Identifying Social-Epidemiological Roles Associated with Viral Exposure Using Regular Equivalence Blockmodeling

Tyler M. Barrett[1], Charles Kevin Tiu[2], Michelle Pender[3], Jean Yves Rabezara[4], Prisca Rahary[4], Lin-Fa Wang[2], Randall A. Kramer[3,5], Voahangy Soarimalala[4,6], Peter J. Mucha[7], James Moody[8], Charles L. Nunn[1,3]

[1]Department of Evolutionary Anthropology, Duke University, Durham, NC, USA

[2]Programme in Emerging Infectious Disease, Duke-NUS Medical School, Singapore

[3]Duke Global Health Institute, Duke University, Durham, NC, USA

[4]Association Vahatra, Antananarivo, Madagascar

[5]Nicholas School of the Environment, Duke University, Durham, NC, USA

[6]Institut des Sciences et Techniques de l'Environment, Université de Fianarantsoa, Fianarantsoa, Madagascar

[7]Department of Mathematics, Dartmouth College, Hanover, NH, USA

[8]Department of Sociology, Duke University, Durham, NC, USA

Abstract

Identifying individuals at high risk of infection is critical to guiding interventions during infectious disease outbreaks. Network centrality can characterize infection risk, but its utility varies across pathogens and social systems. We applied a new approach – equivalence-based blockmodeling – to identify social-epidemiological roles associated with viral exposure among 1,297 adults in northeast Madagascar. Roles were determined based on social networks derived from reported shared free time and exchanges of food and farmwork. We identified three distinct roles, including individuals with many reciprocated social ties (Popular), individuals who sent many ties with few reciprocated (Core), and individuals who had few connections (Periphery). To assess whether the roles covaried with viral exposure, Phage ImmunoPrecipitation Sequencing was performed using dried blood spot samples and VirScan to determine exposure to 337 virus species and subtypes. Membership in the Popular (β [95% CI]: 0.29 [-0.06-0.63]) and Core (β [95% CI]: 0.36 [0.12-0.60]) roles was associated with greater viral exposure than membership in the Periphery role. Roles performed better at predicting exposure than single measures of centrality. Equivalence-based blockmodeling extends the utility of network centrality for characterizing infection risk and provides new insights into how social roles may relate to both pathogen exposure and susceptibility to infection.

Introduction

The social patterning of infectious disease depends on *whom* you interact with and *how* you interact with them. Advances in network analysis have highlighted the importance of this individual-level heterogeneity for epidemic dynamics. For example, superspreading – where individuals disproportionately infect many other individuals in a population – helped drive the explosive spread of SARS-CoV-1 during the 2002-2004 pandemic [1, 2]. Superspreading has also been important for understanding the transmission of SARS-CoV-2, Ebola, and measles [3-5]. Network centrality-based approaches have since been used to identify individuals with superspreading potential and/or high risk of infection, i.e., individuals who have many connections (e.g., degree centrality) or specific types of connections (e.g., betweenness centrality) that are known to facilitate parasite transmission [6-12].

Simulation-based studies have shown that an individual's risk of infection and time to infection over the course of a simulated epidemic are positively associated with multiple measures of centrality, including degree, strength, betweenness, and eigenvector centrality [6, 7, 13-18]. Observational studies of infection patterns across a diverse range of host and parasite species also show that more central individuals tend to have a higher probability of infection [19-23]. Thus, centrality measures can be powerful tools for identifying individuals at high risk of infection and are often proposed as a method of guiding the distribution of vaccines and testing resources during outbreaks [7, 18, 24]. However, the predictive performance of centrality is often influenced by characteristics of the parasite and the social system being modeled. For example, centrality-based targeting of vaccines in wealthy countries appears less effective than other targeting methods when parasites have a high R_0 [18].

On many social networks, parasites can spread through multiple transmission pathways. These different transmission pathways can involve different types of social relationships and processes, such as sexual contact versus needle sharing in the spread of HIV or shared environmental exposures versus social contact in the spread of gastrointestinal parasites [23, 25-30]. These processes often involve different probabilities of transmission along network edges that can greatly influence disease dynamics. In the case of HIV, for example, needle sharing is associated with a higher risk of transmission than sexual intercourse, and among sexual contacts, transmission from males to females is higher than the reverse [25]. Thus, centrality on its own misses important dimensions of social relationships like the differential risk of transmission within and among groups and the social factors that drive transmission dynamics. Understanding these social roles can help guide public health interventions beyond simple estimates of superspreading potential based on single-relation centrality scores.

Prior network studies of infectious disease transmission have attempted to account for these differences in transmission probability. For example, to better understand the heterogeneous effects of non-pharmaceutical interventions during the COVID-19 pandemic, one study varied the relative edge weights of social contacts within and outside of the household and found that social distancing was most effective when within household contacts were weighted either very high or very low relative to external contacts [31]. Another approach is to create an abstracted measure of social activity that aggregates interaction information across a whole network. For example, social fluidity is a measure that accounts for both the density and weight of edges and can be used to assess the gregariousness of a social group and its effect on transmission across different domains of interaction [32].

A third approach – and the focus of this paper – is role analysis via equivalence-based blockmodeling. This approach clusters individuals based on equivalent patterns of relations across multiple networks or types of relations [33-36]. Roles therefore reflect distinct patterns of *how* individuals are connected across networks rather than *whom* specifically they are connected to. In infectious disease ecology and network epidemiology, role analysis can identify classes of nodes with equivalent risk profiles across multiple relations simultaneously. While traditional centrality scores used in epidemiology capture one dimension of transmission risk within one relation (e.g., degree being most relevant for local transmission and betweenness for network-wide transmission), assigning roles based on blockmodels allows a researcher to differentiate subsets of nodes with similar connection profiles and to identify how distinct profiles might contribute to different parasite transmission outcomes. In the social sciences, role analysis has helped answer research questions about how informal social roles are related to marriage exchange, the accumulation of political power, political polarization, and the formation of social hierarchies [37-40]. To date, role analysis has had less influence in ecology and evolutionary biology, but it has been used to study food webs in Malaysia and Florida as well as the social structure of rhesus monkeys (*Macaca mulata*) in Pakistan and Cayo Santiago [41, 42]. To our knowledge, role analysis has not yet been applied to disease ecology or network epidemiology.

Here, we characterize distinct social roles based on friendship, food exchange, and farm co-working in three rural agricultural villages in northeast Madagascar using regular equivalence-based blockmodels. We then assess whether these roles are associated with different patterns of viral exposure using a multiplex serological assay (Phage ImmunoPrecipitation Sequencing [PhIP-Seq]) and the VirScan library [43, 44]. VirScan captures the presence of antibodies to over 300 virus species and subtypes, allowing for the broad characterization of an

individual's history of viral exposure. Prior studies have used VirScan to determine differences in viral exposure profiles between Batwa hunter-gatherers and Bakiga agriculturalists in Uganda, finding that the hunter-gather population had relatively higher seropositivity of double-stranded DNA viruses [45]. Given the durability of social roles [46, 47], we expect the roles identified via equivalence-based blockmodels in rural Madagascar to reflect relatively long-term and distinct patterns of engaging in social life that influence an individual's viral encounters.

Our study had two primary objectives. First, we aimed to provide a case study of how regular equivalence blockmodeling can provide a richer characterization of a social system beyond individual centrality measures. Second, we aimed to assess whether the roles identified can be considered social-epidemiological roles. Within this objective, we hypothesized that individuals who were most active and heterogenous in their social interactions (i.e., had high centrality across networks and sent and received many different social ties) would be exposed to more viruses. We also expected that the social roles would perform better at predicting viral exposure than centrality scores alone, given that roles reflect multiple transmission pathways.

Methods

Data Collection

Survey data and dried blood spot samples were collected from 1,297 adults (≥18 years old) in three villages (Village A, B, and C) bordering Marojejy National Park in the SAVA region of northeast Madagascar. All data were collected between October 3, 2019 and August 4, 2022, and the study was approved by the Duke University IRB (protocol no. 2019-0560) and the Malagasy Ethics Committee for Biomedical Research within the Ministry of Public Health

(114/MSANP/SG/AGMED/CERBM). All study participants provided informed consent and were compensated for their study participation.

