HHS Public Access

Author manuscript

Behaviour. Author manuscript; available in PMC 2019 November 01.

Published in final edited form as:

Behaviour. 2018 January ; 155(7-9): 759–791. doi:10.1163/1568539X-00003471.

Incorporating genomic methods into contact networks to reveal new insights into animal behavior and infectious disease dynamics

Marie L.J. Gilbertson, **Nicholas M. Fountain-Jones**, **Meggan E. Craft**

Department of Veterinary Population Medicine, University of Minnesota, Minneapolis, Minnesota 55455, USA

Abstract

Utilization of contact networks has provided opportunities for assessing the dynamic interplay between pathogen transmission and host behavior. Genomic techniques have, in their own right, provided new insight into complex questions in disease ecology, and the increasing accessibility of genomic approaches means more researchers may seek out these tools. The integration of network and genomic approaches provides opportunities to examine the interaction between behavior and pathogen transmission in new ways and with greater resolution. While a number of studies have begun to incorporate both contact network and genomic approaches, a great deal of work has yet to be done to better integrate these techniques. In this review, we give a broad overview of how network and genomic approaches have each been used to address questions regarding the interaction of social behavior and infectious disease, and then discuss current work and future horizons for the merging of these techniques.

Keywords

Social network; network modeling; phylodynamics; genetics; strain sharing; transmission tree; wildlife

Interplay between behavior and infectious disease

The study and utilization of mathematical models has been revolutionary to the field of disease ecology. Mathematical models can be used to examine mechanisms of pathogen transmission and maintenance, and to make predictions about pathogen spread and management (Lloyd-Smith et al., 2009; Keeling & Rohani, 2011). Initially, models of infectious disease transmission assumed no contact heterogeneity among hosts (Anderson & May, 1991; Keeling & Rohani, 2011). This simplicity can be useful for creating transparent, generalizable models, but may not be appropriate for all systems, as host contact heterogeneity can have significant impacts on disease dynamics (Keeling & Eames, 2005; Meyers, 2007; White et al., 2017). For example, “superspreaders” can have important effects on epidemic outcomes, with their presence resulting in disease outbreaks that are less frequent but more severe than predicted by homogeneous mixing of populations (Lloyd-Smith et al., 2005; Lloyd-Smith et al., 2009). By incorporating contact heterogeneity into model assumptions, contact or “social” networks, applied to disease models, can be

important for conveying a more informative picture of pathogen transmission dynamics (Craft & Caillaud, 2011; Godfrey, 2013; Craft, 2015; White et al., 2017).

Contact networks represent connections between individuals or groups of individuals based on a variety of definitions of "contact," and these contacts are frequently influenced by social behavior. Network models can therefore require more intensive data collection compared to other techniques to model disease spread, such as compartmental models, but can also provide unique insight into disease dynamics thanks to their incorporation of contact heterogeneity. While social network models have been much used in studies of infectious disease in human and livestock systems, they have experienced delayed but growing utilization in wildlife systems (Martínez-López et al., 2009; Godfrey, 2013).

Animal behavior influences contact heterogeneity, more specifically through heterogeneity in rates and patterns of social interaction (VanderWaal & Ezenwa, 2016). Some of the behavioral or social processes that impact contact heterogeneity include individual behavioral phenotypes (Natoli et al., 2005; Dizney & Dearing, 2013), the social structure of a population (Nunn et al., 2015; Sah et al., 2017), seasonal or temporal behavior changes (Chen et al., 2014), and behavioral responses to anthropogenic influences (Gottdenker et al., 2014; Becker et al., 2015). Behavior is therefore important for pathogen transmission, and particularly so for directly transmitted pathogens, where environmental contamination or vectors do not dilute the effect of animal interactions (Godfrey, 2013; VanderWaal & Ezenwa, 2016; White et al., 2017).

While network approaches have made much progress in shaping our understanding of the complicated, dynamic interplay between behavior and infectious disease, they also reveal new, more nuanced questions, demonstrating the need for additional data and methods to better illuminate the processes at work within disease ecology. Genomic methods, applied to host or pathogen, are becoming increasingly recognized for their power to provide new insights into disease ecology (Archie et al., 2009). Genetic and genomic tools can allow us to study and infer pathogen transmission (Metzker et al., 2002; Kao et al., 2014), reconstruct epidemics (Biek et al., 2007; Bird et al., 2007; Sharp & Hahn, 2010), and reveal landscape and environmental factors important to transmission (Blanchong et al., 2007; Archie et al., 2009). Genomic methods are also becoming more available and affordable, putting them within the reach of more researchers. Within the scope of wildlife studies, genomic tools may also provide important opportunities to maximize the use of non-invasive sampling techniques. Non-invasive techniques are expanding in their range and utility, and may be particularly important for studying populations of endangered or at-risk species, where capture and handling of individuals carries increased risk to overall population health (de Carvalho Ferreira et al., 2014; Hoffmann et al., 2016; Smiley Evans et al., 2016).

To improve predictions and accuracy, network models should be informed by as much of the available data as possible (Welch et al., 2011). While genetic characterization of pathogens can suggest that cases of infection are closely related, many techniques (e.g. strain typing, microsatellite markers, etc.) provide only coarse resolution for inferring transmission (Kao et al., 2014). Whole genome sequencing (WGS), on the other hand, provides significantly more refined data for resolving transmission relationships (Kao et al., 2014). In the context

of investigating the dynamics between pathogen transmission and host behavior, it may seem counterintuitive to study pathogens in order to investigate host social structure, but the idea is not necessarily a new one (Welch et al., 2011), and doing so could provide important insights into behavior and disease ecology when integrated with contact networks. Given the increasing availability of genomic sequencing, as well as the expanding utility of non-invasive sampling in wildlife systems, there is an opportunity to incorporate pathogen genomic data into contact networks.

In this review, we discuss how contact networks and genomic methods provide insights into animal behavior and infectious disease transmission, and horizons for the integration of these methods in the future. More specifically, we discuss historical uses of social network analysis and network modeling, and how they are able to highlight the interplay between social behavior and infectious disease. We then describe how pathogen genetic and genomic techniques have been utilized to make inferences about host behavior and pathogen transmission. Finally, we explore the merging of contact network and genomic methods, and illustrate future directions of network modeling in light of the increasing availability of sequencing tools. Given that much can be gained by utilizing methods from other systems, this review will not limit examples to those from the animal or wildlife literature, but will draw from the human infectious disease literature as well.

Contact networks elucidate dynamics of behavior and disease ecology

The concepts and methodology behind the use of contact networks in animal systems have been well-reviewed elsewhere (Croft et al., 2008; Godfrey, 2013; Craft, 2015; White et al., 2017), so we will give a brief overview of these topics with a focus specifically on the contributions of contact networks to studying the interplay between animal behavior and infectious disease. For those readers that would like a brief “primer” on contact network approaches, additional background information can be found in Box 1.

The use of contact networks could be thought of as being composed of two main “branches,” social network analysis (SNA) and network modeling, though these methods are by no means mutually exclusive. Broadly, both SNA and network modeling involve building a social network based on behavioral or observational data. SNA then generally assesses network and node-level metrics that describe connectivity and modularity (Croft et al., 2008; Perkins et al., 2009), while network modeling often describes or simulates an outbreak of an infectious disease on the network. Network modeling is similar in concept to compartmental “Susceptible-Infectious-Recovered”-type (SIR) dynamic models, but instead of assuming homogeneous mixing, network modeling allows for the incorporation of heterogeneous contact patterns. Of course, our framework for describing contact network approaches is a somewhat artificial simplification, and many studies will not obviously fall into one category or the other. Network approaches could also be viewed as a spectrum ranging from describing networks for making observations and formulating hypotheses, to using statistical models to begin testing these hypotheses, through to simulation models of epidemics on networks to make predictions and test understanding. For the purposes of this review, we will discuss network approaches in the context of observational or descriptive approaches (SNA) and simulation-based approaches (network modeling), but with the acknowledgement

that network approaches do not always fall neatly into such a dichotomy, and often involve a continuum of approaches.

In the context of infectious disease research, SNA often compares a node's (individual or group of individuals) position in the network and node-level characteristics to the disease status of the node (Drewe, 2010; MacIntosh et al., 2012; Godfrey, 2013; VanderWaal et al., 2014). This approach can provide insight into how social behavior impacts pathogen transmission or prevalence, identify the importance of different behaviors for pathogen transmission, and elucidate the impact of disease prevention and management decisions. For example, in studying how social behavior impacts pathogen prevalence, MacIntosh et al. (2012) found that high social rank in Japanese macaques (*Macaca fuscata yakui*) was associated with greater nematode parasite species richness, as well as a higher probability and intensity of infection with a potentially pathogenic parasite. SNA can also be used to prioritize individuals to target for disease management interventions, or identify interventions that are less likely to be effective for disease control, based on population social structure (Porphyre et al., 2008; Rushmore et al., 2013).

Similarly, simulating epidemics of pathogens on social networks can accomplish a variety of objectives related to animal behavior and infectious disease. For example, it may not be feasible, or even ethical, to test pathogen spread and control hypotheses with real-world experiments, but modeling these experiments can provide important insights (Lloyd-Smith et al., 2009; Craft, 2015; White et al., 2017). Simulating epidemics on networks can thereby shed light on topics such as how disease is or is not maintained in populations (Craft et al., 2009), how population social structure impacts pathogen transmission (Craft et al., 2011), and what techniques or individuals to target for preventive or epidemic intervention measures (Rushmore et al., 2014; Pellis et al., 2015). For example, network models of pathogen transmission in chimpanzees found that targeting well-connected individuals allows for significant reductions in vaccinations required to prevent disease outbreaks (Rushmore et al., 2014).

