Host–parasite network structure is associated with community-level immunogenetic diversity

Shai Pilosof¹, Miguel A. Fortuna², Jean-François Cosson³, Maxime Galan³, Chaisiri Kittipong⁴, Alexis Ribas⁵, Eran Segal⁶, Boris R. Krasnov¹, Serge Morand⁷,⁸,⁹ & Jordi Bascompte²

Genes of the major histocompatibility complex (MHC) encode proteins that recognize foreign antigens and are thus crucial for immune response. In a population of a single host species, parasite-mediated selection drives MHC allelic diversity. However, in a community-wide context, species interactions may modulate selection regimes because the prevalence of a given parasite in a given host may depend on its prevalence in other hosts. By combining network analysis with immunogenetics, we show that host species infected by similar parasites harbour similar alleles with similar frequencies. We further show, using a Bayesian approach, that the probability of mutual occurrence of a functional allele and a parasite in a given host individual is nonrandom and depends on other host-parasite interactions, driving co-evolution within subgroups of parasite species and functional alleles. Therefore, indirect effects among hosts and parasites can shape host MHC diversity, scaling it from the population to the community level.

¹ Mitrani Department of Desert Ecology and Albert Katz International School for Desert Studies, Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion 8499000, Israel. ² Integrative Ecology Group, Estación Biológica de Doñana (EBD-CSIC), Seville 41092, Spain. ³ INRA, UMR CBGP (INRA/IRD/Cirad/Montpellier SupAgro), Campus International de Baillarguet, CS 30016, 34988 Montferrier-sur-Lez cedex, France. ⁴ Department of Helminthology, Faculty of Tropical Medicine, Mahidol University, Bangkok 10400, Thailand. ⁵ Biodiversity Research Group, Faculty of Science, Udon Thani Rajabhat University, Udon Thani 41000, Thailand. ⁶ Department of Computer Science and Applied Mathematics, Weizmann Institute, Rehovot, 76100, Israel. ⁷ Centre National de la Recherche Scientifique—Institut des Sciences de l’Évolution, Université Montpellier 2, Montpellier 34095, France. ⁸ Centre de coopération International en Recherche Agronomique pour le Développement, Animal et Gestion Intégrée des Risques, Campus de Baillarguet, F-34093 Montpellier Cedex 5, France. ⁹ Centre d’Infectiologie Christophe Mérieux du Laos, PO Box 3888, Samsenthai Road, Vientiane, Lao PDR. Correspondence and requests for materials should be addressed to S.P. (email: shainova@gmail.com).
Speciation interactions are a source of selective pressures that can maintain genetic variability in host populations, parasitism being a classic example. In particular, this is true for the major histocompatibility complex (MHC) in vertebrates. The MHC is a highly polymorphic region in the genome comprising closely linked genes. These genes encode cell-surface glycoproteins that recognize and present foreign antigens, such as those of invading parasites, to T lymphocytes thereby initiating an immune response. Thus, at the host individual level, certain MHC alleles confer resistance/susceptibility to some parasites but not to others.

Parasite-mediated selection drives MHC polymorphism at the host population level because of the evolutionary interplay between parasites and MHC-encoded proteins within a host individual. Indeed, several studies have shown that in a host population, parasite diversity—the number and identity of parasite species and their abundance in a host population—is correlated with MHC diversity—the number of MHC alleles in a host and their identity—but not with genetic diversity assessed with neutral markers such as microsatellites or mitochondrial genes.

At the species level, studies of host interspecific variation in MHC are scarce but have confirmed a positive correlation between parasite species richness and MHC richness across species. However, these studies were based on data collected from the literature, dealt with host species that were not simultaneously surveyed both for parasites and MHC alleles, and did not consider the functional diversity of alleles. More importantly, the way interactions among hosts and parasites are distributed was ignored and thus no study examined the association between multiple host and parasite interactions and MHC diversity. Although few studies have explored the effects of multiple parasites on MHC diversity, they were conducted within a population of a single host.

