Endemic infection can shape exposure to novel pathogens: Pathogen co-occurrence networks in the Serengeti lions

Abstract
Pathogens are embedded in a complex network of microparasites that can collectively or individually alter disease dynamics and outcomes. Endemic pathogens that infect an individual in the first years of life, for example, can either facilitate or compete with subsequent pathogens thereby exacerbating or ameliorating morbidity and mortality. Pathogen associations are ubiquitous but poorly understood, particularly in wild populations. We report here on 10 years of serological and molecular data in African lions, leveraging comprehensive demographic and behavioural data to test if endemic pathogens shape subsequent infection by epidemic pathogens. We combine network and community ecology approaches to assess broad network structure and characterise associations between pathogens across spatial and temporal scales. We found significant non-random structure in the lion-pathogen co-occurrence network and identified both positive and negative associations between endemic and epidemic pathogens. Our results provide novel insights on the complex associations underlying pathogen co-occurrence networks.

Keywords
*Babesia*, calicivirus, canine distemper virus, co-infection, community assembly, coronavirus, feline immunodeficiency virus, parvovirus.

INTRODUCTION
Identifying and determining the nature of interactions between multiple pathogens is increasingly considered critical to understanding infectious disease dynamics (e.g. Pedersen & Fenton 2007; Graham 2008; Telfer et al. 2010; Johnson et al. 2015; Gorsich et al. 2018). Individuals are often co-infected by a diverse infra-community of pathogens, and interactions between pathogens can both alter infection patterns (Cattadori et al. 2008; Lass et al. 2013; Susi et al. 2015) and influence disease outcomes (Moss et al. 2008; Munson et al. 2008; Knowles 2011; Wejse et al. 2015). Pathogens infecting individuals in the first years of life may impact infection by subsequent pathogens (Fenton 2008; Randall et al. 2013; Rynkwicz et al. 2015; Aivelo & Norberg 2018; Budischak et al. 2018). For example endemic pathogens that compete for the same resources as epidemic pathogens and can reduce the likelihood of infection (Randall et al. 2013) or, conversely, facilitate infection via immune suppression (e.g. Geldmacher & Koup 2012). The sequence in which pathogens infect an individual or ‘priority effects’ have been experimentally shown to be important in shaping co-infection dynamics in a variety of systems (e.g. Hoverman et al. 2013; Halliday et al. 2017), yet are rarely demonstrated in non-experimental contexts. How priority effects and pathogen traits (e.g. transmission mode) affect the nature and frequency of associations between endemic and epidemic pathogens, ultimately shaping pathogen infra-communities is a knowledge gap that has significant consequences for understanding patterns of infection (Munson et al. 2008; Telfer et al. 2010; Ezenwa & Jolles 2015; Halliday et al. 2017).

Quantifying associations between pathogens from observational data and inferring interactions from these patterns, however, is a methodological challenge (Fenton et al. 2014). Discriminating between positive (i.e. two pathogens are more likely to occur together) or negative associations (i.e. two pathogens are less likely to occur together) between pathogens in populations is complicated by the short time window that a pathogen is shedding (and thus detectable with molecular methods) and by potentially confounding host immune environments (Tompkins et al. 2011). This is particularly the case for microparasites where pathogen detection often relies on serology, and, thus, without resampling the same individual, the precise timing of exposure cannot be estimated. Detection of pathogens that form chronic infections may be more straightforward as the infection is active for longer periods, but deducing pathogen associations is difficult without extensive longitudinal data (Fenton et al. 2014; Hellard et al. 2015). Identifying whether two pathogens are associated due to host–habitat preferences, the increasing likelihood of exposure with age, or are a product of a negative (e.g. competition) or positive (e.g. facilitation) interactions is
methodologically challenging (Poulin 2007; Johnson & Buller 2011; Fenton et al. 2014; Hellard et al. 2015; Clark et al. 2016). Identifying associations that could represent candidate interactions based on observational data can not only provide a basis for experiments to test potential interactions but also provide novel insights into pathogen infra-community dynamics.

Detecting associations between pathogens is also likely to depend on taxonomic and spatial scales that are seldom considered (Araújo & Rozenfeld 2014; Stutz et al. 2018). Studies commonly aggregate pathogen data to genus level, but associations between pathogens can be subtype or genotype-specific (e.g. Wejse et al. 2015; Benesh & Kalbe 2016; Brook et al. 2017). For example individuals infected with human immunodeficiency virus subtype 1 (HIV-1) are four times more likely to become co-infected with tuberculosis compared to individuals with HIV-2 (Wejse et al. 2015). Beyond subtype or genus, genotype-specific associations have been demonstrated in snails infected by trematodes (Louhi et al. 2015) and in rodents infected by Bartonella bacteria (Brook et al. 2017). Infra-community dynamics are also likely to vary with spatiotemporal scale. In general, associations between free-living species are more apparent at scales where interactions occur compared to broader spatiotemporal scales (Eltonian noise hypothesis; Peterson et al. 2011; Araújo & Rozenfeld 2014), but it remains unclear if this is true for pathogens. Nonetheless, for cross-sectional datasets, important patterns may be missed unless multiple spatiotemporal scales are considered (Ovaskainen et al. 2017). To overcome these challenges, analytical approaches that can quantify associations between pathogens, whilst controlling for potential confounding factors are required to assess the role of associations in shaping pathogen infra-communities.

