INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH AND STUDIES

VOLUME04ISSUE04

''ANALYSIS OF MULTI-STRAIN BARTONELLA PATHOGENS IN NATURAL HOST POPULATION — DO THEY BEHAVE AS SPECIES OR MINOR GENETIC VARIANTS''

P. HOSOLY, D. TEGNE, V. KZIANNABOO, E. PERSTON, O. JAMES

ABSTRACT

Modern advances in genetic analysis have made it feasible to determine the variant form of a pathogen infecting a bunch. Classification of pathogen variants is usually performed by clustering analysis of the observed genetic divergence among the variants. A natural question arises whether the genetically distinct variants are epidemiologically distinct. A broader question is whether or not the various variants constitute separate microbial species or represent minor variations of the identical species. These important issues were addressed within the context of analyzing the dynamics of genetically distinct variants of Bartonella bacteria in gnawing animal hosts. Frequencies of acquiring a brand new variant were measured about the genetic differences between variants successively infecting a private rodent host. Two statistical techniques were introduced for performing such analysis, and also the methodologies were illustrated with a collection of information collected from a specific multi-strain Bartonella system. We dole out a frequency analysis of co-infection patterns and a Markoff chain analysis of panels of successive mixed infection statistic for testing some particular gene-based grouping of the Bartonella variants with a panel of observed disease data from a rodent population. Our analysis suggests that the three genogroups A, B, and C of Bartonella function as independent species but the variants within each genogroup enjoy some cross-immunity against one another. The newly developed methodologies are broadly applicable for analyzing other multi-strain pathogen data which are increasingly collected for diverse infectious diseases.

KEYWORDS

Cross-immunity mixed infection Multi-strain epidemic model Pathogen population Species co-existence

INTRODUCTION

Two questions are central to contemporary communicable disease epidemiology: how infection fluctuations arise and the way pathogen diversity is generated and maintained (Kurtenbach et al., 2006). The prevailing trend towards the outline of temporal and spatial dynamics of various infectious diseases has resulted in many publications, while mathematical modeling of multi-strain pathogens remains a serious challenge. Whereas previously the overwhelming majority of communicable disease models considered one strain, in recent years attention has focused on the interaction of multiple strains or perhaps different pathogen species (Read and Taylor, 2001, Gog and Grefell, 2002, Eames and Keeling, 2006). With recent advances in genetic analysis, it's been found that a disease-causing microbe population might carries with it multiple variants which it's feasible to spot which variants infect a specific host (Tibayrenc and Ayala, 2000).

Traditional definition of bacterial species supported results of DNA–DNA hybridization and outline of phenotypic traits is being replaced more and more often by gene sequence-based criteria. This approach is very important for the rapid characterization of fastidious or non-culturable bacteria. for instance, a comparison of DNA sequence data has served for the identification of Bartonella species more commonly than biochemical reactions due to the fastidious and comparatively inert nature of those bacteria (La Scola et al., 2003). Classification of pathogen variants is often performed by clustering analysis of the observed genetic divergence between the variants (Holmes et al., 1995). it's observed that strains as genetic variants tend to be organized into groups (clusters) (Gog and Grenfell, 2002). The classical example is influenza, where there are several circulating subtypes with many minor variants within each subtype (Andreasen et al., 1997). Cluster structure depends on levels of cross-immunity and, in some cases, on initial conditions specified for the algorithm (Gupta et al., 1998, Calvez et al., 2005). Clusters, once formed, are generally stable, and behave regularly in contrast to the chaotic behavior of the individual strains that allowed Calvez et al. (2005) to propose a

commonality between strain clustering and pattern formation. However, a fundamental question arises whether the genetically distinguished variants correspond to ecologically and epidemiologically distinct variants. Knowledge of the genetic structure and relationships between genetic variants is vital to grasp and predict the responses of pathogen populations to selective pressures imposed by host immunity (Levin et al., 1999, Gog and Grefell, 2002). A related question is whether or not an outsized genetic distance corresponds to the emergence of epidemiologically separate pathogen species within the sense that they need little or no cross-immunity.

Understanding the dynamics of genetically and anti-genically variable pathogens is a particularly important question since it applies to such dangerous diseases as influenza, malaria, and meningitis. The existence of multiple strains of a pathogen can alter host-microbe interactions and might need interesting implications on the epidemiological dynamics of an communicable disease (Read and Taylor, 2001). The critical feature of most multi-strain pathogens is that infection by one strain induces partial immunity to future infections by other strains (Abu-Raddad and Ferguson, 2004). The cross-immunity might cause a competitive interaction between strains, which when in addition to differences in transmission dynamics leads to selection between strains within a number population (Gupta et al., 1998). The mechanisms of cross-immunity are still not well understood for several host-pathogen systems, while a range of modeling approaches are applied to some infections, e.g. influenza, which greatly contributed to our understanding of the evolution of this infection, the interactions between different strains within a bunch population and also the mechanism of antigenic drift (Ackerman et al., 1990, Nelson and Holmes, 2007). Mathematical modeling of co-circulating influenza strains showed a clear stage of a multi-strain equilibrium and sustained oscillations caused by an overshoot within the immunity to a particular strain if cross-immunity between two strains is sufficiently strong (Andreasen et al., 1997). Another example of research during this direction is that the interference between immunologically independent pathogens accountable for fatal childhood diseases (Rohani et al., 2003).

We address the preceding questions with an in depth analysis of the epidemiological structure of a multi-strain system with real field and experimental data on diverse Bartonella strains. The genus Bartonella includes a range of genetically related Gram-negative bacteria that parasitize erythrocytes or epithelial cells of a spread of mammalian hosts including man. Many new species of Bartonella are isolated and characterized from diverse species of rodents and lots of other animals within the past decade (Breitschwerdt and Kordick, 2000, Jardine et al., 2005). thanks to their wide distribution and human disease association, Bartonella species have recently been recognized as emerging pathogens (Anderson and Neuman, 1997, Breitschwerdt and Kordick, 2000). Moreover, Bartonella species can provide a very good model for studying the ecology and evolution of zoonotic bacteria thanks to their wide distribution, high prevalence in animals, diversity of species, evidence for co-existence of multiple strains, and an especially high genetic heterogeneity of the bacterial population.

A longitudinal marked-recapture study conducted in Georgia, the USA in 1996–1997 has allowed examination of the dynamics of Bartonella strains related to one rodent species, the gnawing animal Sigmodon hispidus. the info derived from this study were partially analyzed and published (Kosoy et al., 2004a, Kosoy et al., 2004b). Some results relevant to the present analysis are provided within the following section.

To disentangle the epidemiological interactions of multiple strains observed during this study may require extensive modeling, and therefore the fundamental questions on the rank of the differences between detected Bartonella strains have remained unanswered partly due to the shortage of adequate statistical methodology. Some simplification is crucial for reducing the modeling complexity by initially using simple methods to explore the epidemiological nature of the genetically defined variants.