Interactions between multiple hosts and parasites are often depicted as a host–parasite network. Considering network structure is crucial because, in nature, species do not live in isolation but rather belong to a community where species affect each other directly or indirectly. In this community-wide context, the prevalence of a given parasite species in a given host species may depend on its prevalence in other host species in the community. Therefore, the selective pressure exerted by the parasite may be influenced by indirect species interactions and may itself influence diversity of MHC alleles directly involved in the host-parasite interaction (Fig. 1). That is, the structure of the host–parasite network may correlate with MHC diversity among host species. Using a unique data set on a community of rodents, their helminths parasites and their MHC alleles, we tackle here, for the first time, this largely unexplored research topic by integrating complex network theory with immunogenetics.

We show that host species infected by similar parasites harbour similar MHC functional alleles with similar frequencies after controlling for phylogenetic effects. Also, the indirect effects among hosts and parasites influence the probability of mutual occurrence of a functional MHC allele and a parasite in a given host individual. Therefore, the ecology of host–parasite interactions and the evolution of MHC are intertwined, and indirect effects among hosts and parasites can shape host MHC diversity, scaling it from the population to the community level.

Results
Data set. An extensive helminth survey was conducted in six localities in southeast Asia (Supplementary Fig. 1) on 950 specimens of 11 murinae rodent species within the framework of the Community of Rodents and their Pathogens project (www.ceropath.org). Of these, 440 individuals were parasitized by 26 helminth taxa (22 identified to species level and 4 to family level) (Supplementary Tables 1–3).

Infection usually depends on species co-occurrence in space and time. Also, hosts of the same species but from different populations may be exposed to different selective pressures, thereby creating variation in MHC due to local adaptation. We therefore first analysed data using rodent individuals captured in a single locality (Loei). Despite geographical distance between the localities, they all belong to the same Indochina bioregion and are rather homogeneous in their environmental conditions. Furthermore, all rodent species from other localities occurred in Loei (Supplementary Table 1). Therefore, we pooled data from the six localities and considered rodents as belonging to a single community to increase statistical power. This did not change our results qualitatively (see below) and we thus focus on the results of analyses done on pooled data.

Each individual rodent was genotyped using a portion of the MHC class II DRB exon-2 functional gene that recognizes extracellular parasites such as helminths. Foreign antigens are recognized by particular amino acids of the MHC molecule called

Figure 1 | An example for the interplay between host–parasite and host-supertype networks. (a) On the left, host–parasite interactions between helminths and rodent species (ecological network). Arrow width is proportional to the prevalence of a given parasite on a host. On the right, the immunogenetic host-supertype network in which interactions represent the frequencies (indicated by line width) of supertypes harboured by individuals of a given host that are infected by each parasite (indicated by different colours). (b) A theoretical example showing how in a community-wide context, species and MHC supertypes affect each other indirectly with possible cascading indirect effects. Imagine that a mutation for resistance (depicted by a red star) occurs in an allele from supertype no. 3 in the population of host no. 1 (step 1), causing the prevalence of parasite no. 1 to decrease in that host (step 2). Parasite no. 1 now switches to host no. 2 (step 3), thus reducing the prevalence of parasite no. 2 in host no. 2 due to competition (step 4).
the antigen-binding site. Some MHC alleles code for MHC molecules that bind an overlapping gamma of peptides. This happens when the antigen-binding sites of the coded molecules possess amino-acid sequences with similar physical and biochemical properties. Such alleles are thus functionally similar and can be classified into supertypes based on their antigen-binding site amino-acid sequences. Because we were interested in functional diversity, and to avoid sparse networks caused by an excess of alleles with very few interactions, we first clustered functionally similar alleles into supertypes following refs 22,23 (see Methods section and Supplementary Fig. 2). This resulted in 24 supertypes but we omitted from statistical analysis one supertype (no. 23) that was not harboured by any infected individual, and thus 23 supertypes remained for analysis.