Surveys were administered in the local Malagasy dialect by a trained research team to collect information about participants' socioeconomic characteristics and social networks. An initial set of participants who owned agricultural land were identified in each village. We then used snowball sampling to recruit participants who were named as someone with whom a participant spent free time, exchanged food, or exchanged farmwork [48]. Each participant reported their age, gender, highest level of education, occupation, and household size.

We used two measures to assess participants' market-based wealth [26]. First, we used an inventory of the durable goods an individual's household owned, including whether their household owned a mobile phone, television, bicycle, motorcycle, generator, refrigerator, and computer [49, 50]. We then created an index with values ranging from 0 (no durable goods owned) to 7 (all durable goods owned). Second, we created an index of the material used to construct the floor, walls, and roof of a participant's house, ranging from more gatherable materials (e.g., dirt, bamboo, raffia palm) to more expensive purchased materials (e.g., bricks, concrete, metal roofing sheets) [26, 51]. Each house component was ranked along this spectrum and then the values were standardized. The standardized values for each house component were summed to create the final index (house lifestyle index) with higher values indicating greater market-based wealth and lower values indicating less market-based wealth.

We used five name-generating survey questions to construct the social networks used in this analysis [52]. All participants were asked to name up to five people with whom they spent their free time, provided and received food assistance, and provided and received farmwork assistance. Each of the five directed networks were included in the role analysis. To create single

summary measures of centrality, we created a weighted network for each of the three villages where edge weights ranged from 0 (no relation) to 5 (all relations). In summary networks, where each of the five relations were represented as an edge weight ranging from zero to five, Village A (estimated village population size = 2,700) had 262 nodes and 952 edges, Village B (estimated village population size = 900) had 435 nodes and 1,642 edges, and Village C (estimated village population size = 1,900) had 600 nodes and 3,049 edges. Summary networks have been characterized in more detail in prior work [53].

Phage ImmunoPreccipitation Sequencing

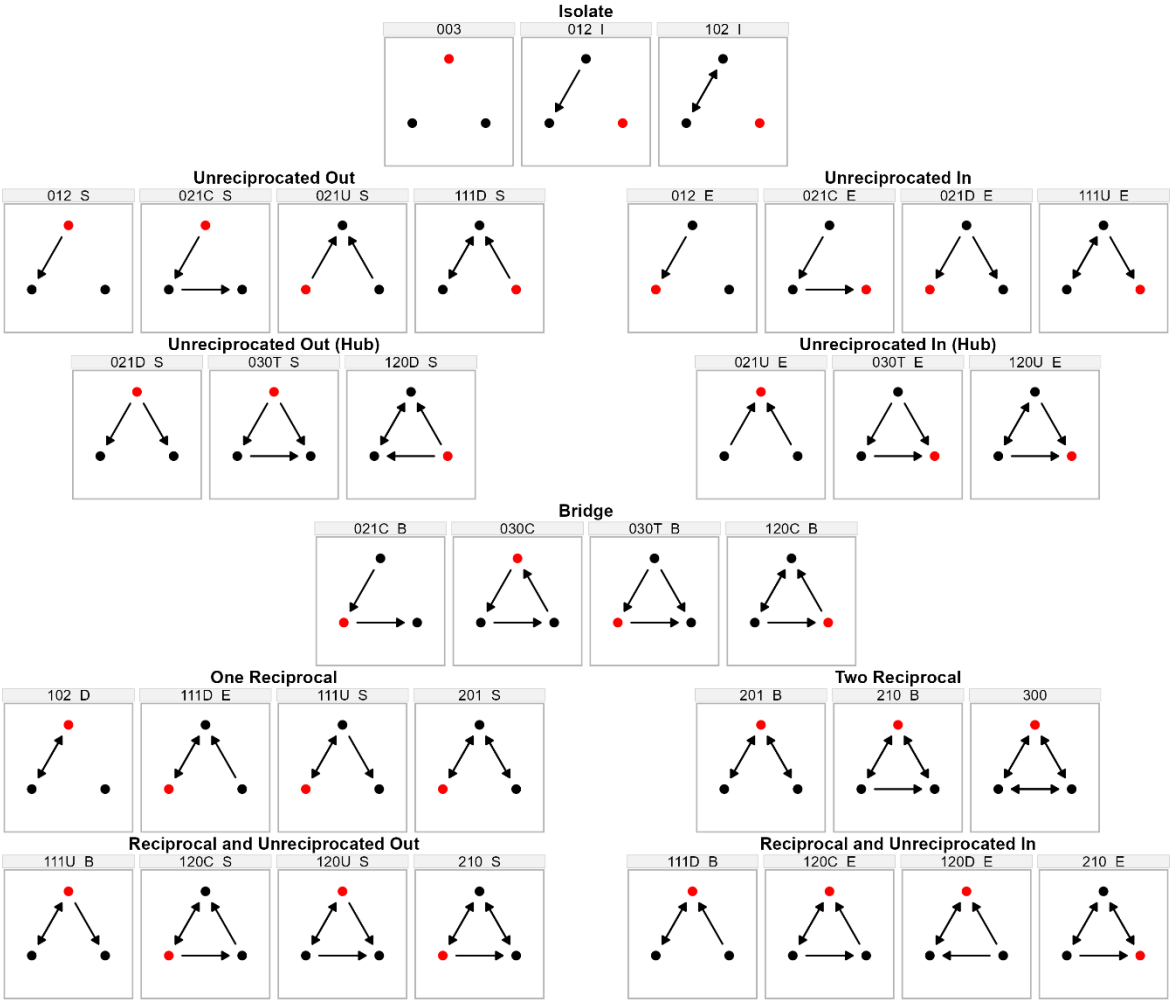
Dried blood spot samples were collected on Whatman 903 filter paper cards and stored at -80°C until analysis. Blood samples from the filter paper cards were extracted using a solution of 0.05% Tween-20 in Phosphate Buffered Saline (PBS). Punched out blood spots were incubated with the extraction solution overnight. Debris was then excluded by centrifugation, and the resulting supernatant was used for subsequent analysis.

PhIP-Seq VirScan was performed using previously described protocols [54, 55]. Briefly, the extracted samples were incubated with the VirScan bacteriophage library [44]. The resulting mixture was then immunoprecipited with Protein A and Protein G Dynabeads. A wash step was then performed, which was followed by nested PCR reactions to prepare the library for high throughput sequencing. This process extracts the gene of interest (coded in the bacteriophage genome) and appends a barcode sequence to allow for the multiplexing of samples in a sequencing run. High throughput sequencing was performed with BGI DNBSeq-G400 150-bp paired-end (BGI Hong Kong). Resulting sequences were demultiplexed and analyzed using in-house scripts described previously [54].

Data Analysis

All data analyses were conducted with R version 4.4.0 [56]. To identify social roles, we used the role_analysis function in the *ideanet* package [57]. We operationalized roles using regular equivalence, which characterizes individuals who have similar patterns of social ties but who are not necessarily connected to the same people [58]. For any directed network, there are 36 positions a node can occupy in a triad **(Figure 1)**. For each individual participant, we generated a vector of counts of the numbers of times that they occupied each of these 36 triad positions in each of the five networks [34]. We then augmented this triad position vector with centrality measures computed for each individual participant in each network. This augmentation included a wide range of transmission-relevant measures: total degree, indegree, outdegree, betweenness, power, eigenvector, and closeness centrality. Similarity in participants' triad position and centrality vectors was determined using hierarchical clustering, and the final roles were determined by maximizing cluster modularity [36]. We then characterized the resulting blockmodel ("image graph") for each village and identified a typology of role categories across villages.

Figure 1. Typology of triad positions. Each panel shows one of 36 triad positions in 16 directed triads, with the focal node in red.



