Combining spatial analysis and host population genetics to gain insights into the mode of transmission of a pathogen: The example of Toxoplasma gondii in mice

Lokman Galal, Claire Stragier, Farid Boumédiène, Azra Hamidović, Océane Maugrion, Marie-Laure Dardé, Aurélien Mercier

Combining spatial analysis and host population genetics to gain insights into the mode of transmission of a pathogen: the example of *Toxoplasma gondii* in mice

Lokman Galal[a,*], Claire Stragier[b], Farid Boumédiène[a], Azra Hamidović[a], Océane Maugrion[a], Marie-Laure Dardé[a,c], Aurélien Mercier [a,c]

[a] *INSERM UMR_S 1094, Neuroépidémiologie Tropicale, Laboratoire de Parasitologie-Mycologie, Faculté de Médecine, Université de Limoges, Limoges 87025, France*

[b] BIOPASS (IRD-CBGP, ISRA, UCAD), Campus de Bel-Air, BP 1386, CP 18524 Dakar, Senegal.

[c] *Centre National de Référence Toxoplasmose/Toxoplasma Biological Resource Center, CHU Limoges, 87042 Limoges, France*

*Corresponding author: Lokman Galal, INSERM UMR_S 1094, Neuroépidémiologie Tropicale, Laboratoire de Parasitologie-Mycologie, Faculté de Médecine, Université de Limoges, Limoges 87025, France. Tel.: +33 (0) 5 19 56 42 64 ; Fax: +33 (0) 5 55 43 58 21.

E-mail address: lokmanmagalal@gmail.com

Abstract

Toxoplasma gondii is an ubiquitous highly prevalent zoonotic protozoan. Cats are the definitive hosts, while all other warm-blooded animals are intermediate hosts for this parasite. Commensal rodents, being the main prey of cats, are probably the major reservoir for *T. gondii* in the domestic environment. Rodents can acquire infection after ingestion of oocysts that have sporulated in the environment. However, experimental evidence shows that vertical transmission can be sufficient for the perpetuation of transmission between generations of mice. In natural settings, the relative epidemiological importance of vertical transmission over oral transmission is a matter of debate and raises the question of the possibility of a *T. gondii* cycle in the absence of cats. In the present study, we took advantage of an extensive survey of commensal rodents in Dakar, Senegal, where the house mouse is the predominant putative reservoir of *T. gondii*. Mice genotypes and spatial location through GPS referencing of all trapping localizations were investigated in relation to *T. gondii* infection in eight sites of the city of Dakar and on Goree Island. In each sampling site, the occurrence of over-prevalence zones of *T. gondii* infection was investigated through Kulldorf's statistic using SaTScan software. Genetic structure and relatedness between mice were investigated within each over-prevalence zone, in order to find clues of transmission between related mice. Within each of the four over-prevalence zones identified across nine sites, infected mice belonged to more than one genetic group. No association between the degree of relatedness and the occurrence of *T. gondii* infection could be detected. These findings suggest an environmental source of infection for mice associated with localized putative foci of environmental contamination and support an oral route of infection for mice from Dakar rather than a cycle based on vertical transmission. However, further investigations based on a denser sampling in different epidemiological contexts are recommended.

Toxoplasma gondii; *Mus musculus*; spatial analysis; relatedness; transmission; vertical.

1. Introduction

Toxoplasmosis is an ubiquitous parasitic zoonosis, caused by an obligate intracellular protozoan parasite *Toxoplasma gondii*. In humans, *T. gondii* infection is often subclinical, except in

some risk groups like the developing foetus (in case of congenital infection) and immunocompromised patients, for which toxoplasmosis can have severe health consequences (Montoya and Liesenfeld, 2004). Felids are the definitive hosts, with the domestic cat being the unique definitive host in the domestic environment, while all other warm-blooded animals are intermediate hosts for this parasite (Frenkel et al., 1970; Dubey, 2009). Birds and mammals, including humans, often develop dormant tissue cysts after ingestion of oocysts shed in the environment by cats in the form of contaminated faeces (Dubey, 2009). Commensal rodents, being the main prey of cats (Turner and Bateson, 2013), are probably the most important reservoirs for *T. gondii* in the domestic environment (Dubey et al., 1995; Hejlícek et al., 1997).