Here, we consider the question of whether strains from genetically distant clusters may be considered belonging to distinct taxonomic units of Bartonella within the sense that they represent different biological traits including little or no cross-immunity. A related

issue is whether or not strains within a cluster are minor variants representing population variations within one taxonomic unit a minimum of with relation to the presence of cross-immunity between these strains. Experimental studies could also be potentially limited because it could be impossible to estimate the degrees of overlapping ecological niches and cross-immunity among growing Bartonella genetic variants. an alternate way is to pick parameters that may characterize ecological interactions between distant variants and their qualitative analysis. Specifically, we propose measuring frequencies of acquiring a replacement variant in relevance the genetic differences between variants during a successively infected individual rodent host. To perform such analysis, we introduce two statistical approaches for studying the questions discussed earlier and illustrate the methodologies employing a set of knowledge collected from a specific multi-strain Bartonella system.

An outline of the remainder of the paper follows. within the Bartonella data, we briefly summarize the monitor program during which the Bartonella data was collected within the field. The blood sample of an infected host may contain one strain of Bartonella, or it should contain multiple co-existing Bartonella strains. Mixed infection data provide a chance to assess the hypothesis of limited, or absent, cross-immunity between two strains of the pathogen. Some biological hypotheses underlying mixed infections are discussed in Hypotheses for mixed infections. In Assessing the independent-variant hypothesis, we propose a frequency analysis method to assess the no cross-immunity hypothesis. within the Bartonella field studies, some rats were trapped repeatedly (Fig. 1), yielding data on their mixed infection histories. The degree of cross-immunity may additionally be investigated by studying the dynamical patterns of mixed infections. In Assessing the species-variant hypothesis by Markoff process analysis, we propose a Markoff chain approach to assess the cross-immunity structure of a multi-strain system and illustrate the tactic using the prevailing Bartonella data. Throughout the paper, we describe newly developed statistical methods within the context of analyzing the Bartonella data. However, the proposed methods are often equally applicable to other pathogens. We conclude briefly finally.

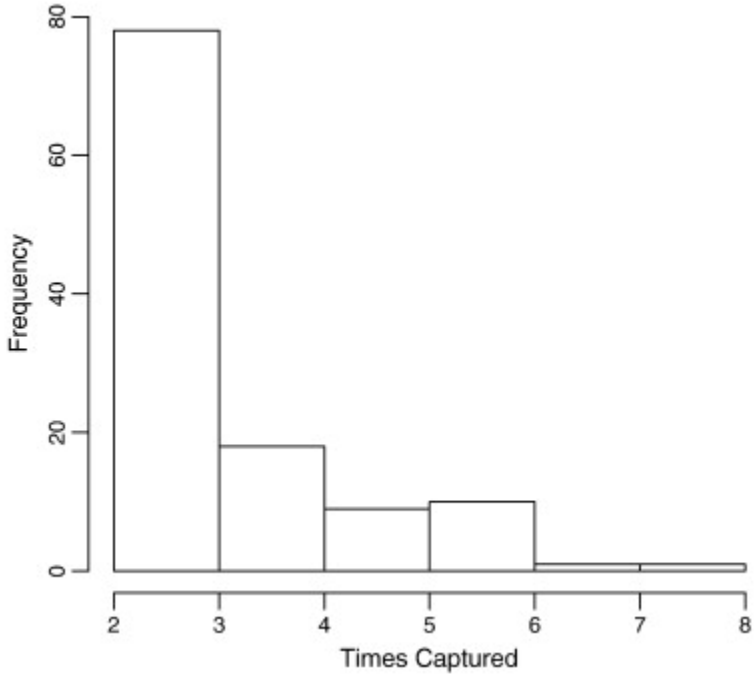

Fig. 1. Histogram of variety of captures of multiply trapped cotton rats.

THE BARTONELLA DATA

The temporal dynamics of Bartonella infections during a population of cotton rats were determined by repeatedly capturing and sampling individual animals near Social Circle, Walton Co., Georgia, USA (Kosoy et al., 2004a). Captured animals were bled, marked with a uniquely numbered ear tag, and after recovery from anesthesia, released at their site of capture. During the longitudinal study conducted over a 17-month interval, 613 samples from 314 individual animals were collected. The procedures used for the isolation of Bartonella are published previously (Kosoy et al., 1997); see Telfer et al. (2007) for an additional isolation approach in a very study of the ecological differences among co-existing Bartonella species in rodents in England. All cotton rats captured four or more times were Bartonella-culture positive a minimum of once. Prevalence of Bartonella infection increased to > 90% among juvenile and sub-adult rats before declining to < 40% among the largest-oldest individuals. Bacteremia levels ranged between

40 and 4.0 × 106 colony-forming units per 1 ml of blood. Bartonella-like colonies, supported their morphology and bacterial microscopy were harvested and ready for PCR analyses using gltA (citrate synthase gene) because the most potent target. Although we used only some of the citrate synthase gene (gltA) for phylogenetic analysis, this gene has been shown to be a reliable tool for distinguishing between closely related Bartonella genotypes (Birtles and Raoult, 1996, La Scola et al., 2003). When bacterial colonies with different morphologies or growth characteristics were obtained from a personal blood sample, additional serial 4-fold dilutions were continued until the numbers of colonies of every phenotype may well be counted. the looks (“smooth” or “rough”), shape, and size of Bartonella-like bacterial colonies were assessed daily. Although the event of observable Bartonella-like colonies may appear in 5–7 days, plates were held and observed for a period of up to twenty days as different phenotypes of Bartonella developed into colonies at different rates (Kosoy et al., 1999). When any apparent variation in colony morphology was distinguished among colonies, a representative single colony was passed onto a brand new agar plate. When bacterial colonies with different morphologies or growth characteristics were obtained from a private blood sample, dilutions were continued until colonies of every phenotype might be enumerated. Our ability to differentiate variants of Bartonella supported colony phenotype (rough or smooth, color, shape, and size) was assessed against the results from genetic analyses. additionally, 128 phenotypically undistinguished Bartonella colonies were sequenced as negative controls, and every one of them were identical by using the cistron.

The gltA alignment was performed for every isolate and DNA sequence similarities were calculated using DNASTAR Lasergene 7 software. Comparison of gltA gene sequence data was accustomed classify all Bartonella variants found within the Bartonella-cotton rat system. The classification of the Bartonella variants into the three major genogroups, namely, A, B, and C, and therefore the further subdivision into A1–A5, B1–B5, and C1–C2 were supported cluster analysis of genetic distances between variants measured by the speed of nucleotide substitutions within the gltA. Each genogroup contained from

2 to 4 unique variants. Results of the similarities were accustomed construct a matrix within which the odds of similarity between all Bartonella variants were included (see Table 1). Although the prevalence of bacteremia thanks to different genogroups/variants of Bartonella was temporally variable, variants of genogroup A predominated during each sampling period (Kosoy et al., 2004a); see Fig. 2 which shows some evidence of synchrony, among the varied variants, that's further quantified by the matrix of contemporaneous correlations in Table 2. Furthermore, Fig. 2 displays the time plot of the monthly counts of co-infections by variants from the identical genogroups, further as that from different genogroups, which shows that these co-infection counts seem to be synchronous; indeed, they're significantly, positively correlated with the whole monthly counts of all infections at 5% significance level, with correlations capable 0.717 and 0.663, respectively, while they seem to be uncorrelated with one another, with correlation 0.282 (p-value 0.78). Note that the co-infection counts are rather small, although the synchrony alluded to earlier generally holds when the counts are further counteracted at individual genogroup level (plots not shown).