We took two approaches to assess the relationship between social roles and viral exposure profiles. First, we assessed bivariate relationships between role category and infection with viruses that had >10% seroprevalence in the sample. To do so, we fit bivariate Bayesian logistic regression models with weakly informative priors (normal distribution with a mean of zero and standard deviation of one) using the brm function in the *brms* package [59]. Second, we assessed the importance of role category as a predictor of viral species and subtype richness using a multi-model comparison and averaging approach with the *MuMIn* package [60]. Each model used a

zero-inflated Poisson distribution to account for zero-inflation in richness. We compared the predictive performance of models containing different sets of predictor variables (age, gender, education, occupation, household size, durable goods ownership, housing material, village, individual centrality measures [strength, betweenness, eigenvector, and closeness], and role category) using sample-size adjusted Akaike Information Criterion (AICc) [61, 62]. This approach allowed us to assess the direction of the relationship between role categories and richness. It also provided a means to compare the importance of role categories relative to demographic characteristics and individual centrality measures for predicting richness, with importance measured as the summed AICc for all models that included that variable. All continuous variables were standardized prior to the analysis, and weighted coefficients were computed across a set of models with a delta AICc less than two.

Results

Participant Demographic Characteristics

Participant demographic characteristics for each village are presented in **Table 1**. Overall, participants' median (IQR) age was 33 (24-48) years, and 46% (n = 593) were women. Nearly all participants (96%, n = 1,237) had at least some education, and 12% (n = 158) had above a secondary-level education. Most (90%, n = 1,166) reported farming as their main occupation. Median (IQR) household size was 4 (3-5) people, and participants had a median (IQR) household lifestyle index of -1.55 (-1.55-1.84), indicating that their homes tended to be constructed of materials readily available in the environment (e.g., wood, ravinala) rather than purchased materials (e.g., metal, concrete). Participants reported that their household owned a median (IQR) of 1 (0-2) durable goods.

Table 1. Participant demographic characteristics by village.

Demographic Characteristics	Village A (n = 262)	Village B (n = 435)	Village C (n = 600)	All Villages (n = 1,297)
Age, median (IQR)	40 (30-53)	30 (22-44)	32 (24-47)	33 (24-48)
Unknown	1	1	0	2
Gender, n (%)				
Women	120 (46%)	182 (42%)	291 (49%)	593 (46%)
Men	142 (54%)	253 (58%)	309 (52%)	704 (54%)
Education, n (%)				
None	14 (5.3%)	21 (4.9%)	14 (2.4%)	49 (3.8%)
Primary	158 (60%)	179 (41%)	243 (41%)	580 (45%)
Secondary	67 (26%)	178 (41%)	254 (43%)	499 (39%)
Higher	23 (8.8%)	54 (13%)	81 (14%)	158 (12%)
Unknown	8 (3%)	0 (0%)	3 (0.5%)	11 (0.8%)
Occupation, n (%)				
Farmer	256 (98%)	388 (89%)	522 (87%)	1,166 (90%)
Non-Farmer[1]	6 (2.3%)	46 (11%)	76 (13 %)	128 (9.9%)
Unknown	0 (0%)	1 (0.2%)	2 (0.3%)	3 (0.2%)
House Lifestyle Index, median (IQR)	-1.55 (-1.55-1.84)	0.63 (-0.34-1.84)	0.63 (-1.55-1.84)	-1.55 (-1.55-1.84)
Durable Goods Owned, median (IQR)	1 (0-2)	1 (0-2)	1 (0-2)	1 (0-2)
Household Size, median (IQR)	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)

[1]Non-farming occupations included carpenter, taxi driver, laborer, shopkeeper, teacher, student, and unemployed.

Role Analysis

Eight distinct roles were identified for Village A. Five roles were identified for Village B, and four roles were identified for Village C **(Figure 2)**. The role structure across all three villages suggested three primary categories of roles: (1) individuals who were "Popular" and named by many individuals across nearly all roles in a village, (2) individuals who were "Core" to a village network but who mostly named others in the village rather than being named themselves, and (3) individuals on the "Periphery" of a village network and minimally connected to the Popular

individuals. Age, gender, education, and occupation were relatively similar across the three categories **(Table 2)**.

Figure 2. Role analysis summary. Panel A shows the blockmodels for the three villages, with nodes representing roles in the Popular (triangle), Core (square), and Periphery (circle) role categories. Edges between nodes indicate whether members of one role nominated members of another role as people with whom they spend free time (purple), exchange food (green), and co-farmwork (orange). Panel B shows the standardized frequency at which members of a role category occupied different types of triad positions averaged across relations and villages. Panel C shows the standardized centrality scores for members of each role category averaged across relations and villages.

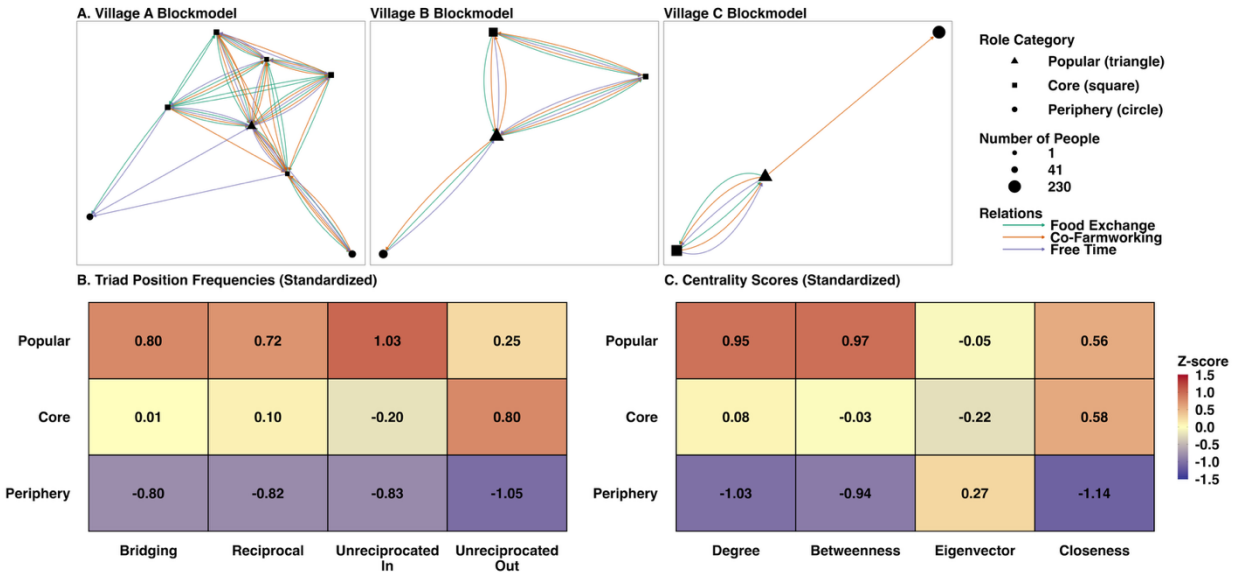


Table 2. Participant demographic characteristics by role category.

Demographic Characteristics	Periphery (n = 445)	Core (n = 476)	Popular (n = 376)	All Roles (n = 1,297)
Age, median (IQR)	32 (22-46)	33 (24-48)	35 (26-49)	33 (24-48)
Unknown	1	0	1	2
Gender, n (%)				
Women	209 (47%)	220 (46%)	164 (44%)	593 (46%)
Men	236 (53%)	256 (54%)	212 (56%)	704 (54%)
Education, n (%)				
None	22 (5%)	18 (3.8%)	9 (2.4%)	49 (3.8%)
Primary	193 (44%)	217 (46%)	170 (45%)	580 (45%)
Secondary	169 (38%)	178 (38%)	152 (41%)	499 (39%)
Higher	55 (13%)	59 (13%)	44 (12%)	158 (12%)
Unknown	6	4	1	11 (0.8%)
Occupation, n (%)				
Farmer	389 (88%)	427 (90%)	350 (93%)	1,166 (90%)
Non-Farmer[1]	55 (12%)	47 (9.9%)	26 (6.9%)	128 (9.9%)
Unknown	1	2	0	3 (0.2%)
House Lifestyle Index, median (IQR)	-0.34 (-1.55-1.84)	0.63 (-1.55-1.84)	0.63 (-1.55-1.84)	-1.55 (-1.55-1.84)
Durable Goods Owned, median (IQR)	1 (0-2)	1 (0-2)	1 (1-2)	1 (0-2)
Household Size, median (IQR)	4 (3-5)	4 (3-5)	4 (3-5)	4 (3-5)

[1]Non-farming occupations included carpenter, taxi driver, laborer, shopkeeper, teacher, student, and unemployed.