Recent applications of network theory to parasite community ecology provide an opportunity to move beyond the pairwise associations between two pathogens (Clark et al. 2016; Aivelo & Norberg 2018; Stutz et al. 2018). Network measures have frequently been used to study food webs but are increasingly applied to pathogen infra-communities where nodes are pathogens, and edges represent pathogen co-occurrences within the host (Vaumourin et al. 2015). Networks are modular if pathogens co-occur more frequently in particular groups, ‘nested’ if pathogens frequently share interaction partners across the network, or ‘segregated’ if the inverse is true (Strona & Veech 2015; Ulrich et al. 2017). If, for example networks are segregated, targeted control of one ‘keystone’ pathogen may lead to co-extinction of other pathogens in a module (Pedersen & Fenton 2007; Säätäry et al. 2013). If a network is nested, perturbations to the pathogen infra-community may spread throughout the network (Griffiths et al. 2014).

Although pathogen co-occurrence networks are valuable for quantifying broad structural patterns, they do not account for environmental or host factors, pathogen traits or differences in spatial or temporal scale. Joint species distribution models (JSDMs) fill this gap by simultaneously assessing environmental influences and interspecific co-occurrences across multiple scales using hierarchical Bayesian mixed models (Warton et al. 2015; Ovaskainen et al. 2017). Here we use both co-occurrence networks and JSDMs to examine the structure of pathogen-pathogen networks and quantify pathogen associations while controlling for environmental/host factors and scales. We include information on pathogen traits such as transmission mode to assess what role they played in the distribution of each pathogen. We collate 10 years of cross-sectional data on endemic and epidemic pathogens in 105 African lions (Panthera leo) as well as extensive host and environmental data from the Serengeti Lion Project (SLP, Packer et al. 2005). The SLP datasets provide a unique opportunity to understand pathogen co-occurrence networks in a wild population while controlling for group, individual and environmental characteristics. We use this data to ask the following interlinked questions at two levels of taxonomic resolution:

(I) To what degree is the pathogens’ co-occurrence network of Serengeti lions nested or segregated?

(II) After accounting for environmental/host factors and spatiotemporal scale, is the type of endemic pathogen an individual is infected by early in life associated with exposure to epidemic pathogens later in life?

(III) Are there significant endemic–endemic or epidemic–epidemic pathogen co-occurrences?

Because we could not directly determine the order of infection events from cross-sectional data in isolation, we used age-prevalence relationships in combination with the natural history of each pathogen to estimate probable timing of events. We describe an analytical pathway that can assess broad network structure and quantify pathogen associations across multiple scales that can generally be applied to understand infectious disease dynamics. The co-occurrence network detects clusters of pathogen sharing amongst individuals and screens for disconnected nodes (pathogens that rarely co-occur with others), while the JSDM approach was used to quantify pathogen–pathogen associations. To assess the plausibility of these putative interactions, we compare our findings to similar mammalian pathogens in experimental studies. Detecting pathogen co-occurrences not only provides novel insights into pathogen infra-community dynamics but also helps aide surveillance efforts in the field and generate testable hypotheses that can be answered in laboratory experiments.

METHODS

Pathogen data

Serological testing and quantitative PCR (qPCR) were performed to detect endemic and epidemic pathogens from blood samples taken from lions in the Serengeti National Park, Tanzania from 1984 to 1994. In total, 394 individuals were sampled throughout this period, but our analysis was restricted to the 105 individuals tested for the full suite of 10 pathogens (Table 1: pathogen natural history; Table S1: number of individuals tested per year included in the analyses). Nomadic individuals (i.e. lions that were not resident in any pride) were excluded due to the difficulty of assigning environmental variables (see Confounding variables below). Serological data on
canine distemper virus (CDV), feline calicivirus, parvovirus and coronavirus has been published previously, except Rift Valley Fever (RVF) (Packer et al. 1999, see Table S2 for assay details). To detect RVF exposure we conducted a plaque reduction virus neutralizing test (PRNT) that quantified virus neutralizing antibodies from serum following Scott et al. (1986) protocol.

We used qPCR to identify nucleotides for feline immunodeficiency virus (FIV_{Ple}) and the protozoan pathogens in this study (Table 1). Three distinct subtypes of FIV_{Ple} co-circulate in Serengeti lions (Troyer et al. 2005, 2011; Antunes et al. 2008) and thus subtype specific qPCR was performed (see Troyer et al. 2004, 2005 for qPCR protocols). The resultant 300 base pair sequences from the pol gene were aligned and assigned to 21 operational taxonomic units/genotypes based on a 95% molecular similarity threshold (see Fountain-Jones et al. 2017 for details). Lions also commonly get infected by a rich protozoan fauna including Babesia and Hepatozoon genera. We developed quantitative PCR protocols using density gel gradient electrophores is to identify each protozoan species (see Munson et al. 2008).

