), we tested whether the identity of focal female genotype, the stimulus female genotype, and their interaction, affected male mating success. Note that for males we did not include an effect of the male’s own genotype (DGEs) because all males in this experiment were from a single genotype. We again included a random effect of group identity. Model fit was assessed using R package DHARMa 0.2.7, and our data were found to be better fit using a Poisson-distributed GLMM (Bolker et al. 2009). Fixed effects were tested using Type III Wald $\chi^2$ tests (Bolker et al. 2009). Non-significant interaction effects were reduced to their main effects, and their significance assessed using Type III Wald $\chi^2$ tests.

To better understand how social group dynamics influenced males’ mating success, we tested how our five measures of social network position, and their interactions with the social group composition of females, impacted mating success. Similar to our analyses investigating female fitness traits, each measure of network position was tested in a model on its own. Each model contained a three-way interaction between males’ network position, the focal female genotype, and the stimulus female genotype. Non-significant three-way and two-way interactions were broken down into
their main effects, and the significance of effects within these models was assessed using Type III Wald $\chi^2$ tests. All analyses were performed in R version 3.6.2 (Team 2019).

RESULTS

Positive frequency-dependent selection for female fitness varies with indirect genetic effects

We found evidence of DGEs influencing females’ lifetime offspring production ($\chi^2 = 7.965, P = 0.019$), but not lifespan ($\chi^2 = 0.714, P = 0.700$). We found only marginal evidence of IGEs influencing females’ lifespan on their own ($\chi^2 = 4.614, P < 0.100$), but no evidence of IGEs influencing offspring production ($\chi^2 = 1.038, P = 0.595$). A female’s genotypic frequency in a group however was a significant predictor of both females’ offspring production ($\chi^2 = 3.850, P < 0.050$) and lifespan ($\chi^2 = 9.150, P = 0.002$; Figure 1). More specifically, inspections of least-squares means indicated that stimulus females (i.e., the common genotype in each group) produced 22 more offspring, and lived 8.5 days longer, compared to focal females (i.e., the rare genotype in each group), indicating the presence of positive frequency-dependent selection. Interestingly, the magnitude of frequency-dependent selection for lifespan also depended on the genotype of the social partner(s) individuals were paired with (IGE-by-frequency interaction, $\chi^2 = 7.188, P = 0.027$; Figure 1).
This effect was not observed to influence offspring production (IGE-by-frequency interaction, $\chi^2 = 0.100, P = 0.951$). We also did not detect the presence of a three-way interaction between DGEs, IGEs, and genotypic frequency in a group; or two-way interactions between DGEs and IGEs or genotypic frequency (all $P > 0.05$).

**Selection on female network position varies with indirect genetic effects for lifespan only**

We found no effects of network position or its interaction with IGEs on offspring production (all $P > 0.05$). We did however, find effects of network position influencing females’ lifespan. Specifically, the number of interactions females engaged in (instrength and outstrength) and how central females were to the structure of their social groups (eigenvector centrality) were strong predictors of females’ lifespan (instrength, $\chi^2 =$
16.966, \( P < 0.001 \); outstrength, \( \chi^2 = 15.145, P < 0.001 \); eigenvector centrality, \( \chi^2 = 15.443, P < 0.001 \); Figure 2). Inspection of least-squares means indicate that engaging in more social interactions (instrength and outstrength) and being more central to social group structure (eigenvector centrality) were correlated with lower lifespans. Additionally, how females’ network positions influenced their lifespan was also dependent on the genotype of their social group partner(s) (IGE-by-instrength, \( \chi^2 = 18.411, P < 0.001 \); IGE-by-outstrength, \( \chi^2 = 15.643, P < 0.001 \); IGE-by-eigenvector centrality, \( \chi^2 = 17.797, P < 0.001 \); Figure 2).