Individuals in the Popular role category occupied reciprocal and unreciprocated-in triad positions at a higher frequency than the other role categories, indicating that they rarely named someone else without also being named themselves. They also occupied more bridging triad positions and had higher betweenness centrality than the other role categories. Together with their relatively high degree centrality, these patterns suggest that members of the Popular role category were top of mind when participants were asked with whom they spend free time, exchange food, and participate in co-farmworking. These patterns also suggest that these participants' connections spanned social groups within the villages.

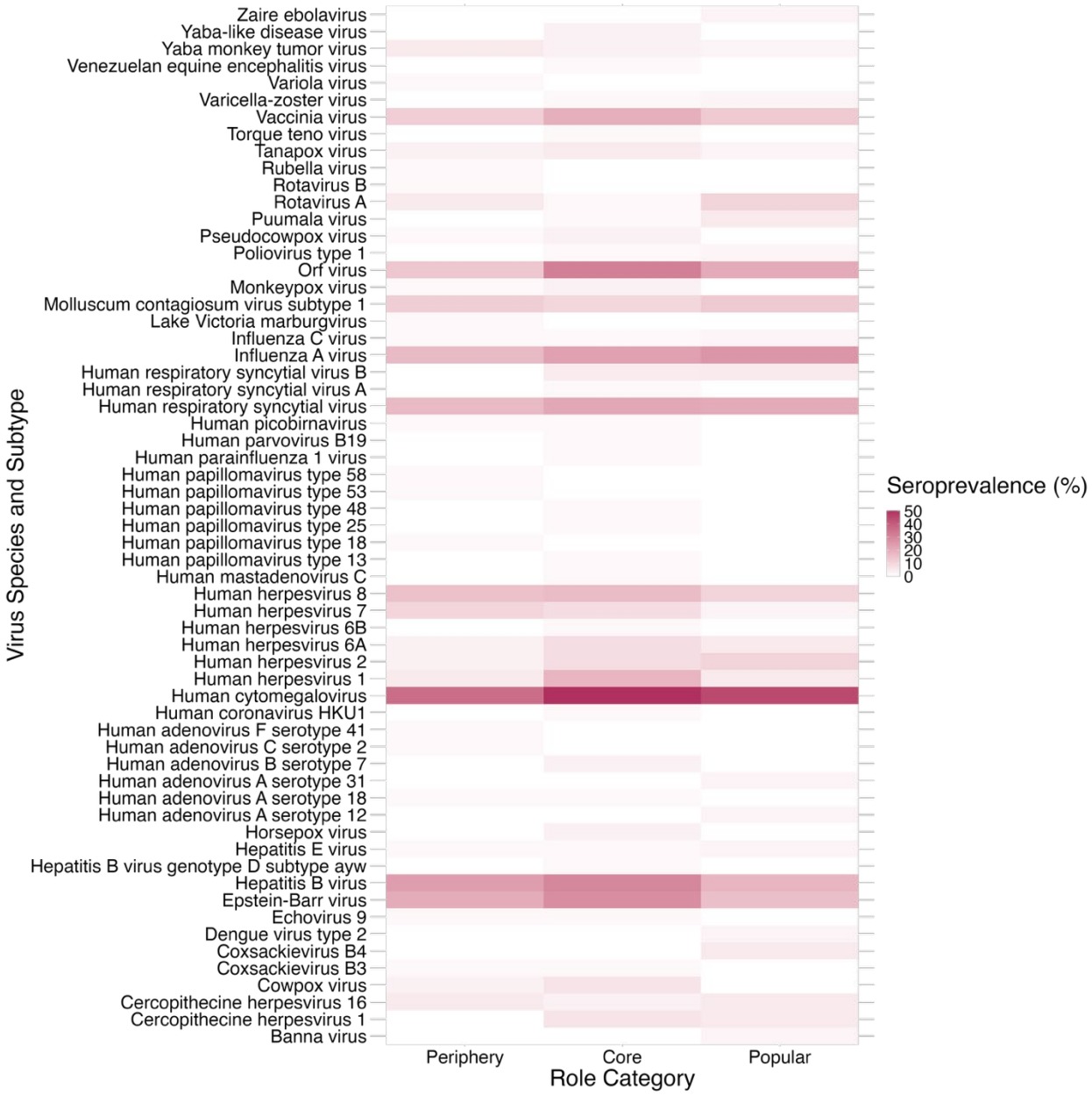
Members of the Core role category occupied bridging, reciprocal, and unreciprocated-in triad positions at a comparatively lower frequency. Instead, they occupied unreciprocated-out triad positions at a much higher frequency. This pattern indicates that they named many other people in the network, but few people named them in return. They did have closeness centrality scores comparable to members of the Popular role category, which suggests they were only a short path away from all other nodes in the network.

In general, members of the Periphery role category were represented in each type of triad position at a lower frequency than members of both the Core and Popular role categories. They also had much lower degree, betweenness, and closeness centrality scores, but they did have relatively high eigenvector centrality, particularly for the food and co-farmworking networks. This pattern suggests that they engaged in exchange with a few highly influential individuals in each village, which likely brought them into the networks despite their minimal engagement with most members of a village. Standardized triad position frequencies and centrality scores for each village, relation, and role category are shown in **Figures S1-S6**.

Association Between Role Categories and Viral Exposure

For a subset of participants from Village B and Village C with VirScan data (n = 156), we assessed the relationship between role categories and past viral exposure. Of the 337 virus species and subtypes in the VirScan library, antibodies to 62 (18%) virus species and subtypes were identified in this sample **(Figure 3)**. Ten of the 62 virus species and subtypes (16%) had a seroprevalence greater than 10%.

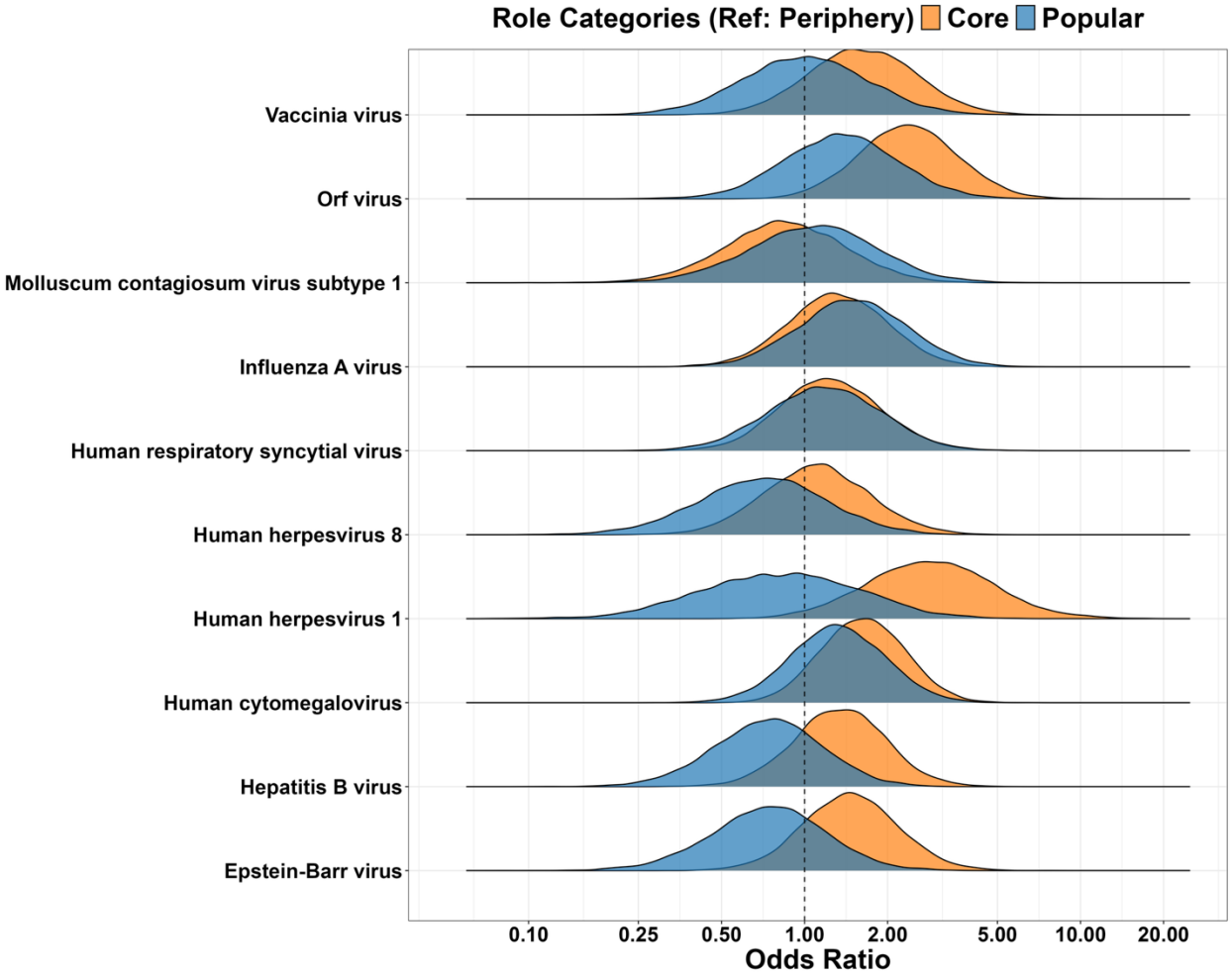
Figure 3. Seroprevalence by role category for the 62 virus species and subtypes identified in the sample. Darker red tiles indicate higher seroprevalence.