Table 1. Matrix indicating genetic similarities among the unique genogroups/variants identified by gltA gene sequence analysis supported 661 isolates of Bartonella from cotton rats.

	A1	A2	A4	A5	B2	B3	B4	B5	C1	C2
A1	1.0000	0.9704	0.9615	0.9970	0.9290	0.9290	0.9349	0.9349	0.8876	0.8905
A2		1.0000	0.9852	0.9704	0.9231	0.9231	0.9290	0.9290	0.8846	0.8876
A4			1.0000	0.9615	0.9231	0.9231	0.9231	0.9231	0.8817	0.8846
A5				1.0000	0.9290	0.9290	0.9349	0.9349	0.8905	0.8935
B2					1.0000	0.9970	0.9911	0.9941	0.9142	0.9172
B3						1.0000	0.9911	0.9941	0.9142	0.9172
B4							1.0000	0.9970	0.9142	0.9172
B5								1.0000	0.9142	0.9172
C1									1.0000	0.9970
C2										1.0000

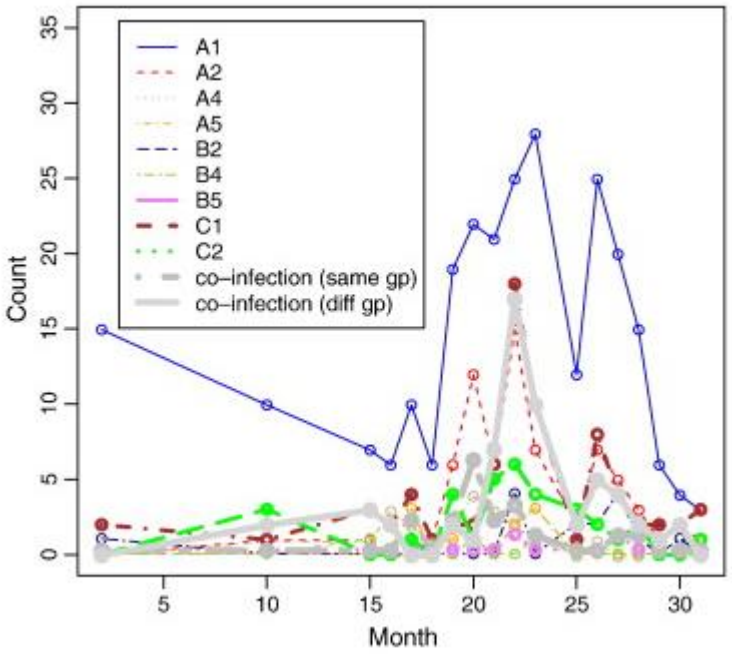

Fig. 2. Time plot of monthly counts of infections by genotypes.

Table 2. Contemporaneous Spearman correlations of counts of infections for several genotypes. All variants of B are combined into one group because of their low counts. Similarly, counts of A1 infections are omitted. Correlations that are significant at 5% level are boldfaced.

	A2	A4	A5	B	C1	C2
A2	0.77					
A4	0.30	0.23				
A5	0.33	0.31	0.37			
B	0.54	0.53	−0.35	−0.04		
C1	0.55	0.46	0.22	0.47	0.34	
C2	0.69	0.68	0.24	0.42	0.20	0.39

The longitudinal study in Georgia has also demonstrated that multiple gltA variants were often (20% of individuals) isolated from one Sigmodon hispidus blood sample; a maximum of 5 variants was recovered from a personal during its sampling history (see Assessing the independent-variant hypothesis within which we feature out a frequency analysis). During this study, a number of these marked rats, 117 in total, were trapped repeatedly and infrequently irregularly, thereby leading to 117 unequally spaced statistic data of mixed infection patterns. The panel of your time series of a succession of mixed infections is analyzed via Markov process analysis in assessing the species-variant hypothesis by Markov process analysis.

Here, we develop some exploratory methods for elucidating the epidemiological character of the genetically classified variants. Specifically, the statistical analysis attempts to resolve the subsequent two questions:

• Are the observed co-infection frequencies in keeping with the hypothesis that the three major genogroups A1–A5, B1–B5, and C1–C2 are epidemiologically distinct species?

• Are the observed temporal patterns within the mixed infections per the hypothesis that the three major genogroups A1–A5, B1–B5, and C1–C2 are epidemiologically distinct

species whereas the within-genogroup variants enjoy some cross-immunity?

Our analyses reported later suggest that the answers to those questions are affirmative, thereby corroborating the usefulness of genetic clustering during this instance as a tool for identifying epidemiologically meaningful grouping of multi-strain pathogens. The new tools developed herein are useful for studying other multi-strain disease-causing agents.

Hypotheses for mixed infections

previous experimental investigations demonstrated that probably all revealed Bartonella variants are typical for the cotton rats since bacteremia might be readily produced only in laboratory-bred cotton rats and not in the other tested rodent species when strains belonging to genogroups A1, B1, and C1 were used for inoculation (Kosoy et al., 2000). this means that an association with a selected rodent species cannot function a criterion for the classification during this case. As was stated earlier, the classification of the Bartonella variants into three genogroups was supported the similarities between genetic variants. Inter-genogroup sequence similarities ranged from 88.2% to 93.5%, whereas sequence similarity among variants within each genogroup ranged from 96.2% to 99.7% (Table 1). However, it's unclear whether or how the genetic differences between these unique genetic variants or genogroups affect the dynamics of infection. One hypothesis is that the three genogroups A, B, and C constitute separate bacterial species each of which admits several variants, which there's no specific immunity between species but some cross-immunity among variants within a species.

This hypothesis is partially supported by previously published field and experimental data suggesting that these genogroups are biologically different:

1) Cross-immunity between strains A1, B1, and C1 might not occur (Kosoy et al., 1999). However, the degree of cross-immunity between closely related variants within the identical genogroup is unknown (e.g., A1 and A2).

2) There exists a large range of variation within the level of bacteremia and growth ki-

netics in cotton rats inoculated with strains A1, B1, and C1. (Kosoy et al., 1999). as an example, inoculation of rats with strain A1 resulted within the highest concentration of microorganisms in their blood, both at low and high doses.

3) The minimal infectious dose for producing bacteremia in cotton rats also significantly varied between the strains with very cheap dose found for C1 strain (Kosoy et al., 2000).