**Figure 2:** Female individuals’ lifespan correlated with the number of social interactions they engaged in (A,B) and how central they were to the structure of their social groups (C). How the network positions of instrength, outstrength, and eigenvector centrality correlated with lifespan varied depending on an IGE of which genotype individuals were paired with (the color of each each line. Lines and shading indicate best-fit linear relationships and standard errors). Genotype labels indicate the maternal parent genotype crossed with ("x") the paternal genotype to create replicate heterozygous genotypes (Mackay et al. 2012). Note that genotype numbers are arbitrary and do not describe similarities or differences between genotypes.

**Selection on male network position varies with indirect genetic effects**

The genotype of focal and stimulus females, and an interaction between them, were not found to affect male mating success (all \( P > 0.05 \)). We also found no evidence of main effects of males’ network positions on mating success (all \( P > 0.05 \)). We did however
find that how many social interactions males initiated (outstrength) predicted mating success, but this effect was dependent on both the female stimulus genotype (outstrength-by-stimulus genotype, $\chi^2 = 6.713, P = 0.035$), and an interaction between the female stimulus and female focal genotypes in the males’ group (outstrength-by-stimulus genotype-by-focal genotype, $\chi^2 = 4.692, P = 0.030$; Figure 3).

![Figure 3: How the number of interactions a male initiated (outstrength) affected their mating success depending not only on the stimulus female genotype present (colored lines), but also the focal female genotype present (dashed lines). Lines indicate best-fit linear relationships. Genotype labels indicate the maternal parent genotype crossed with (“x”) the paternal genotype to create replicate heterozygous genotypes (Mackay et al. 2012). Note that genotype numbers are arbitrary and do not describe similarities or differences between genotypes.](image)

This result indicates that selection on network position for males depended on the genotypic makeup of conspecifics in their social group. We did not find evidence of any other three- or two-way interactions between males’ network positions and the identities of the focal and stimulus genotypes they were paired with (all $P > 0.05$).

**DISCUSSION**

Indirect genetic effects of social partners on individuals’ fitness are filtered through effects on individuals’ phenotypes, upon which social selection then acts (Wolf et al.
An individual’s social network position is a prime target for a phenotype that social partners can influence, as measures of network position are inherently dependent on interactions with social groupmates (Proulx et al. 2005; Krause et al. 2007; Croft et al. 2008; Wey et al. 2008; Sih et al. 2009). Previous research demonstrated that indirect genetic effects can indeed influence individuals’ network positions, specifically network positions that depend on indirect social interactions (Chapter 2). Here, we additionally found that indirect genetic effects shape variation in social selection. Females’ genotypic frequency in a social group affected multiple components of fitness. Specifically, we found positive frequency-dependent selection, whereby the more common female genotype in a group -i.e., the stimulus genotype- lived longer and produced more offspring compared to focal females. We also found selection acting on measures of females’ network positions, where females who engaged in more social interactions and were more central to group structure had shorter lifespans. For both our findings of frequency-dependent selection, and selection on females’ social network positions, we also found that these patterns of selection varied depending on the genotypic composition of females that individuals were paired with. Similarly, we found that selection on males’ tendency to initiate social interactions depended on the genotypic composition of females in their social groups.

When patterns of selection vary depending on the context they are expressed in, this provides a key mechanism for adaptively maintaining variation (Bailey and Zuk 2012; Caruso et al. 2017; Wice and Saltz 2021). Positive frequency-dependent selection typically reduces phenotypic or genotypic variation, because the more common phenotype or genotype enjoys a fitness benefit and becomes even more common in
subsequent generations (Ayala and Campbell 1974; Brisson 2018). But because we found the strength of positive frequency-dependent selection varied depending on the social context - i.e., genotypic composition of social groups- individuals experienced, these variable patterns of selection have the potential to adaptively maintain variation. Similarly, genotypic frequencies within groups are likely to change and evolve over multiple generations. Feedbacks between variable patterns of social selection, and constantly changing genotypic compositions of social groups, can further shape variation in social group structure (Araya-Ajoy et al. 2020). These findings thus highlight the key role of indirect genetic effects for fitness in shaping genetic variation within and among populations.