Bivariate Bayesian logistic regression models suggest that individuals in the Popular and/or Core role categories had higher odds of infection than individuals in the Periphery role category with all but one of the ten viruses that exhibited seroprevalence >10% (the exception was molluscum contagiosum virus subtype 1). However, this pattern should be interpreted with

relatively low confidence, as 80% credible intervals spanned zero in all models. Individuals in the Core role category had a greater effect size than individuals in the Popular role category for seven of the ten viruses **(Figure 4)**. This difference in effect size was most pronounced for herpesviruses, particularly herpes simplex virus 1 (HSV-1) and Epstein-Barr virus (EBV).

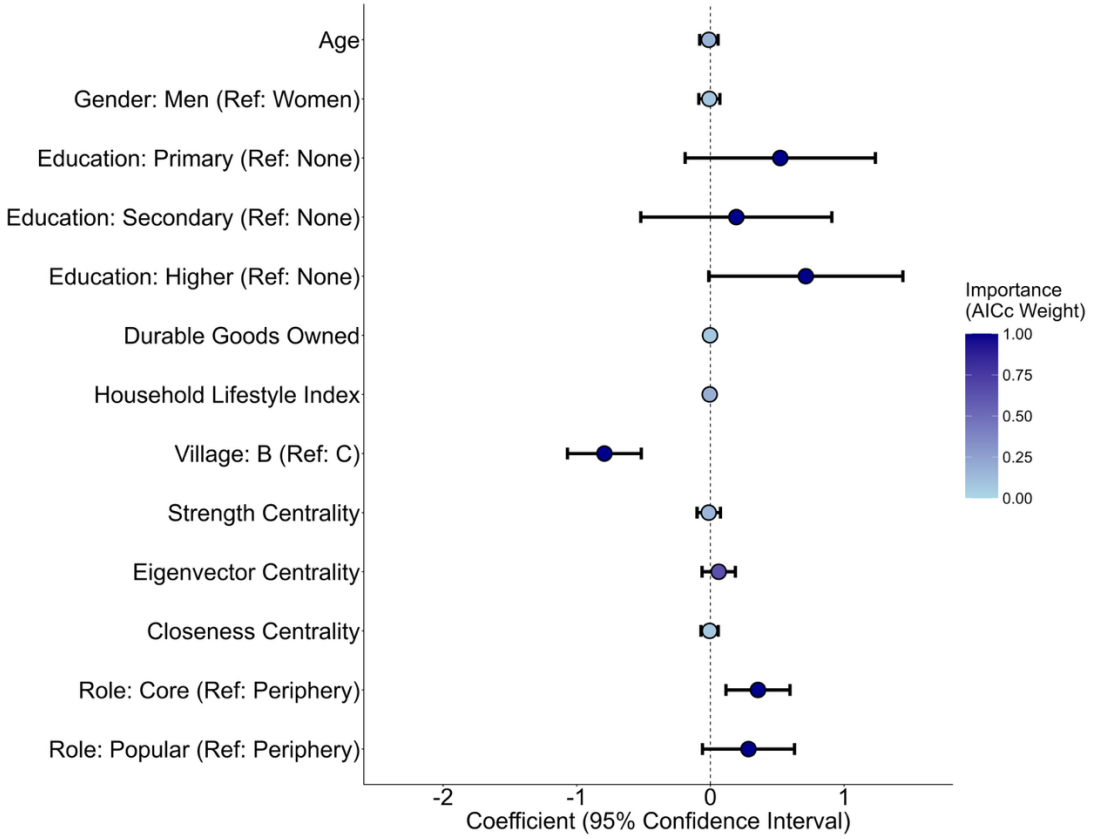
Figure 4. Odds of viral exposure by role category. Posterior distribution (80% credible interval) of odds of exposure by role category for virus species with greater than 10% seroprevalence.



Role category, education, and village were the most important predictors of viral species and subtype richness (summed AICc weight = 1) **(Figure 5)**. Membership in both the Core role category (model averaged β [95% CI]: 0.36 [0.12-0.60]) and Popular role category (model averaged β [95% CI]: 0.29 [-0.06-0.63]) was associated with higher richness than membership in

the Periphery role category. Residence in Village B was associated with lower richness than residence in Village C (model averaged β [95% CI]: -0.79 [-1.07--0.52]). Eigenvector centrality was a relatively important predictor of richness (summed AICc weight = 0.64), but all four individual centrality measures had a coefficient close to zero and were less important than the role categories. Strength (summed AICc weight = 0.15) and closeness centrality (summed AICc weight = 0.07) were both of minimal importance for predicting richness, and betweenness centrality was not included in any of the eleven component models used in the model averaging **(Table S1)**. Age, gender, and the two measures of market-based wealth all had a summed AICc weights less than 0.2, and the direction of each relationship was unclear.

Figure 5. Association between predictor variables and virus species and subtype richness. Points and error bars represent β coefficients and 95% confidence intervals from model averaging, and darker blue points reflect variables that were more important (higher AICc weight) for predicting viral species richness.



Discussion

Measures of node centrality can provide important information for predicting which individuals in a social network are most likely to be exposed to a parasite and have the greatest spreading potential [6-12]. In this study, we demonstrated that the utility of single centrality measures can be enhanced by clustering individuals by their similarity in both centrality scores and triad positions across multiple networks or types of relations. We used this approach – role analysis via equivalence-based blockmodeling – to identify social-epidemiological roles associated with viral exposure in three rural villages in northeast Madagascar. We identified three role categories (Popular, Core, and Peripheral) that reflected general patterns in how people tended to spend their free time with others and exchange food and farmwork. We hypothesized that individuals who had high centrality scores and received and sent many ties would be exposed to the most viruses and found partial support for this hypothesis. We found that membership in the Core role category – those who sent many ties with few reciprocated – had the strongest positive association with viral exposure. Members of the Core role category also had relatively high closeness centrality, indicating that they were a short distance from everyone else in the network. As expected, we found that social-epidemiological roles were more predictive of virial exposure than single measures of centrality, suggesting that role analysis via equivalence-based blockmodeling can capture important information about how people interact across relations in ways that are meaningful for understanding infectious disease transmission dynamics.

The role categories identified in our analysis provided distinct insights into viral transmission dynamics in a rural village setting. An individual's vector of triad positions provides a summary of how they interact across different domains of social life [36], with the domains considered here being how individuals socialize in their free time, how they help and are helped

by others when food is scarce, and how they engage in farmwork with others. Given that greater social activity provides more opportunities for infectious disease exposure, we hypothesized that people who engaged in reciprocal relationships across these domains would be more socially active and thus have more opportunities for viral exposure. In contrast, we found that membership in the Core role category – representing individuals who sent many ties across relations that were rarely reciprocated and participated in few reciprocal closed triads – was associated with greater viral richness.