Applications of contact networks are not without limitations, however. Observational data used to develop a social network may be limited (e.g. only able to collect behavioral observations during daytime hours, or only able to place GPS collars on a subset of the population), which can significantly impact conclusions (Croft et al., 2008). In addition, proxies of contact such as shared space use—as determined by telemetry data—may not always represent actual contacts; for example, avoidance behaviors could theoretically allow individuals in close proximity to avoid contacts relevant to pathogen transmission. While degree of home range overlap may correlate with increased contact rates between individuals, these assumptions are rarely tested and may vary seasonally or on an individual basis (Robert et al., 2012). There is therefore a need to improve or refine our detection or definition of transmission-relevant contacts, particularly when considering host species and/or pathogens in which timing of transmission events is not readily observed. In such cases, the increased resolution provided by molecular approaches applied to pathogens may be able to improve our understanding of disease transmission at a fine scale (Kao et al., 2014), and provide clues about microbial transmission in the context of host social structure (Bull et al., 2012; Blasse et al., 2013; Blyton et al., 2014).

Another challenge within the use of contact networks is the fact that social networks are generally dynamic and can vary through time (Craft & Caillaud, 2011; Rushmore et al., 2013; Eames et al., 2015; White et al., 2017). This variability highlights the importance of investigating social networks and infectious disease within appropriate temporal scales, according to the research question (Perkins et al., 2009; White et al., 2017). If the pathogen in question is infectious on the order of days to weeks, contact networks aggregated from year-long observations may not be appropriate and may overestimate pathogen spread. Over the course of a year, a social network is likely far more well-connected than it would be on the days to weeks-long scale on which the chosen pathogen may be operating. However, if the objective is to predict a "worst case scenario" of transmission, an aggregated network may be appropriate. Understanding the different temporal dynamics of host and pathogen within the context of a specific research question is necessary for appropriate study design and interpretation of model simulations. Because pathogen evolution may operate on different time scales from dynamic social networks, when pathogen phylogenies are assessed together with contact networks, they may shed light on the dynamics of transmission within populations (Vasylyeva et al., 2016). Thus, pathogen phylogenetics may serve a complementary function in balancing the complicated temporal dynamics of many social networks.

In summary, contact network approaches have proven useful for illuminating aspects of the relationship between animal behavior and pathogen transmission by describing how social structure, individual behaviors, and intra- and interspecies interactions contribute to disease transmission and maintenance in populations. Network approaches have also assisted in improving understanding of pathogen dynamics and management in populations, highlighting the importance of contact networks as a tool with real-world applications. An ongoing challenge for the utilization of contact networks is the detection of transmission-relevant contacts, for which genomic tools may provide new avenues for resolving these interactions.

Genomic tools to understand behavior and pathogen transmission

In this section, we will be reviewing how genetic and genomic-based methods have been used to provide insight into host social behavior and its impact on infectious disease dynamics. While genetics is, broadly, the study of individual genes, and genomics the study of whole genomes, for simplicity and consistency, we will refer to genetic and genomic-based techniques as "genomics" throughout the rest of this review, unless individual examples call for greater specification. We will use "phylogenetics," the study of evolutionary relationships based on genetics, to encompass both phylogenetic and phylogenomic techniques.

Population genomic tools are well-recognized for their ability to provide important information about host behavior and the consequences of animal behavior on pathogen transmission. For example, Pope et al. (2007) used population genomic tools to show increased badger dispersal after culling, and inferred important consequences of this increased movement for the spread of bovine tuberculosis. Population genomics have also been used to assess landscape barriers to host dispersal, and the importance of these

dispersal behaviors for disease transmission: for example, identifying rivers as semi-permeable barriers to host gene flow, and inferring the impact of rivers on subsequent animal movement behaviors and interactions in the context of pathogen transmission (Cullingham et al., 2009).

Pathogen phylogenetics can also provide insights into animal behavior and disease dynamics, and in some cases give greater clarity than host population genomics alone (Lee et al., 2012). With pathogen phylogenetics, researchers can better understand pathogen transmission dynamics and suggest patterns of host behavior and interactions (Biek et al., 2006; Lembo et al., 2008; Wheeler et al., 2010; Lee et al., 2012; Streicker et al., 2016; Fountain-Jones, Packer, et al., 2017). For example, Lembo et al. (2008) used pathogen phylogenetics to study cross-species transmission and reservoir dynamics of rabies in the Serengeti ecosystem. Their study found more within-species and less between-species transmission than would be expected from random mixing; these findings could be due to the spatial distribution of hosts, or from increased or preferential intraspecific contacts (Lembo et al., 2008). [These hypotheses could be further tested via a network modeling approach, thereby linking the phylogenetic and contact network approaches in the context of non-random contact patterns within and between species.] Pathogen phylogenetics can therefore provide valuable information about transmission dynamics and host contact patterns, including insights into understanding multi-host pathogen transmission, but these approaches could be enhanced even further by incorporating contact network techniques.

Importantly, investigation of pathogen phylogenetics can provide more detailed information that would be lost or otherwise inapparent by assessing host population genomics alone, particularly for directly transmitted pathogens. For example, phylogenetics of feline immunodeficiency virus in mountain lions revealed recent host demographic history that was not detectable with host population genomics (Biek et al., 2006). In addition, Bayesian phylogeographic approaches can yield important information about behavior and transmission dynamics by examining transmission over space and time (Biek et al., 2007; Lemey et al., 2009, 2010; De Maio et al., 2015), and between groups of hosts (Grad et al., 2014). Biek et al. (2007) used a phylogeographic approach to investigate spatial spread of rabies virus in raccoons in the Northeastern United States, finding, for example, that mountain ranges had a significant impact on rabies spread. These mountain ranges likely slowed spatial expansion due to the poor quality of raccoon habitat and reduced dispersal through these areas (Biek et al., 2007), demonstrating the impact of animal interactions with landscape features on the transmission of pathogens.

When available, the assessment of both pathogen and host genomics provides perhaps the most comprehensive information. For example, coupling host population genomic and pathogen phylogenetic methods have provided information about how solitary carnivores move and respond to landscape features like roads; where host phylogenies may suggest little to no movement across roads, pathogen phylogenies suggest these movements happen, but may be temporary or otherwise fail to result in host reproduction (Wheeler et al., 2010; Lee et al., 2012; Fountain-Jones, Craft, et al., 2017). These findings show that movement may be restricted across major landscape features, but can still occur and be adequate for pathogen transmission. Host and pathogen evolution, especially in cases of rapidly evolving

viruses, often operate on different time scales, so assessing both time scales together can provide information that would be lost by assessing either in isolation (Wheeler et al., 2010). Streicker et al. (2016) used the combined assessment of host population structure and pathogen phylogenies to determine that male vampire bats appear to disperse, while females show greater site fidelity, with important consequences for rabies transmission and expansion. This provides another important example of using host and pathogen phylogenetic analyses together to draw conclusions about host behavior, and then applying those findings to predict pathogen transmission dynamics.

While many phylogenetic studies focus on pathogen relatedness and genetic distance to make inferences, a developing tool for investigating disease dynamics is that of "phylodynamics." The term was first defined by Grenfell et al. (2004) as the "melding of immunodynamics, epidemiology, and evolutionary biology," and more specifically, phylodynamics seeks to understand the molecular footprint of epidemiological processes that are difficult to observe (Baele et al., 2017). Phylodynamic approaches have been used to investigate complicated questions such as short-term epidemic dynamics of human immunodeficiency virus in men who have sex with men (Lewis et al., 2008), or the impacts of urbanization on pathogen transmission and evolution in a wildlife system (Fountain-Jones, Craft, et al., 2017). Bayesian methods have been particularly useful within phylodynamics due to their ability to efficiently incorporate complex evolutionary models and uncertainty in parameter estimates (Drummond & Rambaut, 2007).

A developing branch within the broad field of phylodynamics is in the reconstruction of epidemics through the inference of transmission trees. These transmission trees are different from traditional phylogenetic trees, for several important reasons (Figure 1) (Didelot et al., 2017). First, in the context of epidemics of infectious disease, phylogenies consider each pathogen sample to be a tip (also known as a leaf) on the phylogeny, removing any possibility of a sampled pathogen being the ancestor of another sample (Picard et al., 2017). This means that phylogenies are unable to describe who infected whom. Transmission trees, however, allow sampled pathogens to be ancestors of other samples, and thereby may be able to reconstruct an epidemic by inferring who infected whom (Picard et al., 2017). The second major difference between transmission trees and phylogenetic tree is in the timing of coalescence events. Where timing of coalescent events in phylogenetic trees reflects branching events, timing in transmission trees reflects actual transmission events, which may occur at different time points from evolutionary branching events (Sintchenko & Holmes, 2015). It is important to understand these differences between phylogenetic and transmission trees, as the lines between the two are sometimes blurred and terminology used inappropriately, leading to confusion about the information provided by these two different methodologies.

Inferring transmission trees from epidemic data is a developing technique, with a variety of proposed methods (Cottam et al., 2008; Jombart et al., 2011; Ypma et al., 2012; Ypma, van Ballegooijen, et al., 2013; Didelot et al., 2014; Numminen et al., 2014; Hall et al., 2015; De Maio et al., 2016; Klinkenberg et al., 2017). While describing the methodological differences between these approaches is outside the scope of this review (but see Hall et al., 2016), the general process involves inferring transmission trees from pathogen sequence

data, while assuming one of the following: no within-host diversity or mutation, no within-host diversity but with mutation, or both within-host diversity and mutation (Hall et al., 2016). Pathogens with rapid mutation rates, such as RNA viruses, are generally more amenable to this type of epidemic reconstruction (Welch et al., 2011; Worby et al., 2014). Even with RNA viruses, however, it should not be assumed that an epidemic can be perfectly reconstructed (Hall et al., 2016), as all techniques will result in some uncertainty, particularly epidemics with incomplete sampling or slower pathogen evolution (Worby et al., 2014; Vasylyeva et al., 2016).