Experimental studies have shown that rodents are able to get infected by *T. gondii* through several routes (Dubey, 2009). Infection through oocyst ingestion is universally admitted as a conventional infection route for most of intermediate host species including rodents, although a high inoculum dose can be fatal in some susceptible hosts such as the house mouse *Mus musculus* (Owen and Trees, 1998). In addition, several studies have shown that an infected female rodent can vertically transmit the parasite to her offspring, during pregnancy due to transplacental passage of parasites to foetuses, or to neonates during lactation (Dubey and Shen, 1991; Roberts and Alexander, 1992; Zenner et al., 1993; Dubey et al., 1997; Owen and Trees, 1998; Paulino and Vitor, 1999; Marshall et al., 2004; Freyre et al., 2009). Vertical transmission of *T. gondii* through successive generations of chronically infected rodents without exogenous reinfection was demonstrated only in outbred laboratory mice (Beverley, 1959) and wild mice *Mus musculus* from the United Kingdom (Owen and Trees, 1998). Infected mothers transmitted the infection to the majority of their offspring from each litter (Owen and Trees, 1998; Marshall et al., 2004). In contrast, in BALB / c mice, hamsters, and in most of the laboratory rat lineages, vertical transmission was noticed mainly when the infection occured during pregnancy, but seldom in individuals with a chronic infection (Dubey and Frenkel, 1973; Dubey and Shen, 1991; Roberts and Alexander, 1992; Zenner et al., 1993; Dubey et al., 1997; Paulino and Vitor, 1999; Freyre et al., 2009; Hide, 2016). Hence, the modes of transmission of *T. gondii* in natural populations of rodents are still unresolved and are a matter of debate (Dubey, 2009).

Vertical transmission in intermediate hosts, bypassing sexual multiplication in the definitive host, was proposed as a possible explanation for the clonal structure of most *T. gondii* populations (Johnson, 1997; Worth et al., 2013). However, oral transmission could also sustain the clonality of *T. gondii* populations, as cats simultaneously infected with different strains of *T. gondii* are likely to be very rare events in nature (Howe and Sibley, 1995), and hence only selfing of the infecting strain usually occurs in the cat's gut. In addition, the ability of *T. gondii* to modify the behaviour of rodents to presumably facilitate their predation and its trophic transmission to cats (Webster et al., 1994; Berdoy et al., 2000; Gonzalez et al., 2007; Vyas et al., 2007) suggests that oral transmission is of key importance in this current pattern of host-parasite coevolution. Evaluating the relative frequency of vertical transmission over oral transmission in rodents in natural settings may help in the understanding of the mechanisms underlying the clonality of *T. gondii* populations. A previous study relied on comprehensive lambing records to define families within a flock of sheep, in order to test the hypothesis of possible successive vertical transmission of *T. gondii* within families (Morley et al., 2005). However, investigating an association between families and infection occurrence may be much more challenging to test in wild species given the difficulty of reconstructing families *a posteriori*, using molecular techniques for example (Jones et al., 2010).

In the present study, the objective was to provide insight into the transmission pathways of *T. gondii* within wild populations of house mice (*Mus musculus domesticus*) from Dakar, Senegal (West Africa). In Senegal, and more widely in West Africa, the genetic structure of *T. gondii* populations was shown to be strongly clonal, with few clonal lineages representing the majority of parasitic strains (Galal et al., 2018, 2019a). Here, we relied on data on trapping location, microsatellite genotypes and infectious status for *T. gondii* for mice from several sites in Dakar. The spatial heterogeneity of infected cases was investigated to test if the occurrence of *T. gondii* infection could be associated to putative foci of environmental contamination within the population home range. To this end, the occurrence of possible over-prevalence zones of *T. gondii* infection among mice was investigated within each site. In addition, the genetic structure of local mice populations was analysed and the degree of relatedness between mice was calculated, in order to test if the infected individuals were more related to each other than to non-infected mice. We hypothesized that relationships between

genetic structure or relatedness and infection could be expected in the case of vertical transmission of *T. gondii*.

2. Materials and methods

We took advantage of an extensive survey of urban rodents in the city of Dakar (Stragier et al., 2019). Briefly, sampling was carried out in 12 sites of the Cape Verde Peninsula and in the Goree Island (Figure 1), each site being separated from each other by a minimum distance of 600 m and covering a median surface of 0.04 km². Two traps (one wire mesh trap and one Sherman trap) were set per room or courtyard in buildings such as dwelling houses, boutiques, workshops, offices or warehouses, and locations were precisely recorded with a GPS device. This survey led to the sampling of 473 mice, which were genotyped using a set of 15 microsatellite markers. Details about genotyping methodology and basic analysis of genetic data (allelic richness, genotypic differentiation between sampling sites, deviation from Hardy-Weinberg equilibrium) are reported elsewhere (Stragier et al., 2019). Mice were previously screened for chronic infection with *T. gondii* using a real-time PCR assay that target the 529 bp repeat region of *T. gondii* genome (Galal et al., 2019b). Among the 13 sampled sites considered for this study, four were excluded from the analysis. In the first two sites (*Hann-Pecheur* and *Yoff*), the single PCR-positive mouse detected did not enable to investigate the spatial heterogeneity of infected cases. In the two other sites (*International Port of Dakar* and *Parc de Hann*), the sampling coverage of the site was highly unbalanced, with important distances (> 500 meters) between the sampled sectors within the site. Sampling size, number of infected individuals and prevalence data per site are summarized in Table S1.