The research presented here indicates that indirect genetic effects on fitness are filtered through measures of network position, and particularly those measures of network position that are more indicative of individuals’ direct social interactions. More specifically, measures of the number of social interactions individuals engage in (instrength and outstrength) encompass information about only direct social interactions; how central individuals are to the structure of their network (eigenvector centrality) does take into account indirect interactions, though individual’s first-order connections - i.e., the social partners individuals are directly connected to- carry stronger weight in influencing centrality to network structure (Croft et al. 2008; Wey et al. 2008). These more direct measures of network position were found to interact with IGEs to influence multiple components of fitness, while the other two measures of network position we tested (betweenness centrality and clustering coefficient, which are more descriptive of indirect social interactions) did not influence fitness.
One hypothesis for why measures of social network position that encompass direct social interactions were more likely to statistically interact with IGEs to influence individuals’ fitness could be that individuals have more control over their direct social interactions, though this is not always the case; previous studies have found that measures of network position that are more descriptive of indirect social interactions can be more heritable than measures exclusively incorporating direct social interactions (Fowler et al. 2009; Lea et al. 2010; Brent et al. 2013; Wice and Saltz 2021). The fact that more direct measures of social network position statistically interacted with IGEs to influence fitness, could indicate that individuals modified their behaviors in response to their social environment, but only in a way that ultimately impacted their fitness. For example, our findings that the number of social interactions males initiated (outstrength) only affected their mating success depending on the female genotype(s) they were paired with is possibly the result of males changing which female genotype they socially interacted with, in response to how receptive females were to males’ initiated social interactions. This could also explain why the genotype of a single focal female in the group statistically interacted with both the stimulus female genotype and males’ outstrength to impact mating success, if males redirected all of their social attention onto a single individual when the stimulus female genotype in the group was less receptive to mating.

A limitation of our study is that social groups composed of many genetically-identical individuals are not the norm in wild fly populations. However, in groups of many different genotypes, it becomes exceedingly challenging to parse the effects of all pairwise combinations of genotypes on one another’s phenotypes and fitness (Saltz 2013). While social groups composed of many genotypically identical individuals are less
common, there are many examples of highly related individuals occupying the same social groups (Pamilo 1989; De Ruiter and Geffen 1998; Archie et al. 2006; Mathot and Giraldeau 2010). Despite the challenges associated with manipulating the genotypic composition of social groups and testing the effects of genotypes on one another, these findings are a necessary first step toward understanding how variation in social group composition drives variation in social selection.

Our findings of how social group composition and network position influence components of fitness are highly complementary to other studies investigating these processes. Social interactions and individuals’ rank in dominance hierarchies has been shown to influence offspring production (Pusey et al. 1997; Silk et al. 2003; Strauss and Holekamp 2019) and mating success (Billeter et al. 2012; Saltz 2013); and social network positions have been shown to influence lifespan (Stanton and Mann 2012; Lehmann et al. 2016; Ellis et al. 2017; Blumstein et al. 2018; Thompson and Cords 2018), offspring production (Wey et al. 2013; Cheney et al. 2016), and mating success (Formica et al. 2012; Farine and Sheldon 2015; Wice and Saltz 2021) as well. Interestingly, our results are highly comparable to some of these studies, such as findings from Blumstein et al. (2018) that indicated the number of affiliative interactions females engaged in was negatively associated with lifespan in marmots. However, our findings did not necessarily align with the strength and directionality of effects in other studies.

Overall, our findings illustrate the complex relationships between social structure and fitness: when considering how social structure affects selection, it is important to consider the composition of individuals within the social group being studied. Indeed, based on the findings of our study, one of the potential reasons experiments investigating
the social effects on individuals’ fitness come to variable conclusions is that these experiments are performed in systems where the exact genotypic or phenotypic makeup of a group is either unknown or highly variable (Radersma 2020). Our ability to manipulate and finely control the genotypic composition of social groups revealed that both phenotypic and genotypic variation among individuals within groups may drive variation in social selection. Estimating selection across multiple contexts, whether social or non-social, is admittedly challenging (Araya-Ajoy et al. 2020). However, doing so is a necessary next step toward understanding the contexts and conditions in which variation in social structure evolves. While we were able to address how one generation of selection can be variable across social contexts, selection is an ongoing process; feedbacks between evolutionary changes in social group composition and selection acting on social dynamics are likely to further shape how social groups are structured (Araya-Ajoy et al. 2020). Future work should continue to address the drivers of variation in social selection in order to develop a better understanding of how diverse social structures are formed and evolve.