4) Strains of A (A1 and A3) were reported to be found in gnawer embryos and neonates, while strains of genotypes B and C have not been so detected (Kosoy et al., 1998).

5) Strain A, which is that the most typical in cotton rats, was also common in flea vectors (Polygenis gwyni) parasitizing cotton rats. However, another common strain C was absent from P. gwyni, while a rare Sigmodon hispidus strain B was quite common in P. gwyni (Abbot et al., 2007).

Missing information includes the degree of cross-immunity between closely related variants or the similarity of other biological traits within the identical genogroup (e.g., A1 and A2). Furthermore, the mixed infection dynamics were the same as a modified version of the SIR model (Kamo and Sasaki, 2002) taking under consideration the competition of the Bartonella. This hypothesis states that genetically related variants represent one Bartonella species and shall henceforth be cited because the species-variant hypothesis.

An alternative hypothesis states that one vector transmission may carries with it a mix of Bartonella variants (e.g. A1, A2, C1) and every of them may attack a host-only after a random dormant period. This hypothesis was prompted by the observation that some individual mammalian hosts were observed to own mixed infection by different Bartonella variants, over time. for instance, Kabeya et al. (2002) reported evidence of multiple

infections of genetically different Bartonella henselae in naturally infected cats. At each peak of bacteremia, genetically different variants were isolated from the blood of cats showing relapsing bacteremia. The results obtained by Arvand et al. (2006) also suggest that the populations of primary B. henselae isolates are commonly composed of distant genetic variants, which can disappear upon repeated passages among animals. Thus, the generation of genetically independent variants may represent an dodging to bypass the host-specific immune responses. This hypothesis implies that each one variants A1–A5, B1–B5, and C1–C2 are independent variants and shall be said because the independent-variant hypothesis.

Assessing the independent-variant hypothesis

If A1 and A2 infect the rats independently, then the probability of finding A1 and A2 simultaneously in a very rat equals P(A1) times P(A2) where P(A1) is that the probability of finding an A1 strain within the blood of a random rat, and similarly defined is P(A2). We estimate P(A1) by the subsequent weighted sum, namely, the sum of the ratio of hosts infected by A1 alone, plus 1/2 of the ratio of hosts infected by A1 and A2, 1/3 of that of A1A2A5, etc. This deviates from the easy scheme of estimating P(A1) by the ratio of hosts infected by A1 whether or not the infection is single or mixed. the employment of the weighted scheme is justified by the actual protocol used for identifying Bartonella variants infecting a number. Specifically, if a bunch was found to be infected, the blood sample was repeatedly diluted to determine the bacteria load. Bacteria colonies were then cultivated from diluted blood, which usually yielded atiny low number of colonies. THE kind of the Bartonella variant(s) within the blood sample was then determined supported the morphology of the bacteria colonies.

Hence, the variant identification process involved a sampling process. Consequently, the probability P(A1) should be interpreted because the probability of drawing a variant A1 bacterium during a random (diluted) blood sample, within which case it should equal the sum of the probability of one infection by A1 (because the variant A1 is then uniquely identified), plus the probability of a co-infection by A1A2 which the random blood

sample (of the infected rat) contains an A1 bacterium, plus the probability of a co-infection A1A2A5 which the random blood sample contains an A1, etc., with the sum over all distinct infection types involving A1. The probability of finding an A1A2 infected host which the random blood sample of the host contains an A1 Bartonella bacterium equals the probability of finding an A1A2 co-infection times the probability of drawing an A1 bacterium from such a number. The chance of drawing an A1 bacterium within the blood sample of an A1A2 co-infection equals the ratio of the A1 load to a complete load of A1 and A2 within the blood. While experimental results have some information on such a ratio, the massive uncertainty in such information renders it more prudent to adopt the approach of assuming no prior information on the variant loads so we assign the desired contingent probability as 1/2. Altogether, we then estimate the probability of finding an A1A2 infected host which the random blood sample from the host contains an A1 Bartonella bacterium as 1/2 times the ratio of the A1A2 co-infection. Similarly, the probability of finding an A1A2A5 infected host from which an A1 bacterium is drawn will be estimated by 1/3 times the ratio of the A1A2A5 co-infection. This completes our justification of the weighted scheme for estimating P(A1). The estimates of the chances of other variants are listed in Table 3.

Variant	A1	A2	A4	A5	B2	B3	B4	B5	C1	C2
Probability	0.419	0.107	0.033	0.018	0.021	0.009	0.003	0.003	0.099	0.042

Table 3. Probabilities of assorted Bartonella strains within the blood of a bunch

Note that these probabilities sum to the probability that a rat is infected by some Bartonella variant, and hence the sum is a smaller amount than 1. We computed the theoretical probabilities of assorted co-infection patterns under the independent-variant hypothesis and compared them with the observed relative frequencies (Table 4). specifically, we calculated the ratio of the theoretical probability to the observed ratio for various co-infection patterns. Also, we computed the bootstrap 95% in-

tervals for the theoretical probabilities under the independent-variant hypothesis. The bootstrap was done by re-sampling the information cases with replacement, with each data case being one trapping record.

Table 4. Observed relative frequencies of varied co-infections and therefore the corresponding theoretical probabilities calculated under the independence assumption. The last column lists the 95% bootstrap confidence interval of the theoretical probabilities, supported 5000 bootstrap replications. Positively (negatively) dependent mixed infection types are boldfaced (underlined).

Co-infection	Obs. freq.	Theo. prob.	Theo./obs.	95% C.I. theo. prob.
A1A2	0.01863	0.04872	2.615	(0.03838, 0.05978)
A1A2A4	0.00207	0.00158	0.763	(0.00089, 0.00240)
A1A2A5	0.00207	0.00099	0.479	(0.00045, 0.00166)
A1A4	0.00414	0.01351	3.263	(0.00764, 0.01984)
A1A5	0.00207	0.00848	4.096	(0.00396, 0.01395)
A1B2	0.00414	0.00891	2.152	(0.00439, 0.01413)
A1B2C2	0.00207	0.00038	0.183	(0.00016, 0.00067)
A1B3C1	0.00207	0.00025	0.120	(0.00003, 0.00056)
A1B4	0.00207	0.00129	0.625	(0.00000, 0.00300)
A1C1	0.05176	0.04240	0.819	(0.03275, 0.05228)
A1C1C2	0.00207	0.00180	0.869	(0.00111, 0.00256)
A1C2	0.01449	0.01768	1.220	(0.01129, 0.02442)
A2A4	0.00207	0.00379	1.833	(0.00209, 0.00588)
A2A4C2	0.00207	0.00016	0.078	(0.00008, 0.00027)
A2B2	0.00207	0.00250	1.209	(0.00118, 0.00405)
A2C1	0.01242	0.01191	0.959	(0.00841, 0.01582)
A2C2	0.00828	0.00496	0.600	(0.00303, 0.00721)
A4C1	0.00621	0.00330	0.532	(0.00177, 0.00512)
B2C1	0.00207	0.00218	1.052	(0.00105, 0.00361)
B2C2	0.00414	0.00091	0.219	(0.00039, 0.00162)
B3B4	0.00207	0.00002	0.009	(0.00000, 0.00007)
B4C2	0.00207	0.00013	0.064	(0.00000, 0.00032)
B5C1	0.00207	0.00032	0.153	(0.00000, 0.00086)
C1C2	0.00207	0.00432	2.087	(0.00263, 0.00626)