2.1. Spatial clustering analysis

Within each of the 9 sites included in this study, purely spatial cluster analysis was performed to test whether the infected mice were distributed randomly over space and, if not, to identify significant over-prevalence zones using Kulldorf's statistic. An over-prevalence zone was defined as a circular geographic area within a site in which the number of observed cases was significantly higher than the

number expected if cases were randomly distributed in space. Building units where rodent traps were set during the survey were considered as the smallest spatial unit in this study. 'Spatial scan statistics' which relies on Kulldorf's statistic was used to test the null hypothesis that the relative risk (RR) of *T. gondii* infection was the same between any building groups (or collection of building groups) and the remaining building groups of a sampling site. SaTScan software version 9.4.4 (Kulldorff, 1997, 2010), designed specifically to implement this test and using a Bernoulli model, imposed a circular scanning window on the map that moves across space. The area within the circular window, centered on the centroid of each building unit, varied in size from zero to a maximum radius never including more than 50% of the total population within a given site. The SaTScan software tested for possible over-prevalence zones within the variable window around the centroid of each building group. The number of Monte Carlo replications for statistical inference was set to 999, and over-prevalence zones with statistical significance of $p < 0.05$ were reported.

2.2. Genetic structure and relatedness within over-prevalence zones

Within each site exhibiting an over-prevalence zone, analyses of genetic structure and relatedness were performed in relation to *T. gondii* infection.

The possible occurrence of distinct genetic groups of mice was investigated using Discriminant Analysis on Principal Components (DAPC) (Jombart et al., 2010). DAPC was performed using the adegenet package (Jombart, 2008) for the R 3.4.0 software (R Development Core Team, 2009). This Bayesian clustering method is not based on a predefined population genetics model and is thus free from Hardy–Weinberg equilibrium assumptions. Prior clusters were defined by sequential K-means and a range of appropriate clustering solutions based on the Bayesian information criterion (BIC) was defined. The optimal number of clusters selected was the one for which the BIC showed the lowest value and after which BIC increased or decreased by the least amount. Individuals were considered having an admixed genotype when they exhibited no more than 90% of probability of membership in a single cluster (Jombart and Collins, 2015).

Relatedness (r) for all pairs of mice within each site was estimated through maximum-likelihood methods using ML-Relate (Kalinowski et al., 2006). ML-Relate uses the downhill simplex routine to

find the maximum likelihood estimate of *r*. The *r* values of pairs composed of two infected mice (named group A) were compared to those of pairs composed of an infected mouse and a non-infected mouse (named group B). The Mann-Whitney U test was used to determine statistical significant difference between the two groups in XLSTAT 2017 for Windows.

2.3. Concordance between DAPC clustering and relatedness

To evaluate the concordance between DAPC clustering and relatedness, we compared the r values of pairs of mice from identical DAPC clusters (intra-cluster pairs) to those obtained for pairs of mice from distinct DAPC clusters (inter-cluster pairs). Therefore, the r values were organized in groups defined according to the DAPC cluster to which belongs each of the two mice composing a pair (DAPC1/DAPC1, DAPC1/DAPC2, DAPC2/DAPC2, etc.). The distribution of r values between groups was compared using the Kruskal-Wallis test. *Post hoc* pairwise comparisons of the distributions of r values between groups were performed using the Dunn's procedure, which applies the Bonferroni's correction of p-values. This analysis was performed using XLSTAT 2017 for Windows.

3. Results

3.1. Spatial clustering analysis

Purely spatial cluster analysis revealed the occurrence of over-prevalence zones in four sites (Table 1): Goree Island (Fig. S2), *Grand Dakar* (Fig. S3), *Ouakam* (Fig. S5) and *Plateau Reubeuss* (Fig. S7).