This finding suggests a potential alternative hypothesis. Sociality has both costs and benefits for infectious disease risk, with the primary cost being increased probability of exposure [63]. Low social status and chronic psychosocial stress can also weaken the immune system [64, 65]. Individuals who reported more unreciprocated relationships may thus experience stress-induced immunosuppression, and their high social activity – sending many ties – combined with lower social status – relative to members of the Popular role category – may make them particularly susceptible to infection. Our bivariate analysis of the relationship between role categories and seropositivity for specific viruses also lends support to this alternative hypothesis. For example, the differentiation in effect size between the Core and Popular role categories is largest for two herpesviruses (HSV-1 and EBV). Prior studies linking social stressors to circulating EBV and HSV-1 antibody levels have highlighted the importance of psychosocial stress in affecting susceptibility to infection and reactivation [66-71]. Further work is needed to more directly identify the mechanisms by which the social-epidemiological roles identified in this study may affect susceptibility to infectious diseases through stress pathways.

This study also revealed important differences in how measures of centrality relate to viral exposure in a rural village setting. For example, eigenvector centrality is a measure of how

well connected a node is to highly connected nodes, and prior studies have found that this measure of node importance is helpful for understanding parasite transmission [16, 72]. However, members of the Periphery role category had the highest eigenvector centrality scores relative to members of the Popular and Core role categories. In this case, individuals on the social periphery were likely brought into the network by their connection with popular individuals, but their low overall social integration likely resulted in fewer opportunities for viral exposure. In contrast, members of the Core role category had closeness centrality scores comparable to members of the Popular role category, suggesting that they had a short social distance to all other members of the network. These findings suggest that this network-wide reach facilitates exposure to a wide range of viruses in a rural village setting.

Each role category had distinct patterns of centrality scores, but the role categories themselves were more important for predicting viral species and subtype richness than single measures of centrality alone. This finding suggests that role analysis via equivalence-based blockmodeling allows for a more complete representation of the social processes that underlie infectious disease transmission dynamics. For example, members of the Core role category had relatively high closeness centrality scores, and membership in this role category was associated with greater viral species and subtype richness. However, closeness centrality alone performed relatively poorly at predicting richness. Our findings suggest that the most at-risk individuals in a network might not simply be those who are most central but rather those who fit a general profile of being both central (increasing the probability of exposure) and lower social status (increasing susceptibility to infection once exposed).

In network epidemiology, approaches to identify the most at-risk individuals with high spreading potential often emphasize measures of centrality [6]. Our findings suggest that

additional positional network features like an individual's level of engagement in reciprocal relationships provide important information for understanding disease transmission and targeting interventions. The real-world application of this approach is limited by the need for network data, which are often limited in low-resource settings [7]. However, general principles can emerge that guide interventions, and in cases where network data are available to compute centrality scores, no additional data are needed to perform regular equivalence blockmodeling. This approach may be particularly useful in disease ecology, where researchers are increasingly collecting data on multiple types of relations [26-28, 30]. Our study demonstrates how role analysis can be used in combination with these multidimensional datasets to uncover complex transmission dynamics and generate new transmission-relevant hypotheses.

Our study had several limitations. First, although we had comprehensive network data for all three study villages, our analysis of viral exposure was limited to a subset of participants from two of the three villages. This data limitation precluded an analysis of viral exposure in Village A, which had the most complex set of roles. Sample size limitations also required us to analyze roles at a higher level (role category) for Villages B and C, but the more granular roles could be used in studies with more data per role. Second, PhIP-Seq VirScan is a powerful tool for characterizing exposure to a broad set of viruses [43, 44]. However, determining seropositivity for a specific virus requires confirmatory analyses through more traditional methods like enzyme-linked immunosorbent assays (ELISAs). Third, we examined the utility of role analysis for predicting viral exposure, but its utility for other types of parasites might differ. Future work should explore applications across broad categories of parasites, from viruses to helminths.

Conclusion

Infectious diseases often spread through multiple pathways, and analytical techniques for identifying high risk individuals in transmission networks can be limited by an overemphasis on single measures of centrality. We demonstrated the utility of role analysis via equivalence-based blockmodeling for identifying social-epidemiological roles that are associated distinct patterns of exposure to viruses in a rural village setting. The roles we identified performed better at predicting viral exposure than single measures of centrality alone, and the analysis provided insight into the coupled dynamics of susceptibility and exposure in viral transmission. Future studies in disease ecology and network epidemiology could benefit from using role analysis via equivalence-based blockmodeling to summarize increasingly high dimensional network data.

Acknowledgements

We thank the Duke Lemur Center-SAVA Conservation for logistical support and the Malagasy Ethics Panel for permission to conduct the research. We greatly appreciate the three study communities for their participation and hospitality. Funding was provided by the joint NIH-NSF-NIFA Ecology and Evolution of Infectious Disease Program (R01-TW011493-01, DEB-2308460), an NSF Doctoral Dissertation Research Improvement Grant (BCS-2341234), an NSF Predictive Intelligence for Pandemic Prevention (PIPP) Phase I grant (SBE-2200047), and a Duke Provost's Collaboratory.

References

[1] Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E. & Getz, W.M. 2005 Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355-359. (doi:10.1038/nature04153).

[2] Shen, Z., Ning, F., Zhou, W., He, X., Lin, C., Chin, D.P., Zhu, Z. & Schuchat, A. 2004 Superspreading SARS events, Beijing, 2003. *Emerg Infect Dis* **10**, 256-260. (doi:10.3201/eid1002.030732).

[3] Althouse, B.M., Wenger, E.A., Miller, J.C., Scarpino, S.V., Allard, A., Hebert-Dufresne, L. & Hu, H. 2020 Superspreading events in the transmission dynamics of SARS-CoV-2: Opportunities for interventions and control. *PLoS Biol* **18**, e3000897. (doi:10.1371/journal.pbio.3000897).

[4] De Serres, G., Markowski, F., Toth, E., Landry, M., Auger, D., Mercier, M., Belanger, P., Turmel, B., Arruda, H., Boulianne, N., et al. 2013 Largest measles epidemic in North America in a decade--Quebec, Canada, 2011: contribution of susceptibility, serendipity, and superspreading events. *J Infect Dis* **207**, 990-998. (doi:10.1093/infdis/jis923).

[5] Lau, M.S., Dalziel, B.D., Funk, S., McClelland, A., Tiffany, A., Riley, S., Metcalf, C.J. & Grenfell, B.T. 2017 Spatial and temporal dynamics of superspreading events in the 2014-2015 West Africa Ebola epidemic. *Proc Natl Acad Sci U S A* **114**, 2337-2342. (doi:10.1073/pnas.1614595114).

[6] Christley, R.M., Pinchbeck, G.L., Bowers, R.G., Clancy, D., French, N.P., Bennett, R. & Turner, J. 2005 Infection in social networks: using network analysis to identify high-risk individuals. *Am J Epidemiol* **162**, 1024-1031. (doi:10.1093/aje/kwi308).

[7] Evans, M.V., Ramiadantsoa, T., Kauffman, K., Moody, J., Nunn, C.L., Rabezara, J.Y., Raharimalala, P., Randriamoria, T.M., Soarimalala, V., Titcomb, G., et al. 2023

Sociodemographic Variables Can Guide Prioritized Testing Strategies for Epidemic Control in Resource-Limited Contexts. *J Infect Dis* **228**, 1189-1197. (doi:10.1093/infdis/jiad076).

[8] Paull, S.H., Song, S., McClure, K.M., Sackett, L.C., Kilpatrick, A.M. & Johnson, P.T. 2012 From superspreaders to disease hotspots: linking transmission across hosts and space. *Front Ecol Environ* **10**, 75-82. (doi:10.1890/110111).

[9] Rushmore, J., Caillaud, D., Matamba, L., Stumpf, R.M., Borgatti, S.P. & Altizer, S. 2013 Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk. *J Anim Ecol* **82**, 976-986. (doi:10.1111/1365-2656.12088).

[10] Streicker, D.G., Fenton, A. & Pedersen, A.B. 2013 Differential sources of host species heterogeneity influence the transmission and control of multihost parasites. *Ecol Lett* **16**, 975-984. (doi:10.1111/ele.12122).

[11] VanderWaal, K.L., Atwill, E.R., Isbell, L.A. & McCowan, B. 2014 Quantifying microbe transmission networks for wild and domestic ungulates in Kenya. *Biological Conservation* **169**, 136-146. (doi:10.1016/j.biocon.2013.11.008).