We categorised each pathogen as likely endemic or epidemic in the lion population: endemic pathogens were considered to be constantly circulating and often infecting the young while epidemic pathogens sweep through the population every few years infecting all age classes (Packer et al. 1999; Penzhorn 2006; Troyer et al. 2011). Many of the pathogens have been previously classified as endemic or epidemic (Packer et al. 1999). We supported our classification with age-prevalence plots (Fig. S1) and we plotted yearly prevalence (Fig. S2) for the pathogens not previously classified. Pathogens with a high prevalence at a young age (≤ 2 years old) with little fluctuation across all years and age classes were considered to be likely endemic, whereas an increasing age-prevalence relationship and high temporal variation were classified as more likely to be epidemic in this population. Feline coronavirus can have epidemic and endemic cycles, and it is challenging to assess which form the individual was infected with from serological data, but based on age-prevalence relationships we categorised coronavirus as an endemic infection (Fig. S1). Furthermore, we used patterns of age-prevalence to infer the potential timing of infections. As most individual lions were likely to be infected by the pathogens we considered endemic within the first 2 years after birth (Troyer et al. 2011, Fig. S1), we assume that endemic exposure typically occurred prior to exposure by an epidemic pathogen. We partitioned the endemic pathogen data into two sets based on taxonomic resolution (high and medium). The high taxonomic resolution dataset encompassed FIV_{Ple} genotype and Babesia species data, whereas the medium resolution dataset aggregated FIV_{Ple} subtype information and Babesia data to genus level.

### Co-occurrence network

We examined pathogen co-occurrence patterns to evaluate preferential associations among pathogens. We constructed co-occurrence networks for each taxonomic resolution as well as for pathogens tested for using qPCR and by serology in cases combining both lines of diagnostic evidence led to altered network structure. To do so, we first built an $n \times n$ matrix that described presents/absences (i.e. occurrences) of both endemic and epidemic pathogens across individual lions, where $m$ was the number of individual lions and $n$ the number of pathogens. By multiplying it by its transpose, we then created a summary $n \times n$ co-occurrence matrix that described, for each pair of pathogens, the number of observed co-occurrences across all individual lions. Pathogens detected infrequently in this lion population were included in this analysis to help screen for pathogens disconnected in the network. The co-occurrence matrix was used to evaluate which pathogens were carried by the same individuals utilizing a modularity-based ‘greedy’ approach (Clauset et al. 2004). Measures of modularity aim to determine the adequacy of different

---

**Table 1** Traits of both endemic and epidemic pathogens in this study

| Pathogen Type                                      | Type (binary) | Trans. mode (categorical) | One host? (binary) | Immune sup. (binary) | Exposure timing?* (binary) | Test type (binary) |
|----------------------------------------------------|---------------|---------------------------|--------------------|----------------------|---------------------------|-------------------|
| **EPIDEMIC**                                       |               |                           |                    |                      |                           |                   |
| Feline calicivirus (calicivirus)†                   | Virus         | Direct/env                | N                   | NE                   | Epidemic year             | Serology          |
| canine distemper virus (CDV)                        | Virus         | Direct                    | N                   | Yes                  | Epidemic year             | Serology          |
| Feline panleukopenia (parvovirus)                   | Virus         | Vertical, direct/env      | N                   | Yes                  | Epidemic year             | Serology          |
| Rift valley fever (RVF)                             | Virus         | Vector (mosquito)         | N                   | Yes                  | Throughout life#           | Serology          |
| **ENDEMIC**                                        |               |                           |                    |                      |                           |                   |
| Feline enteric coronavirus (coronavirus)†           | Virus         | Direct/env                | N                   | U                    | Epidemic year             | Serology          |
| B. gibsoni                                          | Protozoa      | Vector (tick)             | N                   | NE                   | < 2 years old             | qPCR              |
| B. leo with insertion                               | Protozoa      | Vector (tick)             | N                   | NE                   | Throughout life            | qPCR              |
| B. felis                                            | Protozoa      | Vector (tick)             | N                   | NE                   | < 2 years old             | qPCR              |
| Hepatozoon felis                                    | Protozoa      | Vector (tick)             | N                   | NE                   | < 2 years old             | qPCR              |
| Feline immunodeficiency virus                       | Virus         | Vertical/direct           | Y                   | Yes                  | < 2 years old             | qPCR              |

Notes Trans.mode: Transmission mode (all pathogens can be horizontally transmitted). Immune sup.: Pathogen can suppress the immune system. Vertical: Vertical transmission is also possible. Env: Environmentally persistent. Direct: Transmission through host contact. Immune sup.: Immune suppression.

*Likely time of exposure.
†Determined by age-prevalence relationships (see Methods and Fig. S1) but can have endemic or epidemic variants. U: Unknown NE: No evidence.
#More likely after heavy rainfall (Fig. S2).
classification schemes in representing clusters and divisions in datasets; here, the clusters represented the co-occurrence of pathogens in individual hosts. Estimates of modularity were calculated for each possible classification by comparing the expected fraction of pathogen co-occurrence to random co-occurrences (Newman 2006). The classification with the highest modularity from all the generated classifications was selected.

We then computed a measure of network structure ($\bar{N}$) and modularity index based on node overlap and segregation (Strona & Veech 2015; Ulrich et al. 2017). $\bar{N}$ ranges from scaled from $-1$ (entirely segregated network) to 1 (entirely nested network). These analyses were performed in R using the ‘igraph’ and ‘nos’ libraries (Csárdi & Nepusz 2006). The co-occurrence matrix was obtained from the incidence matrix using the graph.incidence and the bipartite.projection functions. The classification analysis was performed using the fastgreedy.community function in igraph (Csárdi & Nepusz 2006).