Host–parasite and host–supertype networks are correlated. We constructed a host–parasite ‘ecological network’ and a host–supertype ‘immunogenetic network’ in which the value of interactions were set as parasite prevalence or supertype frequency, respectively (matrix cell values in Fig. 2). If parasite-mediated selection is associated with the structure of the host–parasite network, then there should be a correlation between the identity and prevalence of parasite species and the identity and frequency of supertypes that hosts are infected by or harbour, respectively. Yet, since host species share a common evolutionary history, we have to account for the role that host phylogeny plays in the distribution of supertypes among host species (Fig. 2). We applied a partial Mantel’s test to quantitative dissimilarity matrices based on the ecological and immunogenetic networks, while using a matrix of pairwise phylogenetic distances to control for phylogenetic effects, and found a positive and significant correlation between the host–parasite and host–supertype matrices ($r = 0.62$, $P = 0.001$; Supplementary Fig. 3). This means that two host species, infected by similar parasite species with a similar prevalence, tend to harbour the same MHC supertypes with similar frequencies. These results were robust to host abundances, to the number of parasites that infect a host, to taxonomic resolution of parasite species, to different phylogenetic trees and to statistical ‘group effects’ (Supplementary Methods).

Parasite–supertype associations across hosts are nonrandom. We further explored this first evidence using a Bayesian approach that quantified parasite–supertype associations across host species (Fig. 3). This analysis estimates the probability of an individual host infected by a given parasite species to harbour a given supertype. This probability can be interpreted as a proxy for the resistance/susceptibility that a given supertype provides against a given parasite as compared with all other supertypes across hosts. A high probability means that in hosts where the parasite occurs, the supertype also occurs with high probability, and thus individuals harbouring that supertype are susceptible to the parasite. The advantage of this approach is threefold. First, we can gain a new insight into our system: the probability of obtaining a given supertype $s_k$ given a parasite $p_i$, calculated across hosts (see Fig. 3 and Methods section). Second, we can gain a new insight into our system: the probability of obtaining a given supertype $s_k$ given a parasite $p_i$, calculated across hosts (see Fig. 3 and Methods section). Third, the calculation of the probability of a given parasite–supertype association is based on all host–parasite interactions, thus considering the structure of the host–parasite network.

The resulting parasite–supertype network (Fig. 3) contained all 23 supertypes and 26 parasites, with connectance (proportion of realized interactions out of all possible ones) of 72%, indicating that 28% of the potential supertype and helminths associations...
were not realized in any host. The values of the probabilistic associations were not affected by host abundance (Supplementary Methods). The first important result evident from this analysis is that the association between supertypic and parasites is variable across the host community: Whereas the association of some parasites to supertypic is homogenous for all supertypic, other parasites are strongly associated with specific supertypic. For instance, hosts are susceptible to *Syphacia muris* to a similar extent, regardless of the supertypic they harbour but hosts that harbour supertype no. 10 are more susceptible to *Moniliformis moniliformis* than others. Similarly, the resistance some supertypic (for example, supertypic no. 7 and no. 8) provide is homogenous to all parasites, whereas others (for example, supertypic no. 10 and no. 19) do not confer resistance to specific parasites.

This analysis constitutes the first across-species analysis of the associations between multiple hosts, parasites and supertypic, and indicates that the diversity of MHC supertypic across hosts is associated with the structure of the host–parasite network. To further examine whether the observed parasite–supertype probabilistic associations are correlated with the structure of the host–parasite network, we compared the parasite–supertype matrix (Fig. 3) to 1,000 matrices recalculated after reshuffling the host–parasite interactions (\(M_{\text{rand}}\)). If the observed associations between parasites and supertypic (\(M_{\text{obs}}\)) are not correlated with the host–parasite network structure, then they are determined by a random process and should resemble the values obtained in the random matrices. We compared cell values of \(M_{\text{obs}}\) with distributions of their respective cell values from \(M_{\text{rand}}\) obtained by two null models that reshuffle only the values or the values and the structure of the host–parasite matrix. We found that 47% and 55% of the associations, respectively, significantly differed from random (\(P < 0.05\); two-tailed permutation tests), accounting for 66% of the parasite–supertype associations for the two null models combined (Fig. 3 and Supplementary Figs 4–6). These results were robust to host abundance and parasite taxonomic resolution (Supplementary Methods). This provided a second strong evidence for the interplay between host–parasite network structure and the diversity of MHC in different hosts.