As transmission tree techniques have been further developed, however, they are better able to accommodate previous limitations; new techniques have now been used to study endemic (rather than epidemic) pathogens, ongoing epidemics, epidemics with incomplete sampling, and multiple disease introductions (Mollentze et al., 2014; Didelot et al., 2017). Judicious use of transmission tree methods can thus provide important information about transmission dynamics (Ypma, Jonges, et al., 2013; Numminen et al., 2014). For example, transmission trees have been used to assess the impact of wind direction on transmission of avian influenza in the Netherlands (Ypma, Jonges, et al., 2013). Transmission tree techniques can also shed light on the impact of host behavior and movements on pathogen transmission: Mollentze et al. (2014) used transmission tree reconstruction to show that rabies spread in South Africa appears to include the anthropogenic movement of dogs (i.e. in cars). The use of transmission trees is therefore a technique with great potential for understanding disease transmission dynamics in populations.

Epidemic reconstruction with only genetic sequence data, however, is limited in the resolution it can provide (Hall et al., 2016), and researchers should not expect fully resolved transmission events for an entire epidemic. The incorporation of other epidemiological data may be able to clarify some of the uncertainty inherent in inference of transmission trees (Hall et al., 2016), for which contact network approaches, in identifying likely avenues of transmission, may provide a fruitful path forward.

In summary, genomic analyses of hosts, pathogens, and both together have demonstrated their utility in investigating the dynamic relationship between social behavior and pathogen transmission. The emerging field of phylodynamics shows great promise for expanding upon and refining our understanding of transmission dynamics in populations. In particular, the developing techniques for transmission tree reconstruction provide opportunities to investigate pathogen transmission at greater resolution. These techniques, however, could be more powerful if integrated with other methodologies, such as epidemiologic and network approaches, for drawing conclusions about patterns of behavior and pathogen transmission.

Integrating contact networks and genomics: current efforts and future directions

A growing number of studies have begun to integrate pathogen genetic data with contact networks. These studies have provided new insights on identifying risk factors for disease acquisition, studying the social structure of a population, and learning how social structure may predict infection. Previous work has often used molecular techniques that focus on

pathogen strain-sharing or genetic distance between pathogen isolates across populations or individuals, and have largely utilized SNA in incorporating contact network techniques. We will discuss some examples in greater detail to highlight the range of approaches used to integrate contact network and genomic tools for a variety of research questions.

Current research questions

Do node characteristics predict pathogen genetic similarity?—Perhaps the most straightforward integration of social networks and genomics tools is to use SNA to test if node and network-level characteristics predict pathogen strain sharing or genetic similarity of isolates. Villaseñor-Sierra et al. (2007), for example, use this technique to compare node-level characteristics between children with and without group A *Streptococcus*. "Strain sharing" of group A *Streptococcus* was inferred by comparing restriction fragment length polymorphism patterns across isolates, and this study found that spread of group A *Streptococcus* clones was mediated by high connectedness among children, as detected by SNA. This straightforward integration of contact networks and genetic tools highlights the utility of this technique for establishing the importance of social networks for disease spread. This approach may also be useful for initial assessment of pathogen transmission in populations; if social structure is established as an important mediator of pathogen transmission in a given study system, this could justify further, more in-depth analysis, and may subsequently aid in developing targeted control measures.

Which social networks best predict pathogen networks?—A "next step" in advancing the integration of SNA and genomic tools has been to more explicitly compare a social network to pathogen relatedness across a population. This can be done by comparing a population's social network to a "strain-sharing network" based on microbial relatedness, and testing if the social network predicts the microbial network (Bull et al., 2012; Blyton et al., 2014; VanderWaal et al., 2014; Marquetoux et al., 2016; Springer et al., 2016; Fountain-Jones, Packer, et al., 2017). Assessing social networks in the context of a pathogen-sharing network can identify social network positions that are important for transmission (VanderWaal et al., 2014), test if the type of contact used to determine the social network (e.g. movement between farms) can explain pathogen transmission (Marquetoux et al., 2016), or even examine the varying importance of different modes of transmission (Fountain-Jones, Packer, et al., 2017). These strain-sharing networks, however, should not be confused with transmission trees; the former does not explicitly infer who infected whom, while the latter does. In addition, within strain-sharing networks, the molecular representation of strains can vary in resolution, ultimately affecting their power for inferring transmission. Advancements in genomic tools for inferring who infected whom may allow for increased resolution in these types of studies, and thereby potentially shed light on the variable impacts of social or behavioral factors on pathogen transmission.

What behaviors or locations are high risk for transmission?—Integrating contact networks and genomic tools can also be used to identify high risk behaviors or locations for pathogen transmission. For example, Gardy et al. (2011) utilized pathogen WGS and epidemiological data, in conjunction with social network analysis, to determine that a behavioral risk factor (likely crack-cocaine use) was a probable contributing factor in an

outbreak of *Mycobacterium tuberculosis* in people in Canada. In addition, Chamie et al. (2015) assessed genotype sharing of *Mycobacterium tuberculosis* in people in Uganda in the context of SNA; they investigated spatial overlap among genotype-clustered cases to identify likely transmission sites of social importance. Romano et al. (2010), using both social networks and hepatitis C virus gene sequences, assessed factors such as age, risk behaviors, and sexual networks to better understand risk factors even within different subtypes of the same pathogen. All of these studies demonstrate the utility of assessing pathogen genomics in the context of social networks for identifying important behavioral risk factors for transmission, including in instances where genotyping and contact tracing alone are not adequate for establishing pathogen dynamics (Gardy et al., 2011). All of the above example are studies in humans; in wildlife, such work would likely be more difficult as some methods for acquiring specific behavioral data in humans may not translate well to free-ranging animals (e.g. questionnaires are useful for acquiring behavioral data in humans but not applicable for wildlife).

Phylodynamics and future directions

The examples of previous work integrating contact networks and genomic tools have largely focused on using SNA, with limited utilization of higher-resolution molecular data from pathogens (but see Fountain-Jones, Packer, et al., 2017). The insights provided by phylodynamic approaches, however, suggest that integrating contact networks with phylodynamics may be a fruitful path forward. Some phylodynamics studies have begun to incorporate contact networks, including investigating what phylogenies, through phylodynamic analysis, can reveal about underlying host social structure (Leventhal et al., 2012; Robinson et al., 2013; Colijn & Gardy, 2014; McCloskey et al., 2016; Vasylyeva et al., 2016). Several of these studies have concluded that phylogenetic tree structure can provide some information about network structure and resulting transmission dynamics (Leventhal et al., 2012; Colijn & Gardy, 2014; McCloskey et al., 2016; Vasylyeva et al., 2016). For example, Leventhal et al. (2012) used this approach to conclude that random mixing among hosts was unlikely to result in an observed epidemic's phylogenetic tree. Colijn and Gardy (2014) used this approach to conclude that phylogenetic tree structure can help differentiate between transmission patterns (e.g. super-spreader versus homogenous mixing). However, the conclusions that can be drawn from pathogen phylogenies can be limited, and may not yield particularly novel conclusions for researchers debating the value of sequencing pathogen samples. This may be especially true for hosts with dynamic interactions, as particular caution must be used when inferring a population's underlying social structure for dynamic contact networks (Robinson et al., 2013; Vasylyeva et al., 2016). While these uses of phylogenies to describe underlying host contact structure may be helpful for parameterizing a theoretical network on which to simulate outbreaks of disease, other applications of phylodynamics methods may be better suited to integration with contact networks.

Inference of transmission trees may hold promise to fill this gap for further integrating phylodynamic approaches with contact networks. Like the phylodynamic studies which have attempted to infer host population structure from pathogen phylogenies, some studies have investigated what information about the underlying social network can be gleaned from

transmission trees alone (Carnegie, 2017). For example, assessment of HIV transmission trees has been used to infer preferential attachment in the underlying host social network, and this information was used to draw conclusions on the most appropriate management interventions (Leigh Brown et al., 2011). As with the phylogenetic tree approach, using transmission trees to make inferences about underlying population structure has some limitations. For instance, transmission trees have not demonstrated an ability to detect clustering in the underlying contact network (Welch, 2011). However, the potential utility of pathogen transmission trees for improving understanding of the dynamic interaction between host behavior and pathogen transmission has yet to be fully explored and remains a major gap in the current literature. Future work to explore avenues for integrating genomic and network approaches may specifically benefit from examining how contact networks can inform priors in Bayesian transmission tree reconstruction to resolve some of the inherent uncertainty in this method. In addition, transmission trees, being a higher-resolution representation of the transmission network, should be particularly amenable to statistical approaches that can examine relationships between networks, such as generalized dissimilarity modeling (GDM; Ferrier et al., 2007). Alternatively, a phylogeographic generalized linear model approach could be used, as in Lemey et al. (2014). Ultimately, integrating genomic data—including transmission trees—with network data should provide useful information about the importance of different modes of transmission and provide better predictions about pathogen transmission (Eames et al., 2015), further justifying the effort involved in incorporating these additional layers of data.

While the full scope of the applications of transmission trees to contact network studies has yet to be resolved and will vary depending on research question, some general guidelines can be identified. First, arguably the best-suited study systems for these approaches would be those with fast-evolving pathogens like RNA viruses (Didelot et al., 2017). The rapid evolution of these pathogens may allow for improved inference of transmission events, and could shed light on transmission dynamics at the shorter temporal scale on which many host social networks exist. In addition, if using strain-sharing or transmission networks to better understand population transmission dynamics, Blyton et al. (2014) point out that a study system should ideally have high strain alpha-diversity (local diversity) and high strain turnover in order to increase the resolution of transmission observations and to better represent contemporary transmission processes, respectively.