REFERENCES


James, G., D. Witten, T. Hastie, and R. Tibshirani. 2013. An Introduction to Statistical Learning with Applications in R. Springer.

Jezovit, J. A., R. Rooke, J. Schneider, and J. D. Levine. 2020. Behavioral and
environmental contributions to drosophilid social networks. Proceedings of the National Academy of Sciences.


phenotypes and the evolutionary process. III. Social evolution. Evolution 64:2558–2574.


Wice, E. W., and J. B. Saltz. 2021. Selection on heritable social network positions is
context-dependent in Drosophila melanogaster. Nature Communications

Accepted.


SUPPLEMENTARY FIGURES AND TABLES

**Instrength:** The number of received social interactions an individual has, measured as the amount of time other individuals spent interacting with it.

**Outstrength:** The number of initiated social interactions an individual has, measured as the amount of time it spends interacting with other individuals.

**Eigenvector Centrality:** How central and important an individual is to social group structure, measured by the strength of its direct connections, the strength of its partners’ connections, the strength of its partners’ partners’ connections, etc.

**Betweenness Centrality:** How much an individual serves as a bridge between individuals in the group who do not socially interact, measured as the number of shortest paths between all pairwise combinations of individuals that transmit through a single individual.

**Clustering Coefficient:** How cliquish an individual is, measured as the proportion of social interactions that are shared amongst an individual’s own direct social partners.

**Supplementary Figure 1:** Depictions and definitions of the five most commonly studied metrics of social network position (Webber and Vander Wal 2019, Krause et al. 2007, Croft et al. 2008, Wey et al. 2008, Sih et al. 2009, Proulx et al. 2005). The red individual and the black lines in each network denote how each measure of network position manifests for the red individual.
<table>
<thead>
<tr>
<th>Genotype Frequency and Composition Treatments</th>
<th>Lifespan / Lifetime Offspring Production Sample Sizes</th>
<th>Lifespan / Lifetime Offspring Production with Network Data Sample Sizes</th>
<th>Mating Success Sample Sizes</th>
<th>Mating Success with Network Data Sample Sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:9, 208x313 and 786x716</td>
<td>66</td>
<td>18</td>
<td>140</td>
<td>40</td>
</tr>
<tr>
<td>1:9, 208x313 and 315x229</td>
<td>48</td>
<td>38</td>
<td>170</td>
<td>190</td>
</tr>
<tr>
<td>1:9, 786x716 and 315x229</td>
<td>39</td>
<td>39</td>
<td>110</td>
<td>80</td>
</tr>
<tr>
<td>9:1, 208x313 and 786x716</td>
<td>69</td>
<td>49</td>
<td>180</td>
<td>70</td>
</tr>
<tr>
<td>9:1, 208x313 and 315x229</td>
<td>49</td>
<td>22</td>
<td>130</td>
<td>30</td>
</tr>
<tr>
<td>9:1, 786x716 and 315x229</td>
<td>27</td>
<td>27</td>
<td>110</td>
<td>80</td>
</tr>
<tr>
<td>Total</td>
<td>298</td>
<td>193</td>
<td>820</td>
<td>430</td>
</tr>
</tbody>
</table>

**Supplementary Table 1:** This table summarizes the sample sizes for measures of fitness for females (lifespan and lifetime offspring production) and males (mating success), broken down by the specific genotypic frequency and composition treatment of social groups. It also shows what subset of these samples also had social network data generated to pair with the fitness data.