For each co-infection pattern, we will reject the independent-variant hypothesis at 5% significance level if the 95% bootstrap confidence interval doesn't contain the observed ratio. within the case that cross-immunity exists between two variants, the observed frequency of co-infection by the 2 variants is anticipated to be below the theoretical probability. On the opposite hand, if infection by one variant increases the possibility of infection by a second variant, then the observed ratio of the co-infection by the 2 variants is anticipated to be over the theoretical probability. Thus, within the case of rejection of the independence assumption, the position of the frequency as compared to the theoretical probability may shed insight on the connection between the 2 variants under study.

From Table 4, it is inferred that the observed relative frequencies of co-infection by within-genogroup variants A1A2, A1A4, A1A5, A2A4, and C1C2 are smaller than the theoretical probability and lie outside the arrogance interval of the theoretical probability; hence, we will reject the independence assumption for these co-infection patterns. Moreover, because the relative frequencies of those co-infection patterns are smaller than the theoretical counterparts under the independence assumption, there's some evidence that the variants within the A genogroup (and the 2 within the C genogroup) are subject to cross-immunity. Two exceptions to those mentioned are co-infections A1A2A5 and B3B4 that the independence assumption is rejected but the observed ratio is bigger than the theoretical probability. However, these could also be false alarms, especially given the rare occurrences of B3 and B4. Overall, co-infections by variants from the identical genogroup occurred with a frequency of 0.0290 which is significantly below the theoretical probability of 0.0788 (95% confidence interval extends from 0.0658 to 0.0923) under the independence assumption.

Co-infection patterns by some variants of various genogroups, including A1B4, A1C1, A1C1C2, A1C2, A2B2, A2C1, and B2C1, are found to be in line with the independence assumption. On the opposite hand, the observed relative frequencies of co-infections A2C2, A4C1, B2C2, B4C2, B5C1, A1B2C2, A1B3C1, and A2A4C2 are all greater than

their theoretical counterparts under the independence assumption and lie outside the 95% confidence intervals. Overall, co-infections by variants from different genogroups occurred with a ratio of 0.114 which is marginally significantly over the theoretical probability of 0.0965 (95% confidence interval extends from 0.0803 to 0.113) under the independence assumption. Thus, we reject the independence assumption for these co-infection patterns, but now for the possible reason, that infection by a Bartonella variant increases the prospect of being infected by another variant from a special genogroup, perhaps because the system of the host is weakened by an infection of an independent Bartonella species, although this wasn't measured.

In summary, co-infections by variants within the identical genogroup tend to own lower relative frequencies than the theoretical probabilities assuming these within-group variants are independent species. The smaller relative frequencies of the within-group co-infections suggest the presence of cross-immunity between the within-group variants. Alternatively, the lower frequencies are also attributed to strain–strain competition, which, however, is at odds with the synchrony of the varied groups shown in Fig. 2 and Table 2. On the opposite hand, co-infections by between-group variants tend to own relative frequencies like the theoretical probabilities or higher, with A1B2 being a lone exception, thereby suggesting that between-group variants are independent species which infection by one group may slightly increase the possibility of being infected by an independent species because the system of the host is also more prone to the newly encountered pathogen thanks to an on-going infection. The preceding conjecture on the biological mechanism underlying the detected non-independence is partially supported within the next section within which we show strong evidence that a Bartonella infection was found to be less likely followed by another infection by a within-group variant than by a between-group variant.

The aforementioned analysis is predicated on estimating the probability of a selected Bartonella variant by the weighted scheme described at the start of this section. we've also repeated the aforementioned frequency analysis with such probabilities

estimated by a non-weighted scheme, i.e., P(A1) is estimated by the ratio of hosts with one or mixed infection by A1, etc. See Tables S1–S3 within the online Supplementary Material. In general, the analysis supported the non-weighted estimation scheme yields less clear though generally similar conclusions as those inferred from the weighted scheme.

Assessing the species-variant hypothesis by Markov chain analysis

In the previous section, we found strong evidence that the Bartonella variants circulated within the rat system in Georgia are unlikely to be independent variants. Indeed, it seems to support the hypothesis of classifying the A, B, and C genogroups as independent species with the variants within each group enjoying some cross-immunity. But the preceding analysis is predicated on the frequencies of co-infections by various variants. Here, we study the identical problem by temporal analysis of the mixed infections. The key idea is that the species-variant hypothesis implies some correlation structure for infections by the Bartonella variants which will have some observable implications on the successive patterns of mixed Bartonella infections. An example of the monthly mixed infection pattern for a rat trapped multiple times was A1, no Bartonella detected, B2, A2, not trapped, no Bartonella detected; see Fig. 3 for other observed time-series patterns. Cross-immunity between variants from a species may imply that an infection is more likely to be followed by another infection from a special genogroup than from the identical group, after adjusting for his or her epidemiological characteristics (infectivity, transmission rate, and susceptibility). as an example, an A1 infection is also more likely followed by a B1 than an A2, everything else being equal.