Joint species distribution modelling

Joint species distribution models are a flexible multivariate extension of generalised linear mixed models that can examine how environment (and host) shape multiple species simultaneously across biological scales (Ovaskainen et al. 2017; Björk et al. 2018). JSDMs can quantify associations between species across scales using latent factor models to estimate species–species covariance for each random effect (Ovaskainen et al. 2017; Björk et al. 2018). We fitted JSDMs for both high and medium taxonomic resolution datasets, combining information on environmental and host covariates as fixed effects (see Confounding variables below for details), to the occurrence data for each of the pathogens. Pathogens detected fewer than five times were excluded from this analysis leaving 10 pathogens in the medium taxonomic model and 17 in the high-resolution dataset. Including pathogens with fewer than five occurrences may lead to spurious associations (Ovaskainen et al. 2017). We fitted all the JSDMs with Bayesian inference, using ‘Hierarchical Modelling of Species Communities’ (Blanchet et al. 2018). For each analysis, we modelled the response pathogen co-occurrence matrix using a probit model based on the approach outlined in Ovaskainen et al. (2016). In contrast to the network approach, the JSDM co-occurrence matrix is a product of the pathogen-to-pathogen variance-covariance matrix estimated for each random effect (e.g. pride-year) in the model. Each random effect (and thus each estimated co-occurrence matrix) measures a component of the variation in the response that is different than the other random effects and of the set of explanatory variables (fixed effects) considered in the model. In our models, we added individual (e.g. sex and age), pride and environmental characteristics (see Confounding variables below) as fixed effects. Individual sampled, pride-year (i.e. which pride and year the individual was sampled in) and year-landscape (i.e. what year was the individual sampled in the Serengeti) sampled were added as random effects. As pathogen traits may shape the distribution of each pathogen (e.g. similar environmental and host variables may shape tick-borne pathogens), we included traits such as pathogen type (see Table 1) in each analysis. We utilised the default priors (described in full detail in Ovaskainen et al. 2017) and ran the HMSC model twice using 3 million MCMC samples (the first 300 000 of which being burn-in). Each run was carried out using a different seed. Visual inspection of MCMC traces and the Gelman–Rubin diagnostic calculated to assess convergence. In addition, we made sure that the effective sample size (ESS) of each parameter was > 200.

Confounding variables

As part of the SLP, most of the individuals in this study have been regularly observed since birth (Mosser & Packer 2009). We selected 13 predictor variables that we thought were likely to be important for pathogen exposure and thus could confound possible associations patterns (Table 2). We included variables that captured individual variability (e.g. age at sampling), and pride characteristics including environmental variables (e.g. average vegetation cover of the pride’s territory; see Table 2 for measurement details).

RESULTS

The Serengeti lions were exposed to an average of 5 pathogens (two epidemic and three endemic, SD = 1); one individual had been infected by 9 of 10 pathogens (based on medium resolution data, Fig. S3). Cubs between 1 and 2 years old were often already infected with an average of four pathogens (SD = 1), with one 1.5 years old cub positive for 5. All lions were qPCR positive for at least one protozoan species, and 25% of them were infected by all four protozoans tested.

Pathogen co-occurrence networks are highly nested

The high taxonomic resolution summary network indicated a significantly nested architecture ($\bar{N} = 0.74$) with relatively low modularity (modularity index $= 0.393$, $z = 3.307$, $P \leq 0.001$) with three clusters (Fig. 1a). The largest cluster (green nodes) included all of the protozoans, epidemic pathogens and some FIVPle genotypes, whereas the remaining two clusters consisted of FIVPle genotypes (Fig. 1a). When we modelled networks based on diagnostic test, the general pattern did not substantially change, with the exception that RVF clustered separately from the other viruses detected using serology. In both network formulations, phylogenetically similar genotypes of FIV did not cluster together (Fountain-Jones et al. 2017, see Fig. S5). In contrast, the medium resolution network was completely nested with no modularity ($\bar{N} = 1$, modularity index $= 0$, $z = \infty$, $P = 0$) and no significant clusters (Fig. 1b).

Strong associations between endemic and epidemic pathogens

After accounting for environmental, individual and pride factors and scale, the JSDM analysis identified strong associations between pathogens (Fig. 2) that were not detected in the summary co-occurrence network. Including individual, pride-
The strongest associations between endemic and epidemic pathogens were detected at the lowest spatial-temporal resolution (landscape-year). In the high taxonomic resolution model, pathogens separated into two groups with each group having a very similar association profile. One group was characterised by positive associations between two FIVPle genotypes (C1 & B2), coronavirus, *Hepatozoon felis* and calicivirus (Figs. 2c and Fig. S6c). There were strong negative associations between pathogens in each separate group (e.g. CDV and FIVPle C). Generally, the same associations held in the medium resolution model, we detected a positive association between *H. felis* and FIVPle C not found in the high-resolution model indicating that FIVPle subtype, but not genotype, was important for this association (Fig. 2b and Fig. S6b). Strikingly, we found that FIVPle subtypes had contrasting association profiles. At the individual level, FIVPle B and C were negatively associated with each other, and FIVPle C was positively associated with coronavirus, whereas FIVPle B was negatively associated with coronavirus (Fig. 2a and Fig. S6a). Both high and medium taxonomic resolution JSDMs had reasonable explanatory power (Tjur $R^2 = 0.381 & 0.330$, respectively). In both models, the landscape and host factors that explained the distribution of each pathogen were not predicted well by pathogen traits (Fig. S8). See Fig. 3 for a summary of all of the associations detected across scales from our cross-sectional data and Figs S9/S10 for model details.