**Networks are nested.** If indeed there is an interplay between the ecological and immunogenetic networks, then we should be able to predict a structural property of the immunogenetic host–supertype network based on the ecological host–parasite network. One common structural property of ecological networks is nestedness. In a nested host–parasite network, hosts infected by few parasite species (further referred to as ‘specialist’ hosts) tend to be parasitized by complete subsets of the parasites that infect more ‘generalist’ hosts—a common pattern in ecological networks. The structure of the host–parasite network studied here was significantly nested (\(NODF = 51.8, P < 0.001\); ref. 27). Thus, a substantial proportion of interactions exists among a ‘core’ of generalist hosts and parasites, which are probably those most associated with the diversity of MHC. This finding leads us to predict that hosts with low MHC richness (that harbour few supertypic) will harbour subsets of the supertypic harboured by MHC-rich hosts, resulting in a nested host–supertype network. This prediction was confirmed (\(NODF = 50.3, P < 0.001\)).
The nested structure is of particular interest because parasite species differ in their host specificity. Thus, parasites specific to one or few hosts may encounter only a limited spectrum of supertypes. Hence, we can predict a nested structure in the parasite–supertype network because, owing to the nested structure of the host–parasite network, parasites infecting subsets of the hosts that more generalist parasites infect should also encounter subsets of supertypes harboured by those hosts. This prediction was also confirmed (NODF = 78.7, P < 0.001). However, nestedness is a qualitative measure that ignores parasite identity. We can further predict that if specificity is driving the observed patterns, then generalist parasites (infecting multiple hosts) will also be associated with many supertypes. If no such correlation is found, then parasite specificity is not a major factor to consider in our system. An additional analysis of parasite specificity that considered the quantitative nature of the networks and parasite identity showed that parasites specific to particular hosts are not necessarily specific to particular supertypes (Pearson’s correlation; r = 0.2, P = 0.34; Supplementary Fig. 7). This happens because some parasites with high host specificity infect hosts with high MHC diversity (for example, Gnathostoma sp. is specific to Rattus tanezumi that harbours many supertypes; Fig. 2). Taken together, these two analyses of specificity suggest that the interplay between the ecological and immunogenetic networks is not driven by patterns of parasite specificity, but that there is a predictable nested pattern in the selection regime between parasites and supertypes, mediated through hosts.

**Parasite–supertype co-evolution.** From an evolutionary perspective, the molecular affinity between the proteins coded by given supertypes and parasites should lead to tight associations between subgroups of parasites and supertypes in our network. We examined this hypothesis using modularity analysis and found that the parasite–supertype network had four modules (Fig. 3). Supertypes and parasites within the same module interact with each other to a greater extent than with supertypes or parasites from other modules, suggesting a higher potential of parasite–supertype coevolution. The modularity of the observed parasite–supertype association network (Q = 0.37) was significantly higher than that of random networks assembled as above using both null models (P < 0.001 for both null models), indicating that the modular structure represented true co-evolutionary interactions rather than random associations between hosts and parasites.

**Analysis of a single community.** In Loei, there were 342 rodent individuals (abundance ranged between 8 and 84; Supplementary Table 1) belonging to 11 species. These species were parasitized by 18 of the 26 helminths taxa (70%) in the pooled data set.

There was a positive and significant correlation between the two host–parasite and host–supertype matrices, after controlling for phylogenetic effects (r = 0.7, P = 0.001). This correlation was maintained when controlling for host abundances and statistical ‘group effects’ (Supplementary Methods).