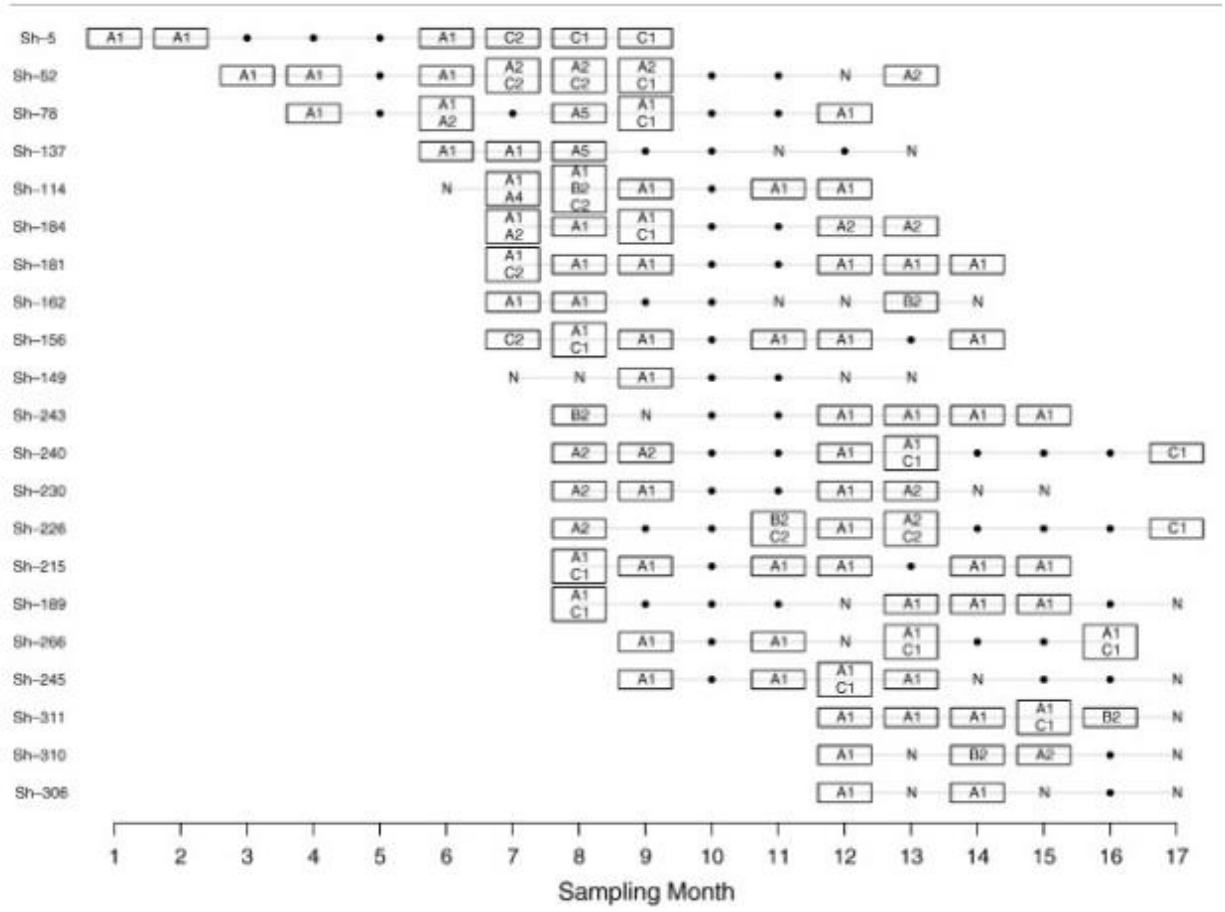

Fig. 3. Re-sampling history, infection course, and genotypic characterization of sequentially recovered Bartonella isolate from 21 cotton rats captured ≥ 5 times during 16 months of trapping (first sampling month is March 1996, but no trapping during December 1996). The genogroups/variants of Bartonella recovered from bacteremic rats at a sample month are shown; the primary symbol indicates the date a rat first entered into the study cohort. “N” stands for no detectable bacteremia at a sample month. Tick marks indicate a rat wasn't recaptured during that trapping session.

The dynamics of the mixed infection pattern is studied via a Markov process analysis of the monthly disease status of a random host. Analysis was done employing a subset of the gnawing animal data where a rat was multiply trapped, leading to 117 time-series data on the succession pattern of mixed infections. Only the subsequent variants were

observed: A1, A2, A4, A5, B2, B3, B5, C1, and C2. However, there are two complications to the current approach. First, a rat may have co-infections, e.g. A1A2, in an exceedingly certain month. This necessitates enlarging the state space of the Markov chain to incorporate all observable co-infection patterns; altogether there are 34 states for the Markoff process. Second, the trapping dataset naturally has many missing data, because the same rat would seldom be trapped each month. Fortunately, maximum likelihood estimation of the transition probability matrix of the Markoff process with extensive missing data are often distributed by the EM algorithm (Dempster et al., 1977). See Rabiner (1989) for details. For data not stricken by missing data, there are other approaches for analyzing mixed infection data, see, e.g. Weinberger et al (2008).

As mentioned earlier, thanks to the presence of co-infections, the states of the Markoff process also include A1A2, etc. altogether 34 (observed) states. it's clearly not revealing to report the estimated 34 by 34 transition probability matrix because it is difficult to investigate such a large matrix. But the most issue concerns the succession frequencies of infections by the identical genogroup or by a distinct genogroup. From this angle, the estimated transition probability matrix will be wont to provide such information as listed in Table 5. Each row in Table 5 gives the conditional probabilities of every of six events given a bunch is infected within the current month by the variant stated within the row heading. The six events concern the disease status of the host within the next month: (i) the host maintains the identical disease status within the next month, (ii) it acquires a co-infection by another variant from the identical genogroup, (iii) it acquires a co-infection by another variant from a unique genogroup, (iv) the first variant is replaced by another variant from the identical genogroup, (v) the initial variant is replaced by another variant from a unique genogroup and (vi) no Bartonella is detected. These six events are labeled as “maintain”, “acq. same”, “acq. diff”, “repl. same”, “repl. diff” and “undetected”, respectively. as an example, from the row with the heading “A1”, we will read that given a rat has an infection by A1 within the current month, the chance that it maintains the identical infection by A1 within the next month equals 0.6254, and therefore the contingent probability that it acquires a co-infection by a variant from genogroup A is

0.0447, etc.

Table 5. Conditional probabilities of varied successive infection patterns within the next month. Each row gives the conditional probabilities of every of six events given a bunch that's infected within the current month by the variant stated within the row heading. The six events concern the disease status of the host within the next month: (i) the host maintains the identical disease status within the next month, (ii) it acquires another variant from the identical genogroup, (iii) it acquires another variant from a special genogroup, (iv) the first variant is replaced by another variant from the identical genogroup, (v) the initial variant is replaced by another variant from a special genogroup and (vi) no Bartonella is detected.

Infection\disease status	Maintain	Acq. same	Acq. diff	Repl. same	Repl. diff	Undetected
A1	0.6254	0.0447	0.0983	0.0800	0.0418	0.1098
A2	0.4042	0.0435	0.0000	0.2082	0.2166	0.1275
A4	0.0000	0.0000	0.1597	0.5679	0.0000	0.2723
A5	0.2097	0.0000	0.0000	0.0000	0.1875	0.6028
B2	0.0000	0.0000	0.0000	0.0000	0.3266	0.6734
B3	0.2085	0.0000	0.0000	0.0001	0.0000	0.7914
B5	0.0000	0.0000	0.0000	0.0000	1.0000	0.0000
C1	0.4246	0.0000	0.0582	0.0012	0.3476	0.1685
C2	0.0000	0.0000	0.1092	0.2434	0.6474	0.0000
Bartonella infection	0.4643	0.0306	0.0762	0.1138	0.1585	0.1567

The information are often further summarized within the last row of the table, which comprises the conditional probabilities of every of the six events providing a number includes a Bartonella infection within the current month. These probabilities are

normalized weighted column sums with the load of every row up to the probability of infection by the variant, labeling that row, at this month, then the weighted column sums are renormalized to form them sum to 1. The numbers within the last row of Table 5 have the subsequent interpretations: on condition that a rat is infected by some Bartonella variant within the current month, the probability that the rat continues to be infected by the identical variant, i.e. unchanged Bartonella infection type, equals 0.4643; that the rat maintains the identical variant and at the identical time acquires another variant of the identical genogroup within the next month has probability 0.0306; that the rat maintains the identical variant and acquires another variant of a spcial genogroup within the next month has chance 0.0762. Thus, the difference of those two probabilities equals 0.0762 − 0.0306 = 0.0456. we've also computed a 95% bootstrap confidence interval for the difference. within the bootstrap, the full statistic of Bartonella infection type (or lack of infection) for every rat forms a unit, and that we bootstrap these statistic units by randomly sampling the panel of your time series of infection type with replacement. for every bootstrap panel of your time series, we estimate the conditional probabilities adore the last row of Table 5. supported 500 bootstrap replications and Efron's percentile method (Efron and Tibshirani, 1993), the bootstrap 95% confidence interval of the difference is (0.0197, 0.944), suggesting that the chance of acquiring another variant from a distinct genogroup is significantly beyond that from the identical genogroup, at 5% significance level.