**DISCUSSION**

Here we demonstrate non-random associations in the pathogens infecting wild African lions, with both negative and positive associations detected between endemic and epidemic pathogens. While there was minimal structure in the co-occurrence network (Fig. 1a), we uncovered structure after accounting for scale and controlling for potentially confounding environmental and host variables (Fig. 2). Using age-prevalence relationships we could assess the likely order of infection using cross-sectional data. We found that the particular endemic pathogen an individual is infected by as a cub may have consequences for which novel epidemic pathogen the individual is infected with later in life (Fig. 3). We emphasise that the approach used here can start to untangle pathogen infra-community relationships and identify potential endemic–epidemic associations in wild populations. These can then be compared with knowledge of pathogen pathogenesis and validated *in-vitro* in a laboratory setting.
laboratory studies of co-infection in lions are rare for good reason, the associations we found have clear precedence in similar pathogens co-infecting humans and represent plausible interactions. Our results not only provide new insights on pathogen community structure in the Serengeti lions but also provide a valuable framework for exploring pathogen co-occurrence networks and infra-community dynamics.

Co-occurrence networks were highly nested with relatively low modularity, particularly at a medium taxonomic resolution. Nonetheless, RVF did cluster separately from the other pathogens tested via serology which potentially indicates that RVF, unlike the other epidemic viruses, has a distinct epidemic cycle with most of the interacting partners being more chronic pathogens. This is supported by the RVF association profile detected in our JSDM analysis and is intuitive given that RVF is the only mosquito-borne pathogen that we sampled. Even though we sampled pathogens considered important for lion health, we lacked data on other potentially pathogenic bacteria, helminths and fungi that the lions were exposed to or potentially infected by. Furthermore, symbiont interactions can also be important in shaping pathogen dynamics (e.g. Halliday et al. 2017) and could be considered in pathogen infra-community studies. These additional taxa may lead to further segregation in the network, as larger and more diverse networks typically show increased modularity and segregation (Thebault & Fontaine 2010; Sauve et al. 2014). Expanding sampling to construct a more complete microbe and macroparasite network would also capture a broader array of potentially facilitative and competitive associations (Ezenwa 2016; Aivelo & Norberg 2018).

After accounting for environment, host and scale, we found that the endemic pathogens were strongly associated with the epidemic pathogens and, based on mammalian laboratory-based experiments, suggest that these patterns represent plausible interactions between pathogens. For example we detected negative associations between endemic pathogens (FIVple B and H. felis) and RVF after accounting for differences between individuals. Co-infections between bunyaviruses like RVF and retroviruses are likely common in humans and wildlife, though there are surprisingly few studies addressing the topic. In contrast, relationships between dengue virus (a flavivirus) and HIV are relatively well understood. Flaviviruses and HIV share similar immune receptors that can inhibit HIV replication and the molecular machinery used to do so may be a viable way to control HIV infection (e.g. Xiang et al. 2009). Given the overall structural similarity of flaviviruses and bunyaviruses (Hernandez et al. 2014), it is possible that a similar mechanism underlies the association in lions between RVF and FIVple that we observed, although we show that this association was subtype specific. If this was true, RVF might inhibit FIVple B infection –
counter to our assumption that endemic pathogens in our system infected each individual first (Fig. 3).

The greatest number of associations between epidemic and endemic pathogens were detected when we included differences across years (year-landscape scale) in our analysis. These associations could represent plausible facilitative or competitive interactions. CDV and Babesia are well-known to interact with high levels of Babesia infection magnifying the impacts of consequent T-cell depletion caused by CDV infection leading to mortality of nearly 40% of the lion population in 1994 (Munson et al. 2008). We found that all tick-borne haemoparasites showed positive associations with CDV including B. leo (with insert) despite its low prevalence in 1993/1994 (Fig. S9). Parvovirus was also positively associated with CDV, but this was likely due to similarities in timings of epidemics with a parvovirus epidemic in 1992 just before the 1994 CDV epidemic (Packer et al. 1999). Parvoviruses are also immune suppressive, and so the timing of the parvovirus outbreak may also have contributed to the CDV/Babesia-induced mortality. The general negative relationship between FIVPle C and CDV/Babesia supports the theory that individuals infected by subtype C were more likely to die in the consequent Babesia/CDV outbreak (Troyer et al. 2011). Thus, this negative association may not be due to competition between pathogens but rather to mortality.

Our approach detected strong associations between the endemic pathogens also. For example there were opposing associations between the FIVPle subtypes and coronavirus (Fig. 2). Negative associations between retroviruses and coronaviruses are rarely reported, yet there are plausible molecular pathways. HIV-1 and human coronaviruses (HCoV) share remarkably similar binding receptors (Chan et al. 2006) and some mild HCoV strains are even considered a viable vaccine against HIV (Eriksson et al. 2006). This may explain the negative association we detected for FIVPle B and coronavirus but does not explain the positive association between FIVPle C and coronavirus we detected across scales. The mechanism driving FIVPle subtype specific relationships with coronaviruses remain unclear, and as coronaviruses infecting lions are also likely to be genetically diverse, examining the genetic structure of coronavirus may help untangle these associations further. In contrast, competitive associations between HIV strains are well characterised with HIV-1 found to outcompete HIV-2 for blood resources (Ariën et al. 2005). For FIVPle, even though co-infection is relatively common (Troyer et al. 2011) competition between subtypes could be important as there is anecdotal cell culture evidence that FIVPle B can propagate more rapidly than FIVPle C (M. Roelke, unpublished data).