While ascertaining the proportion of the population that must be sampled to utilize a transmission tree approach has not been well established to date, pathogens which can be sampled across a greater proportion of the population, perhaps through non-invasive sampling techniques, would be more well-suited to transmission tree approaches. As methods for reconstruction of transmission trees advance, they are becoming more capable of accommodating unsampled individuals (Mollentze et al., 2014; Didelot et al., 2017), so outlining a specific proportion of a population to sample in order to use these methods is currently impractical. Ultimately, such determinations will be dependent on the current state of rapid methodological advances, and will be affected by specific research questions. For example, while low pathogen diversity and/or sampling effort will affect uncertainty to the degree that determining specific transmission events likely happened (rule-in) may be

unrealistic, if the objective is instead to rule out certain transmission events, these systems may be perfectly appropriate for transmission tree approaches (Didelot et al., 2017).

In addition to contact networks, incorporation of host data such as relatedness, space use, or varying definitions of contact (e.g. grooming versus fighting) may also be useful for integrating with transmission tree approaches, depending upon the research question (for two examples, see Figure 2). Furthermore, linking established phylogeographic models (Lemey et al., 2009, 2010) to transmission tree inference will enable more spatially explicit estimates of potential transmission events that can be compared directly to individual space use data. While specific examples of the integration of transmission trees to phylogeographic models and contact networks are, to the best of our knowledge, currently lacking in the primary literature, here we present some proposed pathways to integration, with the acknowledgement that, as phylodynamic approaches continue to advance, the opportunities to incorporate these tools into network studies will certainly evolve as well. In addition, we present advantages and disadvantages of current and proposed future methods for integrating network and genomic approaches in Table 1.

Which behaviors are most important for pathogen transmission?—The previously described examples of integrating contact networks and genomic techniques are limited in their directionality when describing pathogen transmission. In other words, they do not explicitly investigate who infected whom. In some cases, for example the movement of animals between farms, where movements or shipments are considered a “contact,” directionality is more obvious in the contact network (Marquetoux et al., 2016). In addition, detailed behavioral observations can suggest directionality of pathogen transmission between individual animals by investigating, for example, differences in infection incidence between grooming animals and the individuals being groomed (Drewe, 2010). But these are somewhat implied or correlational indicators of directionality, and would be strengthened by more refined inferences of transmission direction. In circumstances where direct observation is difficult and/or detailed epidemiological information is unavailable, phylodynamics-based methods, such as transmission tree inference, may be more capable of determining who infected whom. By incorporating higher resolution transmission data gleaned from these types of approaches, contact network techniques can provide clearer information about the impact of behavioral risk factors on pathogen transmission. For example, comparing a transmission network to various behaviorally-derived networks could provide insight into the types of behaviors that are most important for pathogen transmission (Figure 2b). As discussed above, GDM may be an important tool for this type of analysis, as it can be useful for scrutinizing complex network structures. Additional questions to investigate by comparing socially-derived networks to high-resolution transmission networks could include (but are not limited to): the importance of turnover of social ties, the relative importance of host genetic versus social distance, and the impact of host social structure or movement.

Can we better understand transmission-relevant contact?—The examples given above have all focused on investigating social networks together with assessments of pathogen “networks.” But this is, of course, only applicable in instances in which behavioral or observational data is adequate for describing the social network. Pluciński et al. (2011)

discuss the idea that pathogen relatedness could be used to infer information about the underlying contact network. In support of this idea, genetic relatedness between human commensal oral bacteria has been shown to correlate with social network distance, as determined by questionnaire surveys (Francis et al., 2016). Importantly, these studies highlight the potential for commensal or apathogenic organisms to be used as proxies of contact between individuals or even groups of individuals. For example, can microbiome composition or infection with apathogenic organisms be used to establish risk factors or "rules" to describe a mode of pathogen transmission in a population, and then can those rules predict subsequent transmission with other similar pathogens? Some commensal organisms will, of course, be unsuitable for this kind of transmission analysis. For example, some aspects of the host microbial community can be affected by diet, individual host factors, population of origin, or environmental transmission (Degnan et al., 2012; Blyton et al., 2013; Chiyo et al., 2014). In addition, organisms that transmit among and between multiple hosts would make discerning interactions within a single host species far more challenging; instead, single-host organisms with well-characterized modes of transmission may be most effective at elucidating transmission-relevant contact in a single host species. Rapidly evolving, apathogenic viruses may hold the most promise as proxies of contact, as their diversity and high mutation rates should allow for greater resolution in illuminating transmission events. In addition, identification of such viruses is becoming increasingly achievable due to advances in viromic technologies and non-invasive sampling (Minot et al., 2013; Rasmussen, 2015). The use of such apathogenic infections for detecting transmission-relevant contact may be particularly important for disease systems where the pathogen of interest cannot be easily sampled and sequenced due to, for example, very short shedding periods (e.g. about 1–2 weeks for canine distemper virus, Greene, 2012). Instead, commensal or apathogenic infection information could be utilized to determine a virtual contact network on which to simulate disease outbreaks, potentially without the use of resource intensive observational data typically required to estimate contact rates across populations (Figure 2c). This type of approach could potentially reveal and predict, for example, how animal movement behaviors and natural or anthropogenic landscape features affect disease transmission across populations. Establishing the utility of this type of approach is a clear gap in the literature, and would benefit from future research.

Limitations—As previously noted, transmission tree approaches cannot be expected to reconstruct epidemics with perfect certainty, particularly for slow-evolving pathogens. Incorporating additional bounds and levels of information can, however, improve inferences from phylodynamics and integrative approaches (Welch et al., 2011; Ray et al., 2016). This multimodal approach, which may include data such as host genomics, landscape factors, and epidemic parameters, may be important in studies of pathogens with slow rates of mutation or otherwise limited genetic diversity (Wylie et al., 2005). The current literature on how to utilize a multimodal research framework is limited at this time (Ray et al., 2016), but carries potential for future expansion and may help resolve some areas of uncertainty within phylodynamic approaches. Future work to explore the impacts of incorporating different layers of data—including pathogen genomics—into contact networks should shed light on how these various factors are able to predict and describe the interaction between animal behavior and transmission dynamics.

Of course, within mathematical modeling, there is a perpetual trade-off between precision, generality, and realism (Levins, 1966). In this review, we have discussed the advantages gained by adding complexity to disease models through the use of contact networks and genomics tools, but it must be pointed out that these additions may come at the cost of computational efficiency and generality. While models that are very realistic or precise for a particular study system can provide useful information for managing a particular disease within a specific population, increased complexity within models is not always the best approach (Buhnerkempe et al., 2015). Ultimately, a key challenge for disease modeling, broadly, is in understanding when simplicity or complexity is more appropriate (Buhnerkempe et al., 2015), and this dichotomy should not be forgotten amongst the advancements offered by more precise or realistic data and modeling.

Conclusions

This review has focused on describing how contact networks and genomic tools each shed light on the dynamic interaction between animal social behavior and infectious disease dynamics, and how these tools can be integrated for improved and new insights. These techniques may be particularly useful for wildlife researchers looking to make the most of hard-won pathogen data, and to advance their understanding of pathogen dynamics and social behavior. As genomics approaches become more accessible to more researchers, and the statistical and computational tools used to analyze genomic outputs continue to advance, the opportunities to utilize multimodal approaches in disease modeling will expand. Staying up-to-date with this progress will ultimately allow researchers to use novel techniques to answer complicated questions about animal behavior and its impact on pathogen transmission, and consequently make predictions important to the surveillance, prevention, and management of infectious diseases in populations.

Acknowledgements

M.L.J.G. was supported by the Office of the Director, National Institutes of Health under award number NIH T32OD010993. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. N.M. F-J was funded by National Science Foundation (DEB-1413925) and M.E.C. was funded by National Science Foundation (DEB-1654609), the University of Minnesota's Office of the Vice President for Research and Academic Health Center Seed Grant, and the Cooperative State Research Service, U.S. Department of Agriculture, under Project No. MIN-62–098. Special thanks to L. White, K. Worsley-Tonks, Y. Wang, and two anonymous reviewers for their invaluable input and suggestions.

References

Anderson RM & May RM (1991). Infectious Diseases of Humans: Dynamics and Control - OUP Oxford.

Archie EA, Luikart G, & Ezenwa VO (2009). Infecting epidemiology with genetics: a new frontier in disease ecology. - Trends Ecol. Evol 24: 21–30. [PubMed: 19027985]

Baele G, Suchard MA, Rambaut A, & Lemey P (2017). Emerging concepts of data integration in pathogen phylodynamics. - Syst. Biol 66: e47–e65. [PubMed: 28173504]

Becker DJ, Streicker DG, & Altizer S (2015). Linking anthropogenic resources to wildlife-pathogen dynamics: a review and meta-analysis. - Ecol. Lett 18: 483–495. [PubMed: 25808224]

Biek R, Drummond AJ, & Poss M (2006). A virus reveals population structure and recent demographic history of its carnivore host. - Science 311: 538–541. [PubMed: 16439664]

Biek R, Henderson JC, Waller LA, Rupprecht CE, & Real LA (2007). A high-resolution genetic signature of demographic and spatial expansion in epizootic rabies virus. - Proc. Natl. Acad. Sci. U. S. A 104: 7993–7998. [PubMed: 17470818]

Bird BH, Khristova ML, Rollin PE, Ksiazek TG, & Nichol ST (2007). Complete genome analysis of 33 ecologically and biologically diverse Rift Valley fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. - J. Virol 81: 2805–2816. [PubMed: 17192303]

Blanchong JA, Samuel MD, Scribner KT, Weckworth BV, Langenberg JA, & Filcek KB (2007). Landscape genetics and the spatial distribution of chronic wasting disease. - Biol. Lett 4: 130–133.