The probability that given a Bartonella infection within the current month, the rat isn't any longer infected by that variant but acquires another variant of the identical genogroup within the next month is 0.1138, which the rat isn't any longer infected by that variant but acquires another variant of a unique genogroup within the next month has probability 0.1585. Note that the difference of the 2 probabilities equals 0.1585 − 0.1138 = 0.0447, with the corresponding 95% bootstrap confidence interval being (0.00979, 0.122), again suggesting that the chance of being replaced by a variant of a distinct genogroup is significantly above that of the identical genogroup. Finally, the contingent probability that given a Bartonella infection within the current month, the

probability that the rat has no detectable Bartonella variant within the next month is 0.1567.

Altogether, these results strongly suggest that a Bartonella infection is a smaller amount likely to be followed by an infection by another variant of the identical genogroup than by one in all a unique genogroup. This finding is in line with the species-variant hypothesis that the A, B, and C genogroups are independent subspecies and variants of every genogroup enjoy cross-immunity to some extent, whereas variants of various groups are somewhat independent, although there's a small increase within the cross-group infection rate.

Conclusion

Based on a frequency analysis of co-infections by various Bartonella variants, there's some strong evidence against the hypothesis that every one Bartonella variants are independent species. On the opposite hand, the results of the analysis are in step with the hypothesis that the A, B, and C genogroups function as independent species but the variants within each genogroup enjoy some cross-immunity against one another. there's also some evidence that while the three genogroups are largely independent species, infection by one genogroup may weaken or skew the host immunity which promotes infection by another genogroup.

A second analysis of the panel of your time series of Bartonella infection history for cotton rats that were trapped repeatedly yields results in step with the co-infection frequency analysis. Specifically, an infection is more likely to be followed by another infection by another variant from a special genogroup than from the identical genogroup. These analyses favor the species-variant hypothesis that the three genogroups A, B, and C circulating among the rat system in Georgia (U.S.A.) are more or less independent species, which explains the high prevalence rate of Bartonella infection observed within the studied rat system.

Strain structure is fundamentally important not just for understanding the dynamics of Bartonella but also for several other pathogens. The analysis provided within the current study applies to pathogen-host systems, where individual genetic variants form stable clusters that behave in some ways as independent species. This stability implies that in some cases analysis of multiple separate variants will be replaced by analysis of some clusters (species) that may promote developing predictive models.

Some interesting future work consists of fitting a modified SIR model that accounts for the cross-immunity between different variants, and further assessing the varied hypotheses within such a framework. Another interesting direction of research is to correlate the estimated cross-immunity pattern with the known genetic distance between the variants. Finally, it's of interest to assess the soundness (chaoticity) of the estimated modified SIR model and also the feasibility of the long-run co-existence of various variants.

REFERENCES

Abbot et al., 2007
P. Abbot, A.E. Aviles, L. Eller, L.A. Durden
Mixed infections, cryptic diversity, and vector-borne pathogens: evidence from Polygenis fleas and Bartonella species
Appl. Environ. Microbiol., 73 (2007), pp. 6045-6052
View Record in ScopusGoogle Scholar

Abu-Raddad and Ferguson, 2004
L.J. Abu-Raddad, N.M. Ferguson
The impact of cross-immunity, mutation, and stochastic extinction on pathogen diversity
Proc. R. Soc. Lond. B, 271 (2004), pp. 2431-2438

CrossRef View Record in ScopusGoogle Scholar

Ackerman et al., 1990
E. Ackerman, J. Longini, S. Seaholm, A. Hedin
Simulation model for viral interference in influenza
Int. J. Epidemiol., 19 (1990), pp. 444-454
CrossRef View Record in ScopusGoogle Scholar

Anderson and Neuman, 1997
B.E. Anderson, M.A. Neuman
Bartonella spp. as emerging human pathogens
Clin. Microbiol. Rev., 10 (1997), p. 203219
Google Scholar

Andreasen et al., 1997
V. Andreasen, J. Lin, S. Levin
The dynamics of cocirculating influenza strains to confer partial cross-immunity
J. Math. Biol., 35 (1997), pp. 825-842
View Record in ScopusGoogle Scholar

Arvand et al., 2006
M. Arvand, H. Schubert, J. Viezens
The emergence of distinct genetic variants within the population of primary Bartonella henselae isolates
Microbes Infect., 8 (2006), pp. 1315-1320
Article Record in ScopusGoogle Scholar

Birtles and Raoult, 1996
R.J. Birtles, D. Raoult
Comparison of partial citrate synthase gene (gltA) sequences for phylogenetic analysis

of Bartonella species
Int. J. Syst. Bacteriol., 46 (1996), pp. 891-897
CrossRef View Record in ScopusGoogle Scholar

Breitschwerdt and Kordick, 2000
E.B. Breitschwerdt, D.L. Kordick
Bartonella infection in animals: carriership, reservoir potential, pathogenicity, and zo-onotic potential for human infection
Clin. Microbiol. Rev., 13 (2000), pp. 428-438
View Record in Scopus Google Scholar

Calvez et al., 2005
V. Calvez, A. Korobeinikov, P. Maini
Cluster formation for multi-strain infections with cross-immunity
J. Theor. Biol., 233 (2005), pp. 75-83
ArticleDownload PDFView Record in ScopusGoogle Scholar

Dempster et al., 1977
A.P. Dempster, N.M. Laird, D.B. Rubin
Maximum likelihood from incomplete data via the EM algorithm
(with Discussion)
J. R. Stat. Soc. B, 39 (1977), pp. 1-38
View Record in ScopusGoogle Scholar

Eames and Keeling, 2006
K.T.D. Eames, M.J. Keeling
Coexistence and specialization of pathogen strains on contact networks
Am. Nat., 168 (2006), pp. 230-241
CrossRefView Record in ScopusGoogle Scholar