There were also contrasting associations between the protozoan species. For example the distribution of B. felis was not shaped by any other protozoan and in general, had a narrow association profile (Fig. 2), unlike the other Babesia species. For the individuals co-infected by protozoans, associations involving B. felis were also common, whereas co-infections involving H. felis and the other Babesia species varied in prevalence and composition (Fig. S9). Even though B. gibsoni and B.
show similar age prevalence profiles (Fig. S1), the prevalence of *B. felis* over time was relatively stable compared to the other protozoa (Fig. S10). Differences in the host range for individual *Babesia* species and potential host differences in virulence may partially explain these patterns. For example, *B. felis* has only ever been detected in felids, whereas *B. gibsoni* has a much broader host range including canids (Penzhorn 2006). Generalist pathogens may have greater pathogenicity as there can be reduced selective restraint on virulence particularly in ‘dead end’ hosts (Woolhouse *et al.* 2001). If more pathogenic species are more likely to interact with other pathogens compared to less virulent pathogens, it is an open question in disease ecology. Importantly, patterns like these would be missed without incorporating high-resolution pathogen data.

There are, however, limitations to this approach. The inability to distinguish mortality or correlated exposure (i.e. an individual is infected by multiple pathogens in the same transmission event) from negative or positive associations is one of them, and careful interpretation of associations is necessary. Incorporating approaches such as structural equation models that explicitly include potential mechanisms that underlie candidate pathogen associations (Carver *et al.* 2015) could be a valuable additional step in future pathogen network studies. Another weakness is the inability to estimate the timing of these infections more precisely. For example the negative association between RVF and *H. felis* could be due to temporal differences when ticks and mosquitoes emerge after rains. Years with higher rainfall increase mosquito abundance thus increasing RVF prevalence (Fig. S2), whereas ticks emerge *en masse* when rains follow a dry period potentially increasing *H. felis* prevalence (Munson *et al.* 2008). As rainfall was calibrated to the year of sampling rather than the age of infection (which could differ) the JSDM approach could not capture this variation. Studies using longitudinal data to quantify associations using a similar framework to ours (e.g. Telfer *et al.* 2010; Henrichs *et al.* 2016) will be beneficial as they are likely to provide more robust estimates of the order of infection in wild populations. Furthermore, we cannot quantify the importance of these associations in shaping pathogen distribution across scales compared to processes such as host density. Lastly, incorporating immune function and host resources in both the summary network and JSDM analyses are likely to provide mechanistic insight into pathogen network structure (Griffiths *et al.* 2014). Higher resolution pathogen traits, such as duration of infection, are likely to provide further mechanistic insight into how and why pathogens co-occur as they do in free-living communities (Ulrich *et al.* 2017). However, given the daunting complexity of pathogen infra-community dynamics, our two-step approach can assess broad network structure and identify useful candidate interactions between pathogens thereby reducing some of this complexity.

The high frequency of co-occurrence and co-infection in lions – and the potential for specific associations to cause population decline – highlights the importance of understanding pathogen associations. The lion pathogen co-occurrence network was highly connected with both positive and negative associations between endemic and epidemic pathogens. Our findings indicate that the lion pathogen infra-community is influenced by a number of ecological factors and associations between pathogens. We identify useful associations between pathogens thereby reducing some of this complexity. More broadly, our work demonstrates how different network approaches can be combined to gain insights into the ecological factors underlying pathogen associations and how this can be applied to the study of pathogen communities in wildlife populations. In addition to these biological insights, the study highlights several critical areas for methodological improvement.

Fig. 3 Summary of the strong positive (red lines/arrows) and negative (blue lines/arrows) associations between endemic (grey circles) and epidemic (orange circles) pathogens in the Serengeti lions; dark-grey borders indicate protozoa. The direction of the red or blue arrows indicates the potential sequence of infection events. The black arrow along the *X*-axis represents age; the circles reflect the ages when lions were likely to be infected by each pathogen (based on age-exposure data rather than longitudinal data, see Fig. S1). Dashed circles indicate major co-occurrence clusters identified at the landscape-year scale.
that can currently limit robust inference of pathogen associations from cross-sectional serological and qPCR data. Addressing these limitations is timely, given the ongoing threat of wildlife population decline, creating a need to integrate better molecular, ecological and network information for disease control.

ACKNOWLEDGEMENTS

N.M. F-J. and M.E.C were funded by National Science Foundation (DEB-1413925 and 1654609), the University of Minnesota’ 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. We thank the three anonymous reviewers for their constructive feedback on this paper. We also thank Dr. Linda Munson who led the CDV-Babesia work and Professor CJ Peters for performing the RVF virus neutralisation tests.

AUTHOR CONTRIBUTIONS

NMFJ and MEC designed the study. CP and KT provided data. NMFJ, MJ and FGB conducted statistical analyses. NMFJ wrote the manuscript and all authors contributed to revisions.

DATA ACCESSIBILITY STATEMENT

Data available from the Figshare Repository: https://doi.org/10.6084/m9.figshare.7742900.