Blasse A, Calvignac-Spencer S, Merkel K, Goffe AS, Boesch C, Mundry R, & Leendertz FH (2013). Mother-offspring transmission and age-dependent accumulation of simian foamy virus in wild chimpanzees. - J. Virol 87: 5193–5204. [PubMed: 23449796]

Blyton MDJ, Banks SC, Peakall R, & Gordon DM (2013). High temporal variability in commensal *Escherichia coli* strain communities of a herbivorous marsupial. - Environ. Microbiol 15: 2162–2172. [PubMed: 23414000]

Blyton MDJ, Banks SC, Peakall R, Lindenmayer DB, & Gordon DM (2014). Not all types of host contacts are equal when it comes to *E. coli* transmission. - Ecol. Lett 17: 970–978. [PubMed: 24861219]

Buhnerkempe MG, Roberts MG, Dobson AP, Heesterbeek H, Hudson PJ, & Lloyd-Smith JO (2015). Eight challenges in modelling disease ecology in multi-host, multi-agent systems. - Epidemics 10: 26–30. [PubMed: 25843378]

Bull CM, Godfrey SS, & Gordon DM (2012). Social networks and the spread of *Salmonella* in a sleepy lizard population. - Mol. Ecol 21: 4386–4392. [PubMed: 22845647]

Carnegie NB (2017). Effects of contact network structure on epidemic transmission trees: implications for data required to estimate network structure. - Stat. Med

Chamie G, Wandera B, Marquez C, Kato-Maeda M, Kamya MR, Havlir DV, & Charlebois ED (2015). Identifying locations of recent TB transmission in rural Uganda: a multidisciplinary approach. - Trop. Med. Int. Health 20: 537–545. [PubMed: 25583212]

Chen S, White BJ, Sanderson MW, Amrine DE, Ilany A, & Lanzas C (2014). Highly dynamic animal contact network and implications on disease transmission. - Sci. Rep 4: 4472. [PubMed: 24667241]

Chiyo PI, Grieneisen LE, Wittemyer G, Moss CJ, Lee PC, Douglas-Hamilton I, & Archie EA (2014). The influence of social structure, habitat, and host traits on the transmission of *Escherichia coli* in wild elephants. - PLoS One 9: e93408. [PubMed: 24705319]

Colijn C & Gardy J (2014). Phylogenetic tree shapes resolve disease transmission patterns. - Evol Med Public Health 2014: 96–108. [PubMed: 24916411]

Cottam EM, Thébaud G, Wadsworth J, Gloster J, Mansley L, Paton DJ, King DP, & Haydon DT (2008). Integrating genetic and epidemiological data to determine transmission pathways of foot-and-mouth disease virus. - Proc. Biol. Sci 275: 887–895. [PubMed: 18230598]

Craft ME (2015). Infectious disease transmission and contact networks in wildlife and livestock. - Philos. Trans. R. Soc. Lond. B Biol. Sci 370.

Craft ME & Caillaud D (2011). Network models: an underutilized tool in wildlife epidemiology? - Interdiscip. Perspect. Infect. Dis 2011: 676949. [PubMed: 21527981]

Craft ME, Volz E, Packer C, & Meyers LA (2009). Distinguishing epidemic waves from disease spillover in a wildlife population. - Proc. Biol. Sci 276: 1777–1785. [PubMed: 19324800]

Craft ME, Volz E, Packer C, & Meyers LA (2011). Disease transmission in territorial populations: the small-world network of Serengeti lions. - J. R. Soc. Interface 8: 776–786. [PubMed: 21030428]

Croft DP, James R, & Krause J (2008). Exploring Animal Social Networks - Princeton University Press.

Cullingham CI, Kyle CJ, Pond BA, Rees EE, & White BN (2009). Differential permeability of rivers to raccoon gene flow corresponds to rabies incidence in Ontario, Canada. - Mol. Ecol 18: 43–53. [PubMed: 19140963]

de Carvalho Ferreira HC, Weesendorp E, Quak S, Stegeman JA, & Loeffen WLA (2014). Suitability of faeces and tissue samples as a basis for non-invasive sampling for African swine fever in wild boar. - Vet. Microbiol 172: 449–454. [PubMed: 25017975]

Degnan PH, Pusey AE, Lonsdorf EV, Goodall J, Wroblewski EE, Wilson ML, Rudicell RS, Hahn BH, & Ochman H (2012). Factors associated with the diversification of the gut microbial communities within chimpanzees from Gombe National Park. - Proc. Natl. Acad. Sci. U. S. A 109: 13034–13039. [PubMed: 22826227]

De Maio N, Wu C-H, O'Reilly KM, & Wilson D (2015). New routes to phylogeography: A Bayesian structured coalescent approximation. - PLoS Genet 11: e1005421. [PubMed: 26267488]

De Maio N, Wu C-H, & Wilson DJ (2016). SCOTTI: Efficient reconstruction of transmission within outbreaks with the structured coalescent. - PLoS Comput. Biol 12: e1005130. [PubMed: 27681228]

Didelot X, Fraser C, Gardy J, & Colijn C (2017). Genomic infectious disease epidemiology in partially sampled and ongoing outbreaks. - Mol. Biol. Evol 34: 997–1007. [PubMed: 28100788]

Didelot X, Gardy J, & Colijn C (2014). Bayesian inference of infectious disease transmission from whole-genome sequence data. - Mol. Biol. Evol 31: 1869–1879. [PubMed: 24714079]

Dizney L & Dearing MD (2013). The role of behavioural heterogeneity on infection patterns: implications for pathogen transmission. - Anim. Behav 86.

Drewe JA (2010). Who infects whom? Social networks and tuberculosis transmission in wild meerkats. - Proc. Biol. Sci 277: 633–642. [PubMed: 19889705]

Drummond AJ & Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. - BMC Evol. Biol 7: 214. [PubMed: 17996036]

Eames K, Bansal S, Frost S, & Riley S (2015). Six challenges in measuring contact networks for use in modelling. - Epidemics 10: 72–77. [PubMed: 25843388]

Ferrier S, Manion G, Elith J, & Richardson K (2007). Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. - Diversity and Distributions 13: 252–264.

Fountain-Jones NM, Craft ME, Funk WC, Kozakiewicz C, Trumbo D, Boydston EE, Lyren LM, Crooks K, Lee JS, VandeWoude S, & Carver S (2017). Urban landscapes can change virus gene flow and evolution in a fragmentation-sensitive carnivore. - Mol. Ecol 26: 6487–6498. [PubMed: 28987024]

Fountain-Jones NM, Packer C, Troyer JL, VanderWaal K, Robinson S, Jacquot M, & Craft ME (2017). Linking social and spatial networks to viral community phylogenetics reveals subtype-specific transmission dynamics in African lions. - J. Anim. Ecol 86: 1469–1482. [PubMed: 28884827]

Francis SS, Plucinski MM, Wallace AD, & Riley LW (2016). Genotyping oral commensal bacteria to predict social contact and structure. - PLoS One 11: e0160201. [PubMed: 27684062]

Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, Rempel S, Moore R, Zhao Y, Holt R, Varhol R, Birol I, Lem M, Sharma MK, Elwood K, Jones SJM, Brinkman FSL, Brunham RC, & Tang P (2011). Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. - N. Engl. J. Med 364: 730–739. [PubMed: 21345102]

Godfrey SS (2013). Networks and the ecology of parasite transmission: A framework for wildlife parasitology. - Int. J. Parasitol. Parasites Wildl 2: 235–245. [PubMed: 24533342]

Gottdenker NL, Streicker DG, Faust CL, & Carroll CR (2014). Anthropogenic land use change and infectious diseases: a review of the evidence. - Ecohealth 11: 619–632. [PubMed: 24854248]

Grad YH, Kirkcaldy RD, Trees D, Dordel J, Harris SR, Goldstein E, Weinstock H, Parkhill J, Hanage WP, Bentley S, & Lipsitch M (2014). Genomic epidemiology of *Neisseria gonorrhoeae* with reduced susceptibility to cefixime in the USA: a retrospective observational study. - Lancet Infect. Dis 14: 220–226. [PubMed: 24462211]

Greene CE (2012). Infectious diseases of the dog and cat - Elsevier/Saunders, St. Louis, Mo., 4th ed.

Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, & Holmes EC (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. - Science 303: 327–332. [PubMed: 14726583]

Hall MD, Woolhouse MEJ, & Rambaut A (2016). Using genomics data to reconstruct transmission trees during disease outbreaks. - Rev. Sci. Tech 35: 287–296. [PubMed: 27217184]

Hall M, Woolhouse M, & Rambaut A (2015). Epidemic reconstruction in a phylogenetics framework: Transmission trees as partitions of the node set. - PLoS Comput. Biol 11: e1004613. [PubMed: 26717515]

Hoffmann C, Stockhausen M, Merkel K, Calvignac-Spencer S, & Leendertz FH (2016). Assessing the feasibility of fly based surveillance of wildlife infectious diseases. - Sci. Rep 6: 37952. [PubMed: 27901062]

Jombart T, Eggo RM, Dodd PJ, & Balloux F (2011). Reconstructing disease outbreaks from genetic data: a graph approach. - Heredity 106: 383–390. [PubMed: 20551981]

Kao RR, Haydon DT, Lycett SJ, & Murcia PR (2014). Supersize me: how whole-genome sequencing and big data are transforming epidemiology. - Trends Microbiol 22: 282–291. [PubMed: 24661923]

Keeling MJ & Eames KTD (2005). Networks and epidemic models. - J. R. Soc. Interface 2: 295–307. [PubMed: 16849187]

Keeling MJ & Rohani P (2011). Modeling Infectious Diseases in Humans and Animals - Princeton University Press.