3.2. Genetic structure and relatedness

The occurrence of more than one genetic group was detected in seven of the nine sites included in this study using DAPC (Table 2). In the seven sites where genetic structure was found, most individuals showed clear assignment to a single DAPC cluster, as the percentage of individuals with an admixed DAPC genotype was 18.30% on average (ranged from 0.00% to 36.36% according to site). A more or less marked spatial aggregation of mice of the same DAPC cluster was noticed in most sites

(Fig. S1-S7). This pattern of spatial aggregation by DAPC cluster was strong in *Colobane*, *Grand Dakar*, *Ouakam* and *Gueule Tapée* and less pronounced in Goree Island, *Patte d'oie* and *Plateau Reubeuss* (Fig. S1-S7). Within the four over-prevalence zones previously identified, more than one DAPC cluster was found, and infected individuals also belonged to more than one DAPC cluster (Table 2; Fig.S2, S3, S5 and S7).

In two of the four over-prevalence zones previously identified (in *Grand Dakar* and *Ouakam*), only infected mice were found. We therefore could not compare the r values of pairs composed of two infected mice (group A) to the r values of pairs of mice composed of an infected mouse and non-infected mouse (group B) within these two over-prevalence zones. In the over-prevalence zone occurring in Goree Island, r values ranged from 0 to 0.62 for group A (n=42), with an average of 0.13 and from 0 to 0.91 for group B (n=196), with an average of 0.09. In the over-prevalence zone occurring in *Plateau Reubeuss*, r values ranged from 0 to 0.37 for group A (n=30), with an average of 0.06 and from 0 to 0.25 for group B (n=132), with an average of 0.03. Mann-Whitney test showed no significant difference between group A and group B in both Goree Island and *Plateau Reubeuss*, with p-values of 0.202 and 0.267, respectively.

3.3. Concordance between DAPC clustering and relatedness

As r values were only calculated between pairs of mice from Goree Island and *Plateau Reubeuss*, the comparison between DAPC and relatedness was only possible for these two sites.

In Goree Island, 15 groups were defined for r values according to the DAPC cluster to which belongs each of the two mice composing a pair. Five groups were composed of intra-cluster pairs of mice (which correspond to the five DAPC clusters identified in this site) and 10 groups were composed of inter-cluster pairs of mice. The distribution of r values between groups was significantly different using the Kruskal-Wallis test (p-value < 0.0001). The distribution of r values in groups composed of intra-cluster pairs of mice stochastically dominated (significance threshold of 0.0005 after Bonferroni's correction) groups composed of inter-cluster pairs in all pairwise comparisons, with the exception of an absence of significant difference between group 1 and group 3 (Table S2).

In *Plateau Reubeuss*, six groups were defined for r values according to the DAPC cluster to which belongs each of the two mice composing a pair. Three groups were composed of intra-cluster pairs of mice (which correspond to the three DAPC clusters identified in this site) and three groups were composed of inter-cluster pairs of mice. The distribution of r values between groups was significantly different using the Kruskal-Wallis test (p-value < 0.0001). The distribution of r values in groups composed of intra-cluster pairs of mice stochastically dominated (significance threshold of 0.003 after Bonferroni's correction) groups composed of inter-cluster pairs in all pairwise comparisons (Table S3).

4. Discussion

In this study, we aimed to evaluate whether spatial patterns of *T. gondii* infection in mice may be related or not to the genetic structure of hosts, giving clues about the mode of transmission.

A previous study investigating oocyst spatial distribution in an urban area has demonstrated highly localized foci of oocysts' occurrence, associated with cats' defecation sites (Afonso et al., 2008). Therefore, we hypothesized that mice living on areas of recent oocyst shedding would be predominantly exposed to the oocysts found on the soil, resulting in a spatial aggregation of infected mice. We evaluated whether or not the spatial distribution of infected cases was random, by investigating if infected mice gathered in over-prevalence zones of *T. gondii* infection. Our results showed that the spatial distribution of mice infected with *T. gondii* was not random within several sites in Dakar, as over-prevalence zones of *T. gondii* infection could be identified in four sites. Such a pattern may correspond to (1) an environmental source of infection for mice associated to localized putative foci of environmental contamination, but also to (2) vertical transmission, given that mice are known to live in extended family groups occupying a small territory (Berry and Bronson, 1992; Pocock et al., 2004). The analyses of genetic clustering using DAPC showed that mice within over-prevalence zones belonged to more than one genetic group. This pattern observed in all the four over-prevalence zones identified in this study suggests that infected mice within an over-prevalence zone can exhibit some degree of genetic distance, and that genetic proximity (according to DAPC clustering) is not associated with *T. gondii* infection in mice. In the two sites (Goree Island and

Plateau Reubeuss) having over-prevalence zones with infected and non-infected mice, we also showed that relatedness was not higher in pairs of infected mice than in pairs involving infected and non-infected mice. This result suggests no effect of the relatedness on mice infection. In these two sites, the r values of pairs of mice from identical DAPC clusters (intra-cluster pairs) were generally significantly higher than those obtained for pairs of mice from distinct DAPC clusters (inter-cluster pairs). This comparison enabled to verify the concordance between the two types of genetic analysis. Finally, no clues supporting an association between *T. gondii* infection and genetic structure could be found, suggesting that vertical transmission has no significant role in *T. gondii* infection and that infection occurrence can only be associated to the location.