[12] Weeden, K.A. & Cornwell, B. 2020 The Small-World Network of College Classes: Implications for Epidemic Spread on a University Campus. *Sociological Science* **7**, 222-241.

[13] Bell, D.C., Atkinson, J.S. & Carlson, J.W. 1999 Centrality measures for disease transmission networks. *Social Networks* **21**, 1-21. (doi:10.1016/s0378-8733(98)00010-0).

[14] Griffin, R.H. & Nunn, C.L. 2011 Community structure and the spread of infectious disease in primate social networks. *Evolutionary Ecology* **26**, 779-800. (doi:10.1007/s10682-011-9526-2).

[15] Pierron, M., Sueur, C., Shimada, M., MacIntosh, A.J.J. & Romano, V. 2024 Epidemiological Consequences of Individual Centrality on Wild Chimpanzees. *Am J Primatol* **86**, e23682. (doi:10.1002/ajp.23682).

[16] Romano, V., Duboscq, J., Sarabian, C., Thomas, E., Sueur, C. & MacIntosh, A.J. 2016 Modeling infection transmission in primate networks to predict centrality-based risk. *Am J Primatol* **78**, 767-779. (doi:10.1002/ajp.22542).

[17] Shridhar, S.V., Alexander, M. & Christakis, N.A. 2022 Characterizing super-spreaders using population-level weighted social networks in rural communities. *Philos Trans A Math Phys Eng Sci* **380**, 20210123. (doi:10.1098/rsta.2021.0123).

[18] Singer, B.J., Thompson, R.N. & Bonsall, M.B. 2022 Evaluating strategies for spatial allocation of vaccines based on risk and centrality. *J R Soc Interface* **19**, 20210709. (doi:10.1098/rsif.2021.0709).

[19] Duboscq, J., Romano, V., Sueur, C. & MacIntosh, A.J. 2016 Network centrality and seasonality interact to predict lice load in a social primate. *Sci Rep* **6**, 22095. (doi:10.1038/srep22095).

[20] Marques-Sanchez, P., Martinez-Fernandez, M.C., Leiros-Rodriguez, R., Rodriguez-Nogueira, O., Fernandez-Martinez, E. & Benitez-Andrades, J.A. 2023 Leadership and contagion by COVID-19 among residence hall students: A social network analysis approach. *Soc Networks* **73**, 80-88. (doi:10.1016/j.socnet.2023.01.001).

[21] Rothenberg, R.B., Potterat, J.J., Woodhouse, D.E., Darrow, W.W., Muth, S.Q. & Klovdahl, A.S. 1995 Choosing a centrality measure: Epidemiologic correlates in the Colorado Springs study of social networks. *Social Networks* **17**, 273-297. (doi:10.1016/0378-8733(95)00267-r).

[22] Sandel, A.A., Rushmore, J., Negrey, J.D., Mitani, J.C., Lyons, D.M. & Caillaud, D. 2020 Social Network Predicts Exposure to Respiratory Infection in a Wild Chimpanzee Group. *Ecohealth* **17**, 437-448. (doi:10.1007/s10393-020-01507-7).

[23] VanderWaal, K.L., Atwill, E.R., Isbell, L.A. & McCowan, B. 2014 Linking social and pathogen transmission networks using microbial genetics in giraffe (Giraffa camelopardalis). *J Anim Ecol* **83**, 406-414. (doi:10.1111/1365-2656.12137).

[24] Browne, A., Butts, D., Jaramillo-Rodriguez, E., Parikh, N., Fairchild, G., Needell, Z., Poliziani, C., Wenzel, T., Germann, T.C. & Del Valle, S. 2024 Evaluating disease surveillance strategies for early outbreak detection in contact networks with varying community structure. *Social Networks* **79**, 122-132. (doi:10.1016/j.socnet.2024.06.003).

[25] Adams, J., Moody, J. & Morris, M. 2013 Sex, drugs, and race: how behaviors differentially contribute to the sexually transmitted infection risk network structure. *Am J Public Health* **103**, 322-329. (doi:10.2105/AJPH.2012.300908).

[26] Barrett, T.M., Titcomb, G.C., Janko, M.M., Pender, M., Kauffman, K., Solis, A., Randriamoria, M.T., Young, H.S., Mucha, P.J., Moody, J., et al. 2024 Disentangling social, environmental, and zoonotic transmission pathways of a gastrointestinal protozoan (Blastocystis spp.) in northeast Madagascar. *Am J Biol Anthropol* **185**, e25030. (doi:10.1002/ajpa.25030).

[27] Blyton, M.D., Banks, S.C., Peakall, R., Lindenmayer, D.B. & Gordon, D.M. 2014 Not all types of host contacts are equal when it comes to E. coli transmission. *Ecol Lett* **17**, 970-978. (doi:10.1111/ele.12300).

[28] Bull, C.M., Godfrey, S.S. & Gordon, D.M. 2012 Social networks and the spread of Salmonella in a sleepy lizard population. *Mol Ecol* **21**, 4386-4392. (doi:10.1111/j.1365-294X.2012.05653.x).

[29] Emch, M., Root, E.D., Giebultowicz, S., Ali, M., Perez-Heydrich, C. & Yunus, M. 2012 Integration of Spatial and Social Network Analysis in Disease Transmission Studies. *Ann Assoc Am Geogr* **105**, 1004-1015. (doi:10.1080/00045608.2012.671129).

[30] Rimbach, R., Bisanzio, D., Galvis, N., Link, A., Di Fiore, A. & Gillespie, T.R. 2015 Brown spider monkeys (Ateles hybridus): a model for differentiating the role of social networks and physical contact on parasite transmission dynamics. *Philos Trans R Soc Lond B Biol Sci* **370**, 20140110. (doi:10.1098/rstb.2014.0110).

[31] Nande, A., Adlam, B., Sheen, J., Levy, M.Z. & Hill, A.L. 2021 Dynamics of COVID-19 under social distancing measures are driven by transmission network structure. *PLoS Comput Biol* **17**, e1008684. (doi:10.1371/journal.pcbi.1008684).

[32] Colman, E., Colizza, V., Hanks, E.M., Hughes, D.P. & Bansal, S. 2021 Social fluidity mobilizes contagion in human and animal populations. *Elife* **10**, e62177. (doi:10.7554/eLife.62177).

[33] Boorman, S.A. & White, H.C. 1976 Social Structure from Multiple Networks. II. Role Structures. *American Journal of Sociology* **81**, 1384-1446. (doi:10.1086/226228).

[34] Burt, R.S. 1990 Detecting role equivalence. *Social Networks* **12**, 83-97. (doi:10.1016/0378-8733(90)90023-3).

[35] Doreian, P., Batagelj, V. & Ferligoj, A. 2020 *Advances in Network Clustering and Blockmodeling*.

[36] Rawlings, C.M., Smith, J.A., Moody, J. & McFarland, D.A. 2023 *Network Analysis: Integrating Social Network Theory, Method, and Application with R*.

[37] Bearman, P. 1997 Generalized Exchange. *American Journal of Sociology* **102**, 1383-1415. (doi:10.1086/231087).

[38] McFarland, D.A., Moody, J., Diehl, D., Smith, J.A. & Thomas, R.J. 2014 Network Ecology and Adolescent Social Structure. *Am Sociol Rev* **79**, 1088-1121. (doi:10.1177/0003122414554001).

[39] Moody, J. & Mucha, P.J. 2013 Portrait of Political Party Polarization. *Network Science* **1**, 119-121. (doi:10.1017/nws.2012.3).

[40] Padgett, J.F. & Ansell, C.K. 1993 Robust Action and the Rise of the Medici, 1400-1434. *American Journal of Sociology* **98**, 1259-1319. (doi:10.1086/230190).

[41] Luczkovich, J.J., Borgatti, S.P., Johnson, J.C. & Everett, M.G. 2003 Defining and measuring trophic role similarity in food webs using regular equivalence. *J Theor Biol* **220**, 303-321. (doi:10.1006/jtbi.2003.3147).

[42] Pearl, M.C. & Schulman, S.R. 1983 Techniques for the Analysis of Social Structure in Animal Societies. In *Advances in the Study of Behavior* (pp. 107-146.