Efron and Tibshirani, 1993
B. Efron, R.J. Tibshirani
An Introduction to the Bootstrap
CRC Press, Boca Raton (1993)
Google Scholar

Gog and Grefell, 2002
J.R. Gog, B.T. Grenfell
Dynamics and selection of many-strain pathogens
Proc. Natl. Acad. Sci. The USA, 99 (2002), pp. 17209-17214
View Record in ScopusGoogle Scholar

Gupta et al., 1998
S. Gupta, N. Ferguson, R. Anderson
Chaos, persistence, and evolution of strain structure in antigenically diverse infectious agents
Science, 240 (1998), pp. 912-915
View Record in ScopusGoogle Scholar

Holmes et al., 1995
E.C. Holmes, S. Nee, A. Rambaut, G.P. Garnett, P.H. Harvey
Revealing the history of communicable disease epidemics through phylogenetic trees
Phil. Trans. R. Soc. Lond., B349 (1995), pp. 33-40
View Record in ScopusGoogle Scholar

Jardine et al., 2005
C. Jardine, G. Appleyard, M. Kosoy, D. McColl
Rodent-associated Bartonella in Saskatchewan, Canada
Vector Borne Zoonotic Dis., 5 (2005), pp. 402-409

CrossRefView Record in ScopusGoogle Scholar

Kabeya et al., 2002
H. Kabeya, S. Maruyama, M. Irei, R. Takahashi, M. Yamashita, T. Mikami
Genomic variations among Bartonella henselae isolates derived from naturally infected cats
Vet. Microbiol., 89 (2002), pp. 211-221
ArticleDownload PDFView Record in ScopusGoogle Scholar

Kamo and Sasaki, 2002
M. Kamo, A. Sasaki
The effect of cross-immunity and seasonal forcing in a very multi-strain epidemic model
Physica D, 165 (2002), pp. 228-241
Article Record in ScopusGoogle Scholar

Kosoy et al., 1997
M. Kosoy, R. Regnery, T. Tzianabos, E. Marston, D. Jones, D. Green, G. Maupin, J. Olson, J.E. Childs
Distribution, diversity, and host specificity of Bartonella in rodents from the southeastern u. s.
Am. J. Trop. Med. Hyg., 57 (1997), pp. 578-588
CrossRef View Record in ScopusGoogle Scholar

Kosoy et al., 1998
M. Kosoy, R. Regnery, O. Kosaya, D. Jones, E.L. Marston, J.E. Childs
Isolation and characterization of Bartonella from embryos and neonates of naturally infected rodents
J. Wildlife Dis., 34 (1998), pp. 305-309
CrossRef View Record in ScopusGoogle Scholar

Kosoy et al., 1999
M.Y. Kosoy, R.L. Regnery, O.I. Kosaya, J.E. Childs
Experimental infection of cotton rats with three present Bartonella species
J. Wildlife Dis., 35 (1999), pp. 275-284
CrossRefView Record in ScopusGoogle Scholar

Kosoy et al., 2000
M.Y. Kosoy, E.K. Saito, D. Green, E.L. Marston, D.C. Jones, J.E. Childs
Experimental evidence of host specificity of Bartonella infection in rodents
Comp. Immunol. Microbiol. Infect. Dis., 23 (2000), pp. 221-238
ArticleDownload PDFView Record in ScopusGoogle Scholar

Kosoy et al., 2004a
M.Y. Kosoy, E.L. Mandel, D.C. Green, E.L. Marston, D.C. Jones, J.E. Childs
Prospective studies of Bartonella of rodents. Part I. Demographic and temporal patterns in population dynamics
Vector Borne Zoonotic Dis., 4 (2004), pp. 285-295
CrossRefView Record in ScopusGoogle Scholar

Kosoy et al., 2004b
M.Y. Kosoy, E.L. Mandel, D.C. Green, E.L. Marston, D.C. Jones, J.E. Childs
Prospective studies of Bartonella of rodents. Part II. Diverse infections in an exceedingly single rodent community
Vector Borne Zoonotic Dis., 4 (2004), pp. 296-305
CrossRefView Record in ScopusGoogle Scholar

Kurtenbach et al., 2006
K. Kurtenbach, K. Hanincova, J. Tsao, G. Margos, D. Fish, N. Ogden
Fundamental processes within the evolutionary ecology of Lyme Borreliosis
Nat. Rev., 4 (2006), pp. 660-669

Sept
CrossRefView Record in ScopusGoogle Scholar

La Scola et al., 2003
B. La Scola, Z. Zeaiter, A. Khamis, D. Raoult
Gene-sequence-based criteria for species definition in bacteriology: the Bartonella paradigm
Trends Microbiol., 11 (2003), pp. 318-321
View Record in ScopusGoogle Scholar

Levin et al., 1999
B.R. Levin, M. Lipsitch, S. Bonhoeffer
Population biology, evolution, and infectious disease: convergence and synthesis
Science, 283 (1999), pp. 806-809

View Record in ScopusGoogle Scholar
Nelson and Holmes, 2007
M.I. Nelson, E.C. Holmes
The evolution of epidemic influenza
Nat. Rev. Gen., 8 (2007), pp. 196-205
CrossRefView Record in ScopusGoogle Scholar

Rabiner, 1989
L.R. Rabiner
A tutorial on hidden Markov models and selected applications in speech recognition
Proc. IEEE, 77 (1989), pp. 257-286
View Record in ScopusGoogle Scholar

Read and Taylor, 2001
A.F. Read, L.H. Taylor

The ecology of genetically diverse infections
Science, 292 (2001), pp. 1099-1102
View Record in ScopusGoogle Scholar

Rohani et al., 2003
P. Rohani, C. Green, N. Mantilla-Beniers, B. Grenfell
Ecological interference between fatal diseases
Nature, 422 (2003), pp. 885-888
View Record in ScopusGoogle Scholar

Telfer et al., 2007
S. Telfer, H.E. Clough, L.R. Birtles, M. Bennett, D. Carslake, S. Helyar, M. Begon
Ecological differences and coexistence during a guild of microparasites: Bartonella in wild rodents
Ecology, 88 (2007), pp. 1841-1849
CrossRefView Record in ScopusGoogle Scholar

Tibayrenc and Ayala, 2000
M. Tibayrenc, F. Ayala
Molecular epidemiology and evolutionary genetics of pathogenic microorganisms: analysis and interpretation of information
R.C.A. Thompson (Ed.), Molecular Epidemiology of Infectious Diseases, Arnold, ny (2000), pp. 20-29
Google Scholar

Weinberger et al., 2008
D.M. Weinberger, R. Dagan, N. Givon-Lavi, G. Regev-Yochay, R. Malley, M. Lipsitch
Epidemiologic evidence for serotype-specific resistance to pneumococcal carriage
JID, 197 (2008), pp. 1511-1519
CrossRefView Record in Scopus

AUTHOR AFFILIATION

P. HOSOLY,

D. TEGNE,

V. KZIANNABOO,

E. PERSTON,

O. JAMES