Klinkenberg D, Backer JA, Didelot X, Colijn C, & Wallinga J (2017). Simultaneous inference of phylogenetic and transmission trees in infectious disease outbreaks. - PLoS Comput. Biol 13: e1005495. [PubMed: 28545083]

Lee JS, Ruell EW, Boydston EE, Lyren LM, Alonso RS, Troyer JL, Crooks KR, & Vandewoude S (2012). Gene flow and pathogen transmission among bobcats (*Lynx rufus*) in a fragmented urban landscape. - Mol. Ecol 21: 1617–1631. [PubMed: 22335296]

Leigh Brown AJ, Lycett SJ, Weinert L, Hughes GJ, Fearnhill E, Dunn DT, & UK HIV Drug Resistance Collaboration (2011). Transmission network parameters estimated from HIV sequences for a nationwide epidemic. - J. Infect. Dis 204: 1463–1469. [PubMed: 21921202]

Lembo T, Hampson K, Haydon DT, Craft M, Dobson A, Dushoff J, Ernest E, Hoare R, Kaare M, Mlengeya T, Mentzel C, & Cleaveland S (2008). Exploring reservoir dynamics: a case study of rabies in the Serengeti ecosystem. - J. Appl. Ecol 45: 1246–1257. [PubMed: 22427710]

Lemey P, Rambaut A, Bedford T, Faria N, Bielejec F, Baele G, Russell CA, Smith DJ, Pybus OG, Brockmann D, & Suchard MA (2014). Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. - PLoS Pathog 10: e1003932. [PubMed: 24586153]

Lemey P, Rambaut A, Drummond AJ, & Suchard MA (2009). Bayesian phylogeography finds its roots. - PLoS Comput. Biol 5: e1000520. [PubMed: 19779555]

Lemey P, Rambaut A, Welch JJ, & Suchard MA (2010). Phylogeography takes a relaxed random walk in continuous space and time. - Mol. Biol. Evol 27: 1877–1885. [PubMed: 20203288]

Leventhal GE, Kouyos R, Stadler T, von Wyl V, Yerly S, Böni J, Cellerai C, Klimkait T, Günthard HF, & Bonhoeffer S (2012). Inferring epidemic contact structure from phylogenetic trees. - PLoS Comput. Biol 8: e1002413. [PubMed: 22412361]

Levins R (1966). The strategy of model building in population biology. - Am. Sci 54: 421–431.

Lewis F, Hughes GJ, Rambaut A, Pozniak A, & Leigh Brown AJ (2008). Episodic sexual transmission of HIV revealed by molecular phylodynamics. - PLoS Med 5: e50. [PubMed: 18351795]

Lloyd-Smith JO, George D, Pepin KM, Pitzer VE, Pulliam JRC, Dobson AP, Hudson PJ, & Grenfell BT (2009). Epidemic dynamics at the human-animal interface. - Science 326: 1362–1367. [PubMed: 19965751]

Lloyd-Smith JO, Schreiber SJ, Kopp PE, & Getz WM (2005). Superspreading and the effect of individual variation on disease emergence. - Nature 438: 355–359. [PubMed: 16292310]

MacIntosh AJJ, Jacobs A, Garcia C, Shimizu K, Mouri K, Huffman MA, & Hernandez AD (2012). Monkeys in the middle: parasite transmission through the social network of a wild primate. - PLoS One 7: e51144. [PubMed: 23227246]

Marquetoux N, Heuer C, Wilson P, Ridler A, & Stevenson M (2016). Merging DNA typing and network analysis to assess the transmission of paratuberculosis between farms. - Prev. Vet. Med 134: 113–121. [PubMed: 27836032]

Martínez-López B, Perez AM, & Sánchez-Vizcaíno JM (2009). Social network analysis. Review of general concepts and use in preventive veterinary medicine. - Transbound. Emerg. Dis 56: 109–120. [PubMed: 19341388]

McCloskey RM, Liang RH, & Poon AFY (2016). Reconstructing contact network parameters from viral phylogenies. - Virus Evol 2: vew029. [PubMed: 27818787]

Metzker ML, Mindell DP, Liu X-M, Ptak RG, Gibbs RA, & Hillis DM (2002). Molecular evidence of HIV-1 transmission in a criminal case. - Proc. Natl. Acad. Sci. U. S. A 99: 14292–14297. [PubMed: 12388776]

Meyers LA (2007). Contact network epidemiology: Bond percolation applied to infectious disease prediction and control. - Bull. Am. Math. Soc 44: 63–87.

Minot S, Bryson A, Chehoud C, Wu GD, Lewis JD, & Bushman FD (2013). Rapid evolution of the human gut virome. - Proc. Natl. Acad. Sci. U. S. A 110: 12450–12455. [PubMed: 23836644]

Mollentze N, Nel LH, Townsend S, le Roux K, Hampson K, Haydon DT, & Soubeyrand S (2014). A Bayesian approach for inferring the dynamics of partially observed endemic infectious diseases from space-time-genetic data. - Proc. Biol. Sci 281: 20133251. [PubMed: 24619442]

Natoli E, Say L, Cafazzo S, Bonanni R, Schmid M, & Pontier D (2005). Bold attitude makes male urban feral domestic cats more vulnerable to Feline Immunodeficiency Virus. - Neurosci. Biobehav. Rev 29: 151–157. [PubMed: 15652262]

Numminen E, Chewapreecha C, Sirén J, Turner C, Turner P, Bentley SD, & Corander J (2014). Two-phase importance sampling for inference about transmission trees. - Proc. Biol. Sci 281: 20141324. [PubMed: 25253455]

Nunn CL, Jordán F, McCabe CM, Verdolin JL, & Fewell JH (2015). Infectious disease and group size: more than just a numbers game. - Philos. Trans. R. Soc. Lond. B Biol. Sci 370.

Pellis L, Ball F, Bansal S, Eames K, House T, Isham V, & Trapman P (2015). Eight challenges for network epidemic models. - Epidemics 10: 58–62. [PubMed: 25843385]

Perkins SE, Cagnacci F, Stradiotto A, Arnoldi D, & Hudson PJ (2009). Comparison of social networks derived from ecological data: implications for inferring infectious disease dynamics. - J. Anim. Ecol 78: 1015–1022. [PubMed: 19486206]

Picard C, Dallot S, Brunker K, Berthier K, Roumagnac P, Soubeyrand S, Jacquot E, & Thébaud G (2017). Exploiting genetic information to trace plant virus dispersal in landscapes. - Annu. Rev. Phytopathol 55: 139–160. [PubMed: 28525307]

Pluciński MM, Starfield R, & Almeida RPP (2011). Inferring social network structure from bacterial sequence data. - PLoS One 6: e22685. [PubMed: 21829645]

Pope LC, Butlin RK, Wilson GJ, Woodroffe R, Erven K, Conyers CM, Franklin T, Delahay RJ, Cheeseman CL, & Burke T (2007). Genetic evidence that culling increases badger movement: implications for the spread of bovine tuberculosis. - Mol. Ecol 16: 4919–4929. [PubMed: 17944854]

Porphyre T, Stevenson M, Jackson R, & McKenzie J (2008). Influence of contact heterogeneity on TB reproduction ratio R0 in a free-living brushtail possum *Trichosurus vulpecula* population. - Vet. Res 39: 31. [PubMed: 18275805]

Rasmussen AL (2015). Probing the viromic frontiers. - MBio 6: e01767–15. [PubMed: 26556279]

Ray B, Ghedin E, & Chunara R (2016). Network inference from multimodal data: A review of approaches from infectious disease transmission. - J. Biomed. Inform 64: 44–54. [PubMed: 27612975]

Reynolds JJH, Hirsch BT, Gehrt SD, & Craft ME (2015). Raccoon contact networks predict seasonal susceptibility to rabies outbreaks and limitations of vaccination. - J. Anim. Ecol 84: 1720–1731. [PubMed: 26172427]

Robert K, Garant D, & Pelletier F (2012). Keep in touch: Does spatial overlap correlate with contact rate frequency? - J. Wildl. Manage 76: 1670–1675.

Robinson K, Fyson N, Cohen T, Fraser C, & Colijn C (2013). How the dynamics and structure of sexual contact networks shape pathogen phylogenies. - PLoS Comput. Biol 9: e1003105. [PubMed: 23818840]

Romano CM, de Carvalho-Mello IMVG, Jamal LF, de Melo FL, Iamarino A, Motoki M, Pinho JRR, Holmes EC, de Andrade Zanotto PM, & VGDN Consortium (2010). Social networks shape the transmission dynamics of hepatitis C virus. - PLoS One 5: e11170. [PubMed: 20585651]

Rushmore J, Caillaud D, Hall RJ, Stumpf RM, Meyers LA, & Altizer S (2014). Network-based vaccination improves prospects for disease control in wild chimpanzees. - J. R. Soc. Interface 11: 20140349. [PubMed: 24872503]

Rushmore J, Caillaud D, Matamba L, Stumpf RM, Borgatti SP, & Altizer S (2013). Social network analysis of wild chimpanzees provides insights for predicting infectious disease risk. - J. Anim. Ecol 82: 976–986. [PubMed: 23734782]

Sah P, Leu ST, Cross PC, Hudson PJ, & Bansal S (2017). Unraveling the disease consequences and mechanisms of modular structure in animal social networks. - Proc. Natl. Acad. Sci. U. S. A 114: 4165–4170. [PubMed: 28373567]

Sharp PM & Hahn BH (2010). The evolution of HIV-1 and the origin of AIDS. - Philos. Trans. R. Soc. Lond. B Biol. Sci 365: 2487–2494. [PubMed: 20643738]

Silk MJ, Croft DP, Delahay RJ, Hodgson DJ, Weber N, Boots M, & McDonald RA (2017). The application of statistical network models in disease research. - Methods Ecol. Evol 8: 1026–1041.