The parasite–supertype matrix contained all 18 helminths taxa and 17 supertypes. The connectance of this network was 65%. When compared with the two null models that reshuffle only the values or the values and structure of the host–parasite matrix, we found that 39% and 58% of the associations, respectively, significantly differed from random (P < 0.05; two-tailed permutation tests), accounting for 67% of the parasite–supertype associations for the two null models combined.

The modularity of the observed parasite–supertype association network (Q = 0.37) was significantly higher than that of random networks (N = 1,000) assembled as beforehand using the two null models that shuffle interaction values or values and network structure (P < 0.001 for both). The network had three modules.

The host–parasite, host–supertype and parasite–supertype networks were all significantly nested (NODF = 48.4, 44.4 and 80.8, P < 0.01 for all three networks). As with the pooled data, parasites specific to particular hosts were not specific to particular supertypes (Pearson’s correlation; r = 0.37, P = 0.13).

**Discussion**

This study is the first to consider the structure of the associations between multiple parasites and hosts in the context of functional MHC diversity, leading to novel insights regarding the ecology and evolution of MHC. Specifically, our results lead to the conclusion that considering the population level alone is not sufficient to understand parasite-mediated selection. Even the across-host positive correlation between MHC allele richness and parasite richness demonstrated in earlier studies 11,12 gives only a narrow picture of the complex evolutionary process. Instead, the ecological pattern in which host–parasite interactions are organized is crucial to understand adaptive evolution of MHC.

It is evident from our data that the direct and indirect effects hosts and parasites exert on each other are correlated with more than half of the variability in MHC diversity resulting from parasite-mediated selection. Yet, some parasite–supertype associations were not affected by the host–parasite network structure, indicating that associations can also depend solely on the parasite and supertype. This can happen, for example, when a parasite is specific to a host that has low MHC diversity resulting in independence from other host–parasite interactions.

We note that because the parasite–supertype association network is probabilistic in nature, devising a clear-cut criterion or threshold for what is a ‘biologically meaningful’ parasite–supertype association may not be the best practice, and should also greatly depend on the research question and system. In fact, because this is a result of two interconnected networks (Fig. 1), all associations are ‘meaningful’ to some extent because they all mutually affect each other. This reflects well the true biological complexity in interactions between parasites and supertypes mediated through hosts. Parasite–supertype association networks created for other communities may have other properties than those of the network described here, and this is a rich area for future research.

Nature is dominated by complex dynamics between species co-evolution and the structure of the networks they assemble 28. Here, host–parasite co-evolution is also dependent on MHC. Hosts that evolve resistance remove interactions from the host–parasite network (or weaken existing ones) and parasites that evolve new pathways of infection add interactions to the host–parasite network (or strengthen existing ones) through genetic change in MHC alleles. When such dynamics occur at the community level, they can simultaneously affect multiple hosts, parasites, MHC alleles and the interactions among them. For example, an overall reduction in the genetic resistance in a host population to a parasite owing to changes in allele frequencies will cause an increase in the parasites’ prevalence in the host population. If the parasite is not a strict host specialist, then selection imposed by the parasite on individuals of a second host species may also increase, thereby promoting MHC diversity in that host. This process can be extended to multiple hosts and parasites, eventually resulting in a complex network of indirect effects between hosts, parasites and MHC supertypes, whereby parasite-mediated selection on MHC supertypes of one or few hosts can have cascading effects on the MHC diversity of the whole host community.
environmental changes can alter the interactions between hosts and parasites, perturbing the host–supertype network. Human-driven selection to include multiple host–parasite interactions. For instance, according to the ‘rare-allele advantage hypothesis’29, rare alleles in the host population confer better resistance than more frequent alleles because parasites have not had the time to evolve avoiding mechanisms (resistance). The rise of such an allele in a particular host species would weaken the host–parasite interaction in the network, with cascading effects on other species. With time, the rare allele becomes more frequent in the population and thus its selective advantage decreases as parasites evolve resistance. This would strengthen the host–parasite interaction, again affecting other species in the community. Once the allele becomes rare again, parasite adaptation for it also decreases, causing the selective advantage of the allele to increase again. Because changes to one interaction in the network have cascading effects on other interactions, there should be a dynamical interplay between the structural patterns of the host–parasite and host–supertype networks eventually affecting the co-evolution between parasites and the host. This is supported by the predicted nested patterns of parasite-supertype associations and by the modular network structure that indicates high co-evolutionary potential between particular parasites and supertypes.