REFERENCES

Aivelo, T. & Norberg, A. (2018). Parasite-microbiota interactions potentially affect intestinal communities in wild mammals. J. Anim. Ecol., 87, 438–447.

Antunes, A., Troyer, J.L., Roelke, M.E., Pecon-Slattery, J., Packer, C., Winterbách, C. et al. (2008). The evolutionary dynamics of the lion Panthera leo revealed by host and viral population genomics. PLoS Genet., 4, e1000251.

Araújo, M.B. & Rozenfeld, A. (2014). The geographic scaling of biotic interactions. Ecography (Cop.), 37, 406–415.

Arień, K.K., Abrahá, A., Quitones-Mateu, M.E., Kestens, L., Vanham, G. & Arts, E.J. (2005). The replicative fitness of primary human immunodeficiency virus type 1 (HIV-1) group M, HIV-1 group O, and HIV-2 isolates. J. Virol., 79, 8979–8990.

Benes, D.P. & Kalbe, M. (2016). Experimental parasite community ecology: intraspecific variation in a large tapeworm affects community assembly. J. Anim. Ecol., 85, 1004–1013.

Björk, J.R., Hui, F.K.C., O’Hara, R.B. & Montoya, J.M. (2018). Uncovering the drivers of host-associated microbiota with joint species distribution modelling. Mol. Ecol., 27, 2714–2724.

Blanchet, F.G., Tikhonov, G. & Norberg, A. (2018). HMSC: hierarchical modelling of species community. R package version 2.2-1.

Brook, C.E., Bai, Y., Yu, E.O., Ranaivoson, H.C., Shin, H., Dobson, A.P. et al. (2017). Elucidating transmission dynamics and host-parasite-vector relationships for rodent-borne Bartonella spp. in Madagascar. Epidemics, 20, 56–66.

Budischak, S.A., Wiria, A.E., Hamid, F., Wammes, L.J., Kaisar, M.M.M., van Lieshout, L. et al. (2018). Competing for blood: the ecology of parasite resource competition in human malaria-helminth co-infections. Ecol. Lett., 21, 536–545.

Carver, S., Beatty, J.A., Troyer, R.M., Harris, R.L., Stutzman-Rodriguez, K., Barrs, V.R. et al. (2015). Closing the gap on causal processes of infection risk from cross-sectional data: structural equation models to understand infection and co-infection. Parasit. Vectors, 8, 658.

Cattadori, I.M., Boag, B. & Hudson, P.J. (2008). Parasite co-infection and interaction as drivers of host heterogeneity. Int. J. Parasitol., 38, 371–380.

Chan, V.S.F., Chan, K.Y.K., Chen, Y., Poon, L.L.M., Cheung, A.N.Y., Zheng, B. et al. (2006). Homozygous L-SIGN (CLEC4M) plays a protective role in SARS coronavirus infection. Nat. Genet., 38, 38–46.

Clark, N.J., Wells, K., Dimitrov, D. & Clegg, S.M. (2016). Co-infections and environmental conditions drive the distributions of blood parasites in wild birds. J. Anim. Ecol., 85, 1461–1470.

Clausen, A., Newman, M.E.J. & Moore, C. (2004). Finding community structure in very large networks. Phys. Rev. E, 70, 066111.

Csárdi, G. & Nepusz, T. (2006). The igraph software package for complex network research. InterJournal Complex Syst., 1695, 38.

Eriksson, K.K., Makia, D., Maier, R., Ludewig, B. & Thiel, V. (2006). Towards a coronavirus-based HIV multigene vaccine. Clin. Dev. Immunol., 13, 353–360.

Ezenwa, V.O. (2016). Helminth-microparasite co-infection in wildlife: lessons from ruminants, rodents and rabbits. Parasite Immunol., 38, 527–534.

Ezenwa, V.O. & Jolles, A.E. (2015). Opposite effects of anthelmintic treatment on microbial infection at individual versus population scales. Science, 347, 175–177.

FAO & IIASA (2009). Harmonized World Soil Database. Food and Agriculture Organization, Rome, Italy and IIASA, Laxenburg, Austria.

Fenton, A. (2008). Worms and germs: the population dynamic consequences of microparasite-macroparasite co-infection. Parasitology, 135, 1545–1560.

Fenton, A., Knowles, S.C.L., Petchey, O.L. & Pedersen, A.B. (2014). The reliability of observational approaches for detecting interspecific parasite interactions: comparison with experimental results. Int. J. Parasitol., 44, 437–445.

Fountain-Jones, N.M., Packer, C., Troyer, J.L., VanderWaal, K., Robinson, S., Jacquot, M. et al. (2017). Linking social and spatial networks to viral community phylogenetics reveals subtype-specific transmission dynamics in African lions. J. Anim. Ecol., 86, 1469–1482.

Geldmacher, C. & Koup, R.A. (2012). Pathogen-specific T cell depletion and reactivation of opportunistic pathogens in HIV infection. Trends Immunol., 33, 207–214.

Gorsich, E.E., Etienne, R.S., Medlock, J., Beechler, B.R., Spaan, J.M. & Spaan, R.S., et al. (2018). Opposite outcomes of coinfection at individual and population scales. Proc. Natl Acad. Sci. U. S. A., 115, 7545–7550.

Graham, A.L. (2008). Ecological rules governing helminth-microparasite coinfection. Proc. Natl Acad. Sci. U.S.A., 105, 566–570.