However, the approach used in this study, combining host genetics and spatial analysis of *T. gondii* prevalence, faced some limitations in several sites. In five sites, no over-prevalence zones could be detected, which may be attributed to gaps in sampling, the spatial overlapping (partial or total) of several contamination foci within a site — if we assume a predominant role of oral transmission — or the movements of infected individuals within the site. In addition, no genetic structure could be identified using DAPC in two of these sites to test the effect of genetic structure on the distribution of infected individuals. Furthermore, DAPC is known to group together individuals with similar genetic profiles in terms of allele frequencies, but inferred groups do not necessarily correspond to family groups. Studies in the lab or in the wild have shown that mice populations are subdivided into social breeding groups, termed demes, where gene flow is restricted by behavioural barriers (Petras, 1967; Anderson, 1970; Selander, 1970; De Fries, 1972; Berry and Jakobson, 1975; Singleton, 1983). This social organisation may contribute to the occurrence of a genetic structure within populations of mice. Although populations inferred from DAPC could be composed of related individuals, they may also be composed of unrelated individuals with close genetic profiles. Finally, the significance of relatedness inferences could have been affected by the small sampling from each site when larger sample sizes are usually required to infer relatedness from wild populations (Jones et al., 2010).

Overall, despite the limitations that we have pointed out, the results provided by this study suggest that *T. gondii* infection in mice may be more associated to location than to genetic proximity or

degree of relatedness between mice. Transmission is hence more likely to occur through an oral route than by vertical transmission in mice from Dakar.

This conclusion appears to be in contradiction with a number of field studies in which the results supported a vertical transmission of *T. gondii* in several species of rodents (Murphy et al., 2008; Thomasson et al., 2011; Webster, 1994). In those previous studies, high prevalence levels of *T. gondii* infection were reported among rodents in areas that appear to be relatively free of cats. These findings provide indirect evidence that vertical transmission can be sufficient to the perpetuation of transmission. In line with this, a field study was performed on wood mice *Apodemus sylvaticus* in a location where low a cat density (Bajnok et al., 2015). The wood mice from this location were found to belong to four genetically distinct subpopulations defined with STRUCTURE (Pritchard et al., 2000), distributed in different parts of the study zone. The prevalence of infection was found to be significantly different in each of the subpopulations and was linked to host genotype rather than location of capture. This suggested that parasite infection was associated to subpopulations and non-randomly distributed throughout the populations, which is in accordance with a predominantly vertical route of transmission. However, as for the DAPC is the present study, it is unknown whether or not these populations defined by STRUCTURE truly correspond to family groups. In addition, an important aspect to point out is that all these studies were conducted in the United Kingdom and in areas free (or nearly free) of cats, limiting the diversity of epidemiological situations in the exploration of this issue. Indeed, in situations where cats densely populate an area — as it is the case in Dakar, Senegal (Lahamdi, 1992; Bend, 2006) — and in which rodents are therefore heavily exposed to oocysts, the high predation pressure could reduce the rodents' overall lifespan and the likelihood of reaching sexual maturity to transmit the infection to offspring through vertical transmission (Turner et al., 2013). In particular, assuming a role of parasitic manipulation in facilitating the predation of rodents by cats, a primary involvement of vertical transmission could be detrimental for the sustainability of the cycle, as infected individuals would be less likely to reach sexual maturity and reproduce compared to non-infected individuals as they would die earlier in age.

5. Conclusions and perspectives

Altogether, the factors quoted above emphasize on the importance of considering the specificities of each epidemiological situation in inferring on transmission dynamics from empirical data. Cat and rodent density, rodent host and parasite genetic determinants probably influence the mode of transmission of *T. gondii*, by favouring — or not — vertical transmission (Lélu et al., 2013; Turner et al., 2013). The absence of a validated and standardized methodology to evaluate the relative epidemiological importance of vertical transmission over oral transmission in rodents in natural settings may also explain the inconsistencies between the results of different studies. The findings of the present study provide interesting insights into the transmission routes of *T. gondii* in natural populations of mice, although the approach we used showed certain limitations in a number of situations.