The association between host–parasite network structure and the diversity of MHC supertypes suggests that a perturbation to the structure of the host–parasite network may result in a concomitant perturbation to the host–supertype network. Human-driven environmental changes can alter the interactions between hosts and parasites30. For example, rodents and their associated parasites can invade disturbed areas, possibly creating new interactions between local rodents and invading parasites31. Such alteration to the host–parasite network may have consequences to the selection pressure imposed by parasites and to the evolution of MHC supertypes through indirect cascading effects. Because rodents are carriers of zoonotic parasites and pathogens32,33 (Supplementary Table 3), the ecological-evolutionary dynamics between altered host–parasite interactions and MHC may have implications for risk of zoonotic disease34. Indeed, the rapid evolutionary potential of both hosts and parasites34 makes genetic change possible even on ecological timescales.

This study is the first to integrate tools and theory of network ecology with immunogenetics, providing novel insights into the evolution of functional MHC alleles in the context of community ecology. In particular, our results lead us to conclude that more than half of the variability in MHC diversity may be shaped by host–parasite interactions in the community, with reciprocal effects of the changes in MHC diversity on host–parasite interactions. This scales our view of MHC diversity from the population level to the community level. The complexity inherent to this model of network-based parasite-mediated selection promotes MHC diversity through indirect pathways, and suggests that the structure of host–parasite interactions plays a major role in the evolution of host immunity at the community level.

**Methods**

**Ethics statement.** All work with animals was approved by the French National Research Agency, project ANR 07 BDIV 012. Animals were treated in accordance with the guidelines of the American Society of Mammalogists and with the European Union legislation (Directive 86/609/EEC). None of the rodent species investigated were on the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) list or on the Red List of the International Union for Conservation of Nature (IUCN). Each trapping session was validated by the national, regional and local health authorities, and under an oral agreement of local landowners. Approval notices for trapping and investigating rodents were given by the Ethical Committee of Mahidol University, Bangkok, Thailand number 0517.11.16/661 and by the Ministry of Health Council of Medical Sciences, National Ethics Committee for Health Research Lao PDR, number 51/NECHR based on the validation of the rodents trapping book protocols of CERoPath16. Cambodia has no ethics committee overseeing animal experimentation. In addition, regional approval was obtained from the regional Head of Veterinary Service (Hérault, France) for sampling and killing rodents and harvesting rodent tissues (approval no. B 34-169-1).

**MHC genotyping.** Since previous studies have shown an association between helminthic infections and certain MHC alleles27,29,30, each individual in the data set (Supplementary Data 1) was genotyped using a portion of the MHC class II DRB region. The 2-construct functional gene that recognizes extracellular peptides combines 2 functional domains (2 functional domains: C2 and C4 domains). This was accomplished by PCR and 454 sequencing, and to determine the genotype of each individual. Gene duplication is common in MHC genes, and, in theory, an individual can have more than two alleles. However, this was ruled out for all the 11 species studied18. Thus, individuals of the rodent species used here had a maximum of two alleles amplified by the PCR primers.

**Supertype clustering.** The validated MHC alleles were clustered with similar antigen binding into supertypes as proposed in ref. 22. The clustering to supertypes was based on 22 amino-acid sites detected as antigen-binding site on the basis of X-ray data of peptide class II protein complexes38,39 and/or detected under positive selection using the HyPre programme40 included in the MEGAS software41. Each of the 22 amino acids forming the antigen-binding site was characterized by five physicochemical descriptor variables: z1 (hydrophobicity), z2 (steric bulk), z3 (polarity), z4 and z5 (electronic effects).