Griffiths, E.C., Pedersen, A.B., Fenton, A. & Petchey, O.L. (2014). Analysis of a summary network of co-infection in humans reveals subtype-specific transmission dynamics in African lions. J. Anim. Ecol., 86, 1469–1482.

Halliday, F.W., Umbanhowar, J. & Mitchell, C.E. (2017). Interactions among symbionts operate across scales to influence parasite epidemics. Ecol. Lett., 20, 1285–1294.

Hellard, E., Fouchet, D., Vavre, F. & Pontier, D. (2015). Parasite-parasite interactions in the wild: how to detect them? Trends Parasitol., 31, 640–652.

Henrichs, B., Oosthuizen, M.C., Troskie, M., Gorsich, E., Gondhalekar, C., Beechler, B.R. et al. (2016). Within guild co-infections influence parasite community membership: a longitudinal study in African buffalo. J. Anim. Ecol., 85, 1025–1034.

Hernandez, R., Brown, D.T. & Paredes, A. (2014). Structural differences observed in arboviruses of the alphavirus and flavivirus genera. Adv. Virol., 2014, 259382.

Hoverman, J.T., Hoye, B.J. & Johnson, P.T.J. (2013). Does timing matter? How priority effects influence the outcome of parasite interactions within hosts. Oecologia, 173, 1471–1480.

Johnson, P.T.J. & Buller, I.D. (2011). Parasite competition hidden by correlated coinfection: using surveys and experiments to understand parasite interactions. Ecology, 92, 535–541.
buffer the response of Serengeti predators to El Niño Southern Oscillation. Ecology, 94, 1123–1130.

Strona, G. & Veech, J.A. (2015). A new measure of ecological network structure based on node overlap and segregation. Methods Ecol. Evol., 6, 907–915.

Stutz, W.E., Blaustein, A.R., Briggs, C.J., Hoverman, J.T., Rohr, J.R. & Johnson, P.T.J. (2018). Using multi-response models to investigate pathogen co-infections across scales: insights from emerging diseases of amphibians. Methods Ecol. Evol., 9, 1109–1120.

Susi, H., Barrès, B., Vale, P.F., Laine, A.-L., Mideo, N., Alizon, S. et al. (2015). Co-infection alters population dynamics of infectious disease. Nat. Commun., 6, 9795.

Teller, S., Lambin, X., Bittes, R., Beldomenico, P., Burthe, S., Paterson, S. et al. (2010). Species interactions in a parasite community drive infection risk in a wildlife population. Science, 330, 243–246.

Thebault, E. & Fontaine, C. (2010). Stability of ecological communities and the architecture of mutualistic and trophic networks. Science, 329, 853–856.

Tomkins, D.M., Dunn, A.M., Smith, M.J. & Teller, S. (2011). Wildlife diseases: from individuals to ecosystems. J. Anim. Ecol., 80, 19–38.

Troyer, J.L., Pecon-Slattery, J., Roelke, M.E., Black, L., Packer, C. & O’Brien, S.J. (2004). Patterns of feline immunodeficiency virus multiple infection and genome divergence in a free-ranging population of African lions. J. Virol., 78, 3777–3791.

Troyer, J.L., Pecon-Slattery, J., Roelke, M.E., Johnson, W., VandeWoude, S., Vazquez-Salat, N. et al. (2005). Seroprevalence and genomic divergence of circulating strains of feline immunodeficiency virus among Felidae and Hyaenidae species. J. Virol., 79, 8282–8294.

Troyer, J.L., Roelke, M.E., Jespersen, J.M., Baggett, N., Buckley-Beason, V., MacNulty, D. et al. (2011). FIV diversity: FIV *p4* subtype composition may influence disease outcome in African lions. Vet. Immunol. Immunopathol., 143, 338–346.

Ulrich, W., Kryszewski, W., Sewerniak, P., Puchalka, R., Strona, G. & Gotelli, N.J. (2017). A comprehensive framework for the study of species co-occurrences, nestedness and turnover. Oikos, 126, 1607–1616.

Vaumourin, E., Vourc’h, G., Gasqui, P., Vayssier-Taussat, M., Windsor, D., Anderson, R. et al. (2015). The importance of multiparasitism: examining the consequences of co-infections for human and animal health. Parasit. Vectors, 8, S45.

Warton, D.I., Blanchet, F.G., O’Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C. et al. (2015). So many variables: joint modeling in community ecology. Trends Ecol. Evol., 30, 766–779.

Wejse, C., Patsche, C.B., Kühle, A., Bamba, F.J.V., Mendes, M.S., Lembik, G. et al. (2015). Impact of HIV-1, HIV-2, and HIV-1 + 2 dual infection on the outcome of tuberculosis. Int. J. Infect. Dis., 32, 128–134.

Westgate, M.J. (2016). circleplot: Circular plots of distance and association matrices. R package version 0.4.

Woolhouse, M.E.J., Taylor, L.H. & Haydon, D.T. (2001). Population biology of multihost pathogens. Science, 292, 1109–1112.

Xiang, J., McLinden, J.H., Rydze, R.A., Chang, Q., Kaufman, T.M., Klimzak, D. et al. (2009). Viruses within the Flaviviridae decrease CD4 expression and inhibit HIV replication in human CD4+ cells. J. Immunol., 183, 7860–7869.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.