Sintchenko V & Holmes EC (2015). The role of pathogen genomics in assessing disease transmission. - BMJ 350: h1314. [PubMed: 25964672]

Smiley Evans T, Gilardi KVK, Barry PA, Ssebide BJ, Kinani JF, Nizeyimana F, Noheri JB, Byarugaba DK, Mudakikwa A, Cranfield MR, Mazet JAK, & Johnson CK (2016). Detection of viruses using discarded plants from wild mountain gorillas and golden monkeys. - Am. J. Primatol 78: 1222–1234. [PubMed: 27331804]

Springer A, Mellmann A, Fichtel C, & Kappeler PM (2016). Social structure and *Escherichia coli* sharing in a group-living wild primate, Verreaux's sifaka. - BMC Ecol 16: 6. [PubMed: 26868261]

Streicker DG, Winternitz JC, Satterfield DA, Condori-Condori RE, Broos A, Tello C, Recuenco S, Velasco-Villa A, Altizer S, & Valderrama W (2016). Host-pathogen evolutionary signatures reveal dynamics and future invasions of vampire bat rabies. - Proc. Natl. Acad. Sci. U. S. A 113: 10926–10931. [PubMed: 27621441]

VanderWaal KL, Atwill ER, Isbell LA, & McCowan B (2014). Linking social and pathogen transmission networks using microbial genetics in giraffe (*Giraffa camelopardalis*). - J. Anim. Ecol 83: 406–414. [PubMed: 24117416]

VanderWaal KL & Ezenwa VO (2016). Heterogeneity in pathogen transmission: mechanisms and methodology. - Funct. Ecol 30: 1606–1622.

Vasylyeva TI, Friedman SR, Paraskevis D, & Magiorkinis G (2016). Integrating molecular epidemiology and social network analysis to study infectious diseases: Towards a socio-molecular era for public health. - Infect. Genet. Evol 46: 248–255. [PubMed: 27262354]

Villaseñor-Sierra A, Quiñonez-Alvarado MG, & Caballero-Hoyos JR (2007). Interpersonal relationships and group A *Streptococcus* spread in a Mexican day-care center. - Salud Publica Mex 49: 323–329. [PubMed: 17952239]

Welch D (2011). Is network clustering detectable in transmission trees? - Viruses 3: 659–676. [PubMed: 21731813]

Welch D, Bansal S, & Hunter DR (2011). Statistical inference to advance network models in epidemiology. - Epidemics 3: 38–45. [PubMed: 21420658]

Wheeler DC, Waller LA, & Biek R (2010). Spatial analysis of feline immunodeficiency virus infection in cougars. - Spat. Spatiotemporal Epidemiol 1: 151–161. [PubMed: 21197421]

White LA, Forester JD, & Craft ME (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. - Biol. Rev. Camb. Philos. Soc 92: 389–409. [PubMed: 26613547]

Worby CJ, Lipsitch M, & Hanage WP (2014). Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. - PLoS Comput. Biol 10: e1003549. [PubMed: 24675511]

Wylie JL, Cabral T, & Jolly AM (2005). Identification of networks of sexually transmitted infection: a molecular, geographic, and social network analysis. - J. Infect. Dis 191: 899–906. [PubMed: 15717265]

Ypma RJF, Bataille AMA, Stegeman A, Koch G, Wallinga J, & van Ballegooijen WM (2012). Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data. - Proc. Biol. Sci 279: 444–450. [PubMed: 21733899]

Ypma RJF, Jonges M, Bataille A, Stegeman A, Koch G, van Boven M, Koopmans M, van Ballegooijen WM, & Wallinga J (2013). Genetic data provide evidence for wind-mediated transmission of highly pathogenic avian influenza. - J. Infect. Dis 207: 730–735. [PubMed: 23230058]

Ypma RJF, van Ballegooijen WM, & Wallinga J (2013). Relating phylogenetic trees to transmission trees of infectious disease outbreaks. - Genetics 195: 1055–1062. [PubMed: 24037268]

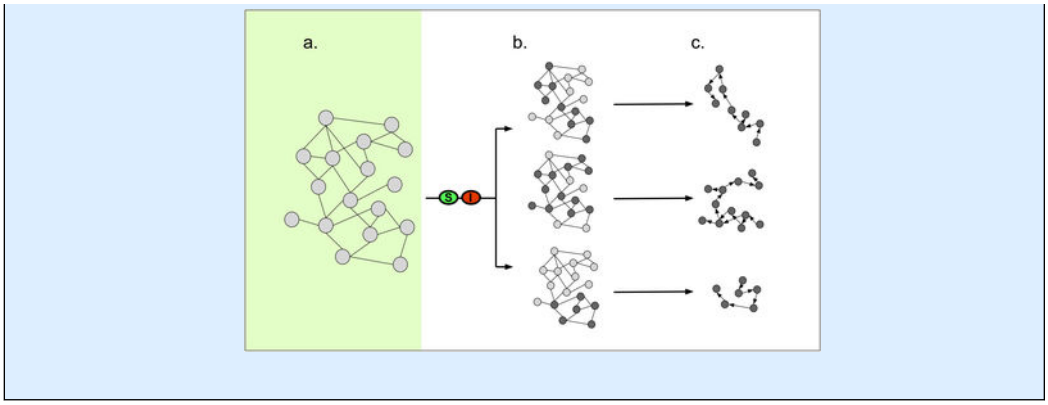
Box 1:

The basics of contact networks

Within network approaches, “nodes” can represent individual humans or animals, or can represent groups of individuals such as herds or even farms. Connections or contacts between nodes are called “edges,” and are defined depending on the research question; for example, in a network where nodes represent individual wild animals, and the pathogen in question is transmitted by sexual contacts, an edge might be drawn between two individuals known to have had sexual contact. Alternatively, in a network of farms, with a pathogen transmitted by direct or indirect contact, an edge might be drawn between two farms if they have shipped animals from one farm to the other. This highlights the importance of carefully defining a “contact” when using network approaches, depending on the system and research question (Perkins et al., 2009).

While social network analysis approaches are often descriptive or observational, network modeling approaches typically involve simulating an outbreak of a pathogen in a population (but see the main text for caveats about this dichotomy). Simulating an outbreak of a pathogen on a contact network, in light of the transmission probability (e.g. properties of the infectious host, susceptible host, and infectivity of the pathogen), can result in the output of a transmission network which can graphically represent who infected whom (Figures a-c). This transmission network is therefore different from the observed contact network, and could be viewed as a subset of the contact network. Network approaches are, however, dependent on extensive behavioral or observational data to create the initial contact network, and can be highly influenced by how contacts are defined and detected in the population (Perkins et al., 2009; Eames et al., 2015). In some cases, the entire population being studied is observed to build the “complete” contact network; in other cases, this is not feasible or practical, and only a portion of the population of interest will be observed. In these cases, the observed population can be used to create a set of “rules” for contact in the population, and then extrapolated to the full population (e.g. with ERGMs, as in Reynolds et al., 2015; Silk et al., 2017). Thus, even networks built on a sample of empirical observations may be generalized to a larger population.

Conceptual flow from a contact network to a transmission network. Areas with a green background represent components informed by empirical data; areas with a white background represent components informed by simulations. Network (a.) shows a contact network that would be defined by, for example, observational data of direct contacts. A disease process, such as an SI (susceptible-infectious) model, (b.) could be applied to the contact network many times (pictured three times here, but realistically, a simulation would be run on the order of 1000 times). In this example, the index case was randomly seeded, and darkened nodes represent those nodes that were infected in the course of a simulation. The who-infected-whom, transmission networks (c.) could be a final output of network model simulations.



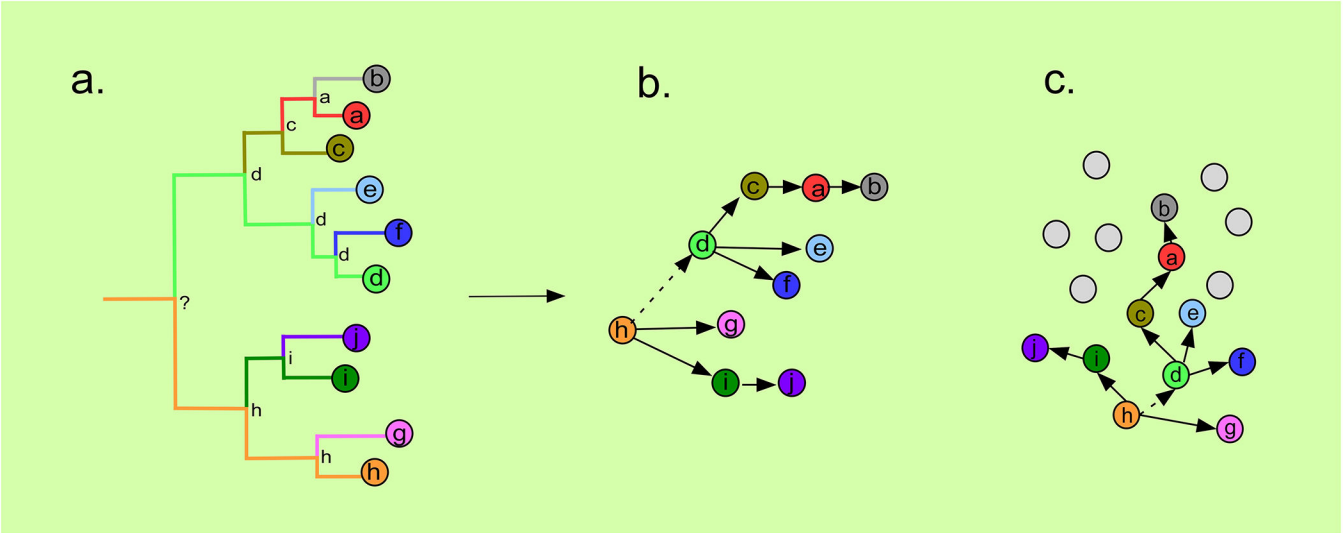


Figure 1.