Estimating the epidemiological importance of vertical transmission based on the levels of prevalence in embryos collected from pregnant dams may not necessarily be the most appropriate approach to address this question. Although a high frequency of *T. gondii* transmission from infected females to foetuses is observed (Beverley, 1959; Owen and Trees, 1998; Marshall et al., 2004), experimental congenital infection in mice was shown to be associated with heavy mortalities among the offspring and does significantly reduce the number reared to maturity in each litter (Beverley, 1959). This may limit the transmission of the parasite between generations of mice. Another appealing approach could be the genotyping of *T. gondii* strains infecting mice of distinct family groups which do not segregate in space to see if they are infected by different strains or not. The limited sensitivity of the available genotyping techniques would probably hamper the application of this approach at the present time as the tissue burden of *T. gondii* is too low for genotyping in most individuals (Galal et al., 2019b). Indeed, among 89 PCR-positive mice (from a total of 671 mice), only 6 *T. gondii* genotypes could be obtained (refer to Supplementary Fig. S1 in Galal et al., 2019b).

Future field studies should rely on a denser sampling in smaller spatial scales, which could allow a more accurate determination of parenthood between trapped individuals. A two steps sampling approach could be followed: the first step would be to conduct an extensive sampling (density of traps similar to the one followed in this study) to detect areas where infected mice appear to aggregate, and the second step would be to maximize the sampling efforts on those areas in order to obtain a robust sampling for kinship analyses in relation to infection. In parallel, fine-scale environmental specificities

that could explain the spatial heterogeneity in the distribution of infected cases should be thoroughly investigated. This part of the survey should include an estimation of the spatial variability in cat density and the identification of cat defection sites as possible foci of environmental contamination with oocysts. Another important aspect is the social organization of wild mice populations and their spatiotemporal evolution which are still poorly understood. A better understanding of the family structure of mice and the interactions between family groups would be useful to validate kinship analysis and to understand how *T. gondii* and other pathogens may spread between mice.

Acknowledgements

We thank Karine Berthier and Carine Brouat for her methodological advices. We thank the French Agence Nationale de la Recherche (ANR project IntroTox 17-CE35-0004) and the Nouvelle-Aquitaine region for funding this research.

References

Afonso, E., Lemoine, M., Poulle, M.-L., Ravat, M.-C., Romand, S., Thulliez, P., Villena, I., Aubert, D., Rabilloud, M., Riche, B., Gilot-Fromont, E., 2008. Spatial distribution of soil contamination by *Toxoplasma gondii* in relation to cat defecation behaviour in an urban area. Int. J. Parasitol. 38, 1017–1023. https://doi.org/10.1016/j.ijpara.2008.01.004

Anderson, P.K., 1970. Ecological structure and gene flow in small mammals. Mammal Review 1, 29–30.

Bajnok, J., Boyce, K., Rogan, M.T., Craig, P.S., Lun, Z.R., Hide, G., 2015. Prevalence of *Toxoplasma gondii* in localized populations of *Apodemus sylvaticus* is linked to population genotype not to population location. Parasitology 142, 680–690. https://doi.org/10.1017/S0031182014001760

Bend, R.L., 1980. Enquête coprologique sur la toxoplasmose dans la population des chats de la ville de Dakar. Université Cheikh Anta Diop de Dakar. Thesis n°6 Dakar 2006.

Berdoy, M., Webster, J.P., Macdonald, D.W., 2000. Fatal attraction in rats infected with *Toxoplasma gondii*. Proc. Biol. Sci. 267, 1591–1594. https://doi.org/10.1098/rspb.2000.1182

Berry, R.J., Bronson, F.H., 1992. Life history and bioeconomy of the house mouse. Biological Reviews 67, 519–550.

Berry, R.J., Jakobson, M.E., 1975. Ecological genetics of an island population of the house mouse (Mus musculus). Journal of zoology 175, 523–540.

Beverley, J.K.A., 1959. Congenital Transmission of Toxoplasmosis through Successive Generations of Mice. Nature 183, 1348. https://doi.org/10.1038/1831348a0

De Fries, J.C., 1972. Behavioural genetics and the fine structure of mouse population: a study in microevolution. Evol. Biol. 5, 279–291.

Dubey, J., 2009. Toxoplasmosis in animals and humans. Boca Raton: CRC Press.