Our final matrix for supertype classification had 276 rows representing unique alleles and 110 columns (5 physicochemical descriptors for 22 sites). We clustered the unique alleles into supertypes using a K-means clustering algorithm and model selection following the procedure described in ref. 23 with the function ‘find.clusters’ from package ‘adegenet’ (version 1.3-9; ref. 42). In brief, the optimal number of supertypes (clusters) is defined as the minimal number of clusters after which the Bayesian Information Criteria (BIC) decreases by a negligible amount as indicated by the curve of BIC values as a function of cluster number (Supplementary Fig. 2). We used the same criterion as recommended in the ‘smoothNgoesup’, in which the selected number of clusters is the one which minimizes the index quantifies the amount of certainty that we can have in our differentiation of clusters. The index quantifies the amount of certainty that we can have in our classification by observing an automated one. If the manual classification is identical to the automated one (clusters in the manual and automated classifications contain the exact same alleles), the index takes its maximum value of 1. If the classifications are totally independent (complete mismatch in the ‘find.clusters’ from package ‘adegenet’ (version 1.3-9; ref. 42), the index value is 0 (see details in ref. 43). The average mutual information across the 500 automatic classifications was 0.85, indicating that our classification was accurate. In Loei, a classification of the 144 alleles found across individuals revealed 18 supertypes, one of which was not harboured by infected individuals and was omitted from analysis. Our manual classification to supertypes is identical to the mean of the automated classification process and fell within its 95% confidence intervals (CIs; mean ± s.d. 23.23 ± 3.7; 95% CI (16.29, 29.56)). Furthermore, we used a measure of normalized mutual information44, based on information theory, to examine the differences between the automated and manual classifications of allele identities. The index quantifies the amount of certainty that we can have in our manual classification by observing an automated one. If the manual classification is identical to the automated one (clusters in the manual and automated classifications contain the exact same alleles), the index takes its maximum value of 1. If the classifications are totally independent (complete mismatch in the ‘find.clusters’ from package ‘adegenet’ (version 1.3-9; ref. 42), the index value is 0 (see details in ref. 43). The average mutual information across the 500 automatic classifications was 0.88, indicating that our classification was accurate. In Loei, a classification of the 144 alleles found across individuals revealed 18 supertypes, one of which was not harboured by infected individuals and was omitted from analysis. Our manual classification to supertypes is identical to the mean of the automated classification process and fell within its 95% CIs (mean ± s.d. 18.28 ± 2.4; 95% CI (12.21, 24.36)). The average normalized mutual information across the 500 automatic classifications was 0.88, indicating that our classification was accurate. Furthermore, a comparison of supertype classification between Loei and the pooled data set resulted in a normalized mutual information of 0.84. This essentially indicated that by knowing the classification in the Loei data set, we can know the classification in the pooled data set and vice versa with 84% certainty, indicating a great overlap in the classification of alleles to particular supertypes even between the data sets.

We matched each individual with their corresponding supertypes such that each individual in the data set was also characterized by a pair of supertypes. To simplify, in this study, the terms genotype, homozygote and heterozygote refer to supertypes, not alleles, and are defined as follows. A supertype is a functional group to which an allele belongs, and a genotype is a combination of two supertypes (in case of a heterozygote) or one supertype (in case of a homozygote) to which the two alleles of an individual belong. For example, if alleles A and B were classified as supertype no. 1, and allele C was classified as supertype...
no. 2, an individual bearing alleles A and B would be a homozygote with a genotype 1-1, whereas an individual bearing alleles B and C would be a heterozygote with a genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2. At the individual host level, susceptibility to helminths can result from one of the two alleles or from a co-dominant effect of both, also depending on the genotype 1-2.
Additional information
Accession codes: The demultiplexed raw sequences generated in this study have been deposited in NCBI Short Read Archive (SRA) under the accession code SRP047116.
Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications
Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Pilosof, S. et al. Host–parasite network structure is associated with community-level immunogenetic diversity. Nat. Commun. 5:5172 doi: 10.1038/ncomms6172 (2014).