[43] Tiu, C.K., Zhu, F., Wang, L.F. & de Alwis, R. 2022 Phage ImmunoPrecipitation Sequencing (PhIP-Seq): The Promise of High Throughput Serology. *Pathogens* **11**, Article 5. (doi:10.3390/pathogens11050568).

[44] Xu, G.J., Kula, T., Xu, Q., Li, M.Z., Vernon, S.D., Ndung'u, T., Ruxrungtham, K., Sanchez, J., Brander, C., Chung, R.T., et al. 2015 Viral immunology. Comprehensive serological profiling of human populations using a synthetic human virome. *Science* **348**, aaa0698. (doi:10.1126/science.aaa0698).

[45] Harrison, G.F., Sanz, J., Boulais, J., Mina, M.J., Grenier, J.C., Leng, Y., Dumaine, A., Yotova, V., Bergey, C.M., Nsobya, S.L., et al. 2019 Natural selection contributed to immunological differences between hunter-gatherers and agriculturalists. *Nat Ecol Evol* **3**, 1253-1264. (doi:10.1038/s41559-019-0947-6).

[46] Doehne, M., McFarland, D.A. & Moody, J. 2024 Network ecology: Tie fitness in social context(s). *Soc Networks* **76**, 174-190. (doi:10.1016/j.socnet.2023.09.005).

[47] Fuhse, J.A. 2009 The Meaning Structure of Social Networks. *Sociological Theory* **27**, 51-73.

[48] Naderifar, M., Goli, H. & Ghaljaie, F. 2017 Snowball Sampling: A Purposeful Method of Sampling in Qualitative Research. *Strides in Development of Medical Education* **14**, e67670.

[49] Bindon, J.R., Knight, A., Dressler, W.W. & Crews, D.E. 1997 Social context and psychosocial influences on blood pressure among American Samoans. *American Journal of Physical Anthropology* **103**, 7-18. (doi:10.1002/(sici)1096-8644(199705)103:1<7::Aid-ajpa2>3.0.Co;2-u).

[50] Henrich, J. 1997 Market Incorporation, Agricultural Change, and Sustainability Among the Machiguenga Indians of the Peruvian Amazon. *Human Ecology* **25**, 319-351. (doi:10.1023/a:1021982324396).

[51] Liebert, M.A., Snodgrass, J.J., Madimenos, F.C., Cepon, T.J., Blackwell, A.D. & Sugiyama, L.S. 2013 Implications of market integration for cardiovascular and metabolic health among an indigenous Amazonian Ecuadorian population. *Ann Hum Biol* **40**, 228-242. (doi:10.3109/03014460.2012.759621).

[52] Kauffman, K., Werner, C.S., Titcomb, G., Pender, M., Rabezara, J.Y., Herrera, J.P., Shapiro, J.T., Solis, A., Soarimalala, V., Tortosa, P., et al. 2022 Comparing transmission potential networks based on social network surveys, close contacts and environmental overlap in rural Madagascar. *J R Soc Interface* **19**, 20210690. (doi:10.1098/rsif.2021.0690).

[53] DeSisto, C.M.M., Binder, R.A., Kauffman, K., Barrett, T.M., Pender, M., Kramer, R.A., Soarimalala, V., Rabezara, J.Y., Rahary, P., Moody, J., et al. Under Review Spreading potential in disease relevant networks: predicting centralities in rural northeast Madagascar.

[54] Tiu, C.K., Chia, W.N., Anderson, D.E., Chee, S.P., Wang, L.F. & Siak, J. 2024 Pan-viral Antibody Repertoire of Aqueous Humor in Cytomegalovirus Uveitis. *Am J Ophthalmol* **266**, 218-226. (doi:10.1016/j.ajo.2024.05.004).

[55] Mohan, D., Wansley, D.L., Sie, B.M., Noon, M.S., Baer, A.N., Laserson, U. & Larman, H.B. 2018 PhIP-Seq characterization of serum antibodies using oligonucleotide-encoded peptidomes. *Nat Protoc* **13**, 1958-1978. (doi:10.1038/s41596-018-0025-6).

[56] 2025 R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

[57] Wolff, T., Morgan, J.H., Varela, G., Lele, K., Bhojani, E., Heraty, E., Pasquale, D., Mucha, P. & Moody, J. 2024 ideanet: Integrating Data Exchange and Analysis for Networks ('ideanet'). (

[58] White, D.R. & Reitz, K.P. 1983 Graph and semigroup homomorphisms on networks of relations. *Social Networks* **5**, 193-234. (doi:10.1016/0378-8733(83)90025-4).

[59] Bürkner, P.-C., Gabry, J., Weber, S., Johnson, A., Modrak, M., Badr, H.S., Weber, F., Vehtari, A., Ben-Shachar, M.S., Rabel, H., et al. 2024 brms: Bayesian Regression Models using "Stan". (doi:10.32614/CRAN.package.brms).

[60] Bartoń, K. 2025 MuMIn: Multi-Model Inference. **R package version 1.48.11**. (doi:10.32614/CRAN.package.MuMIn).

[61] Burnham, K.P. & Anderson, D.R. 2004 Multimodel Inference. *Sociological Methods & Research* **33**, 261-304. (doi:10.1177/0049124104268644).

[62] Symonds, M.R.E. & Moussalli, A. 2010 A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. *Behavioral Ecology and Sociobiology* **65**, 13-21. (doi:10.1007/s00265-010-1037-6).

[63] Kappeler, P.M., Cremer, S. & Nunn, C.L. 2015 Sociality and health: impacts of sociality on disease susceptibility and transmission in animal and human societies. *Philos Trans R Soc Lond B Biol Sci* **370**, 20140116. (doi:10.1098/rstb.2014.0116).

[64] Segerstrom, S.C. & Miller, G.E. 2004 Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychol Bull* **130**, 601-630. (doi:10.1037/0033-2909.130.4.601).

[65] Cohen, S. 1999 Social status and susceptibility to respiratory infections. *Ann N Y Acad Sci* **896**, 246-253. (doi:10.1111/j.1749-6632.1999.tb08119.x).

[66] Kiecolt-Glaser, J.K., Glaser, R., Shuttleworth, E.C., Dyer, C.S., Ogrocki, P. & Speicher, C.E. 1987 Chronic stress and immunity in family caregivers of Alzheimer's disease victims. *Psychosom Med* **49**, 523-535. (doi:10.1097/00006842-198709000-00008).

[67] McDade, T.W. 2002 Status incongruity in Samoan youth: a biocultural analysis of culture change, stress, and immune function. *Med Anthropol Q* **16**, 123-150. (doi:10.1525/maq.2002.16.2.123).

[68] Sorensen, M.V., Snodgrass, J.J., Leonard, W.R., McDade, T.W., Tarskaya, L.A., Ivanov, K.I., Krivoshapkin, V.G. & Alekseev, V.P. 2009 Lifestyle incongruity, stress and immune function in indigenous Siberians: the health impacts of rapid social and economic change. *Am J Phys Anthropol* **138**, 62-69. (doi:10.1002/ajpa.20899).

[69] McClure, H.H., Snodgrass, J.J., Martinez, C.R., Jr., Eddy, J.M., Jimenez, R.A. & Isiordia, L.E. 2010 Discrimination, psychosocial stress, and health among Latin American immigrants in Oregon. *Am J Hum Biol* **22**, 421-423. (doi:10.1002/ajhb.21002).

[70] Panter-Brick, C., Eggerman, M., Mojadidi, A. & McDade, T.W. 2008 Social stressors, mental health, and physiological stress in an urban elite of young Afghans in Kabul. *Am J Hum Biol* **20**, 627-641. (doi:10.1002/ajhb.20797).

[71] Chida, Y. & Mao, X. 2009 Does psychosocial stress predict symptomatic herpes simplex virus recurrence? A meta-analytic investigation on prospective studies. *Brain Behav Immun* **23**, 917-925. (doi:10.1016/j.bbi.2009.04.009).

[72] Gomez, J.M., Nunn, C.L. & Verdu, M. 2013 Centrality in primate-parasite networks reveals the potential for the transmission of emerging infectious diseases to humans. *Proc Natl Acad Sci U S A* **110**, 7738-7741. (doi:10.1073/pnas.1220716110).