Dubey, J.P., Frenkel, J.K., 1973. Experimental *Toxoplasma* Infection in Mice with Strains Producing Oocysts. The Journal of Parasitology 59, 505–512. https://doi.org/10.2307/3278784

Dubey, J.P., Shen, S.K., 1991. Rat model of congenital toxoplasmosis. Infect. Immun. 59, 3301–3302.

Dubey, J.P., Shen, S.K., Kwok, O.C., Thulliez, P., 1997. Toxoplasmosis in rats (*Rattus norvegicus*): congenital transmission to first and second generation offspring and isolation of *Toxoplasma gondii* from seronegative rats. Parasitology 115 (Pt 1), 9–14.

Dubey, J.P., Weigel, R.M., Siegel, A.M., Thulliez, P., Kitron, U.D., Mitchell, M.A., Mannelli, A., Mateus-Pinilla, N.E., Shen, S.K., Kwok, O.C., 1995. Sources and reservoirs of *Toxoplasma gondii* infection on 47 swine farms in Illinois. J. Parasitol. 81, 723–729.

Frenkel, J.K., Dubey, J.P., Miller, N.L., 1970. *Toxoplasma gondii* in cats: fecal stages identified as coccidian oocysts. Science 167, 893–896.

Freyre, A., Fialho, C.G., Bigatti, L.E., Araujo, F. a. P., Falcón, J.D., Mendez, J., González, M., 2009. *Toxoplasma gondii*: congenital transmission in a hamster model. Exp. Parasitol. 122, 140–144. https://doi.org/10.1016/j.exppara.2009.02.004

Galal, L., Ajzenberg, D., Hamidović, A., Durieux, M.-F., Dardé, M.-L., Mercier, A., 2018. *Toxoplasma* and Africa: One Parasite, Two Opposite Population Structures. Trends Parasitol. 34, 140–154. https://doi.org/10.1016/j.pt.2017.10.010

Galal, L., Sarr, A., Cuny, T., Brouat, C., Coulibaly, F., Sembène, M., Diagne, M., Diallo, M., Sow, A., Hamidović, A., Plault, N., Dardé, M.-L., Ajzenberg, D., Mercier, A., 2019a. The introduction of new hosts with human trade shapes the extant distribution of *Toxoplasma gondii* lineages. PLoS Negl Trop Dis 13, e0007435. https://doi.org/10.1371/journal.pntd.0007435

Galal, L., Schares, G., Stragier, C., Vignoles, P., Brouat, C., Cuny, T., Dubois, C., Rohart, T., Glodas, C., Dardé, M.-L., Kane, M., Niang, Y., Diallo, M., Sow, A., Aubert, D., Hamidović, A., Ajzenberg, D., Mercier, A., 2019b. Diversity of *Toxoplasma gondii* strains shaped by commensal communities of small mammals. Int. J. Parasitol. 49, 267–275. https://doi.org/10.1016/j.ijpara.2018.11.004

Gonzalez, L.E., Rojnik, B., Urrea, F., Urdaneta, H., Petrosino, P., Colasante, C., Pino, S., Hernandez, L., 2007. *Toxoplasma gondii* infection lower anxiety as measured in the plus-maze and social interaction tests in rats A behavioral analysis. Behav. Brain Res. 177, 70–79. https://doi.org/10.1016/j.bbr.2006.11.012

Hejlícek, K., Literák, I., Nezval, J., 1997. Toxoplasmosis in wild mammals from the Czech Republic. J. Wildl. Dis. 33, 480–485. https://doi.org/10.7589/0090-3558-33.3.480

Hide, G., 2016. Role of vertical transmission of *Toxoplasma gondii* in prevalence of infection. Expert Rev Anti Infect Ther 14, 335–344. https://doi.org/10.1586/14787210.2016.1146131

Howe, D.K., Sibley, L.D., 1995. *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. J. Infect. Dis. 172, 1561–1566.

Johnson, A.M., 1997. Speculation on possible life cycles for the clonal lineages in the genus *Toxoplasma*. Parasitol. Today (Regul. Ed.) 13, 393–397. https://doi.org/10.1016/s0169-4758(97)01129-0

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Collins, C., 2015. A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0. London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling.

Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11, 94. https://doi.org/10.1186/1471-2156-11-94

Jones, A.G., Small, C.M., Paczolt, K.A., Ratterman, N.L., 2010. A practical guide to methods of parentage analysis. Mol Ecol Resour 10, 6–30. https://doi.org/10.1111/j.1755-0998.2009.02778.x

Kalinowski, S.T., Wagner, A.P., Taper, M.L., 2006. ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. Molecular Ecology Notes 6, 576–579.