Conceptual flow from (a.) a transmission tree to (b.) a transmission network; (c.) depicts a transmission network in the context of the rest of the population, where gray nodes represent uninfected individuals. The green background highlights that this approach is based on empirical data, rather than simulations. Colored circles represent sequenced samples from infected individuals. In (a.), branching events in the transmission tree indicate transmission events, with color changes occurring at these events; lettered labels at internal nodes represent infecting individuals. The colored lines and labels at branching events in the transmission tree highlight a primary difference between phylogenetic trees and transmission trees: pathogens sampled at the “tips” of the transmission tree are allowed to be ancestors of other samples. This then allows for the inference of who infected whom, as demonstrated by the directed networks in (b.) and (c.), but with some uncertainty that cannot be fully resolved (represented here by uncertain transmission from individual “h” to individual “d”).

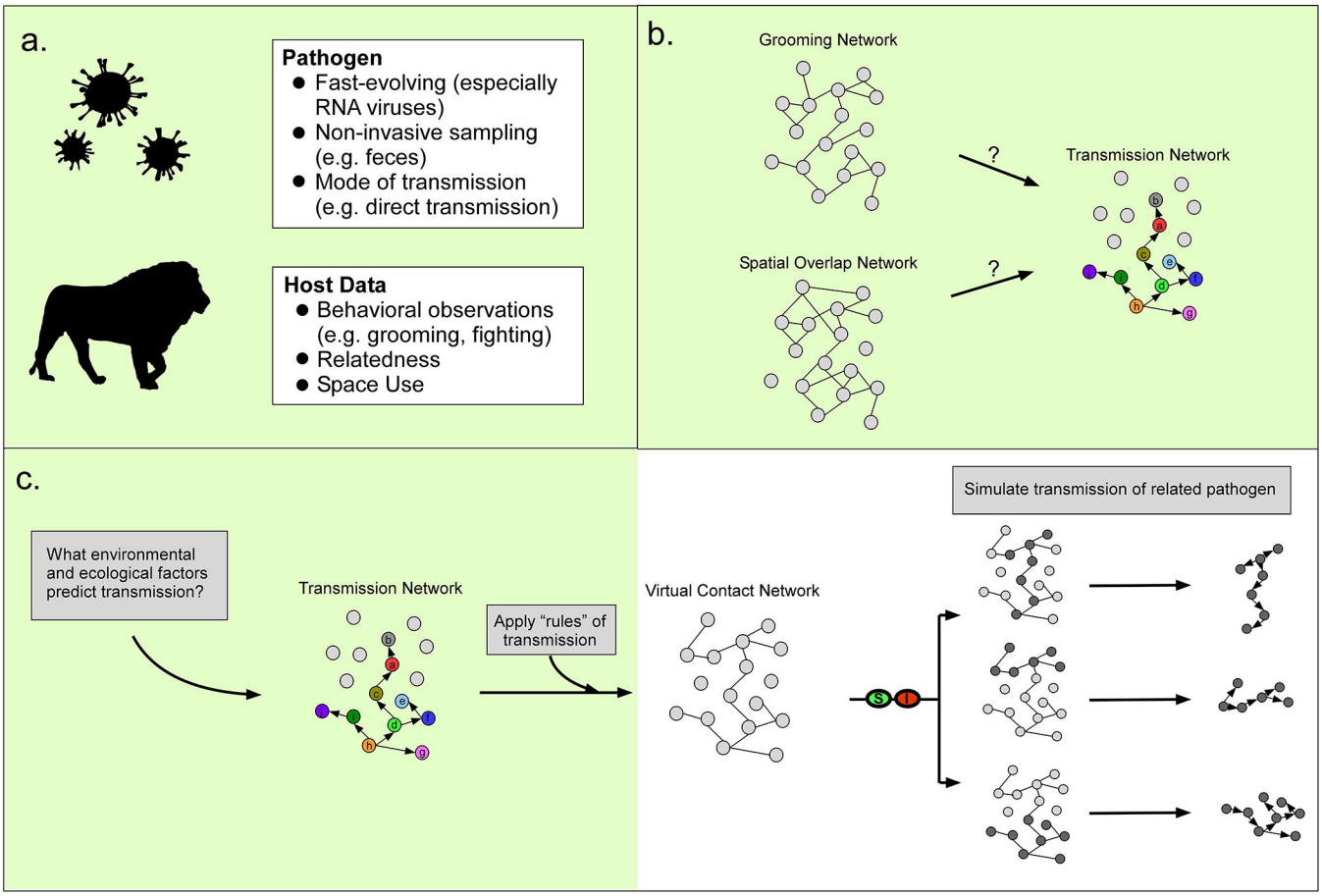


Figure 2.

Two examples of proposed integrations of contact networks and genomic tools, in this case focusing on utilizing transmission networks derived from inference of transmission trees. Areas with a green background represent components informed by empirical data; areas with a white background represent components informed by simulations. Panel (a.) highlights some characteristics of pathogens and types of host data that may be well-suited to transmission tree approaches, but these are by no means all-inclusive. Panel (b.) demonstrates a SNA application, in which observed networks such as grooming or spatial overlap networks, are compared to high-resolution transmission networks to determine what social or behavioral factors have the greatest impact on pathogen transmission. Panel (c.) depicts a network modeling application, in which a transmission network derived from an apathogenic or commensal organism is used to determine environmental and ecological factors that best predict transmission in a population. These "rules" are then used to create a "virtual contact network" on which epidemics with a related (pathogenic) organism could be modeled. This approach could help determine the best preventive or intervention measures to be applied prior to an outbreak of a pathogen of concern.

Table 1.

Advantages and disadvantages of currently used and proposed methods for integrating network and genomic approaches. Methods are listed in the order in which they appear in the text. A * indicates that the given method utilizes Bayesian approaches, which have the advantage of being able to incorporate uncertainty.

Method	Advantages	Disadvantages	References
SNA and pathogen strain-sharing network	Requires less intense pathogen sequencing effort. Able to identify if social structure is important for pathogen transmission, and/or the relative importance of different behaviors or locations. In future, may incorporate whole genome sequencing of pathogens, especially in the context of transmission tree reconstruction for higher resolution transmission networks.	Uses low resolution representation of pathogen relatedness with limited ability to represent direction of transmission. May have reduced ability to refine conclusions about impact of specific behaviors, locations, etc. on transmission.	Villaseñor-Sierra et al., 2007; VanderWaal et al., 2014; Marquetoux et al., 2016; Fountain-Jones, Packer, et al., 2017
Phylogenetic tree structure to describe network structure	Able to identify non-random mixing and differentiate some transmission patterns in underlying social network. In future, may be able to be used to parameterize theoretical networks for simulations and hypothesis testing.	Limited in conclusions that can be drawn. Identifying non-random mixing may not be novel in all systems. Poorly suited to dynamic networks.	Leventhal et al., 2012; Robinson et al., 2013; Colijn & Gardy, 2014; McCloskey et al., 2016; Vasylyeva et al., 2016
Transmission tree structure* to describe network structure	Able to identify non-random mixing in underlying social network, and potentially draw conclusions about management interventions. In future, may be able to be used to parameterize theoretical networks for simulations and hypothesis testing.	Limited in conclusions that can be drawn. Unable to reliably detect clustering in underlying network. High pathogen sequencing effort required. Unlikely to be the most effective integration of genomic approaches and contact networks.	Leigh Brown et al., 2011; Welch, 2011; Carnegie, 2017
Transmission tree reconstruction* with contact networks and/or phylogeographic models informing priors	Contact networks and/or phylogeographic models may help resolve some of the uncertainty involved in transmission tree reconstruction. Transmission trees then provide higher resolution data about transmission events.	Higher effort for sequencing, computational effort for reconstructing transmission trees. Generally requires pathogen with high mutation rate and high-intensity sampling effort.	Theoretical, but see Hall et al., 2016
SNA and pathogen transmission trees*	Transmission trees may better represent directionality of transmission than strain-sharing networks. May be able to identify social structure, behaviors, locations, etc. important for transmission. Higher resolution representation of pathogen relatedness than strain-sharing networks.	Higher effort for sequencing, computational effort for reconstructing transmission trees. Generally requires pathogen with high mutation rate and high-intensity sampling effort.	Theoretical
Transmission tree* of apathogenic infections to describe social network	Apathogenic infections may reveal transmission relevant contact prior to an epidemic with a pathogenic infection. May be useful for describing relevant intervention measures in at-risk wildlife populations. May be able to capitalize on non-invasive sampling.	Higher effort for sequencing, computational effort for reconstructing transmission trees. Generally requires agent with high mutation rate and high-intensity sampling effort. Requires agent with well-characterized mode of transmission. Most appropriate for directly transmitted, single-host pathogens. This approach is currently untested; transmission of pathogenic and apathogenic organisms may be too different to be able to translate mechanisms/risk factors between agents.	Theoretical, but see Plucíński et al., 2011; Bull et al., 2012; Blasse et al., 2013; Blyton et al., 2014; Francis et al., 2016; Springer et al., 2016