Kulldorff, M., 2010. SaTScan-Software for the spatial, temporal, and space-time scan statistics. Boston: Harvard Medical School and Harvard Pilgrim Health Care.

Kulldorff, M., 1997. A spatial scan statistic. Communications in Statistics - Theory and Methods 26, 1481–1496. https://doi.org/10.1080/03610929708831995

Lahamdi, A., 1992. Etude Comparative de deux Techniques Serologiques: ELISA et IFI appliquées au Serodiagnostic de la Toxoplasmose ovine dans les quartiers de Dakar et Banlieue. Université Cheikh Anta Diop de Dakar. Thesis n°38 Dakar 1992.

Lélu, M., Langlais, M., Poulle, M.-L., Gilot-Fromont, E., Gandon, S., 2013. When should a trophically and vertically transmitted parasite manipulate its intermediate host? The case of *Toxoplasma gondii*. Proc. Biol. Sci. 280, 20131143. https://doi.org/10.1098/rspb.2013.1143

Marshall, P.A., Hughes, J.M., Williams, R.H., Smith, J.E., Murphy, R.G., Hide, G., 2004. Detection of high levels of congenital transmission of *Toxoplasma gondii* in natural urban populations of *Mus domesticus*. Parasitology 128, 39–42. https://doi.org/10.1017/s0031182003004189

Montoya, J.G., Liesenfeld, O., 2004. Toxoplasmosis. Lancet 363, 1965–1976. https://doi.org/10.1016/S0140-6736(04)16412-X

Morley, E.K., Williams, R.H., Hughes, J.M., Terry, R.S., Duncanson, P., Smith, J.E., Hide, G., 2005. Significant familial differences in the frequency of abortion and *Toxoplasma gondii* infection within a flock of Charollais sheep. Parasitology 131, 181–185. https://doi.org/10.1017/s0031182005007614

Murphy, R.G., Williams, R.H., Hughes, J.M., Hide, G., Ford, N.J., Oldbury, D.J., 2008. The urban house mouse (*Mus domesticus*) as a reservoir of infection for the human parasite *Toxoplasma gondii*: an unrecognised public health issue? Int J Environ Health Res 18, 177–185. https://doi.org/10.1080/09603120701540856

Owen, M.R., Trees, A.J., 1998. Vertical transmission of *Toxoplasma gondii* from chronically infected house (*Mus musculus*) and field (*Apodemus sylvaticus*) mice determined by polymerase chain reaction. Parasitology 116 (Pt 4), 299–304.

Paulino, J.P., Vitor, R.W., 1999. Experimental congenital toxoplasmosis in Wistar and Holtzman rats. Parasite 6, 63–66. https://doi.org/10.1051/parasite/1999061063

Petras, M.L., 1967. Studies of Natural Populations of Mus. II. Polymorphism at the T Locus. Evolution 21, 466–478. https://doi.org/10.2307/2406608

Pocock, M.J., Searle, J.B., White, P.C., 2004. Adaptations of animals to commensal habitats: population dynamics of house mice *Mus musculus domesticus* on farms. Journal of Animal Ecology 73, 878–888.

Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.

R Development Core Team, 2009. R: a language and environment for statistical computing [WWW Document]. URL https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing (accessed 1.29.18).

Roberts, C.W., Alexander, J., 1992. Studies on a murine model of congenital toxoplasmosis: vertical disease transmission only occurs in BALB/c mice infected for the first time during pregnancy. Parasitology 104 Pt 1, 19–23.

Selander, R.K., 1970. Behavior and genetic variation in natural populations. American zoologist 10, 53–66.

Singleton, G.R., 1983. The social and genetic structure of a natural colony of house mice, *Mus musculus*, at Healesville WildlifeSanctuary. Australian Journal of Zoology 31, 155–166.

Stragier, C., Piry, S., Loiseau, A., Kane, M., Sow, A., Niang, Y., Diallo, M., Ndiaye, A., Gauthier, P., Borderon, M., Granjon, L., Brouat, C., Berthier, K., 2019. Impact of historical and current features of the cityscape on the genetic structure of the house mouse (*Mus musculus domesticus*) in Dakar (Senegal, West Africa). bioRxiv 557066. https://doi.org/10.1101/557066

Thomasson, D., Wright, E.A., Hughes, J.M., Dodd, N.S., Cox, A.P., Boyce, K., Gerwash, O., Abushahma, M., Lun, Z.-R., Murphy, R.G., Rogan, M.T., Hide, G., 2011. Prevalence and co-infection of *Toxoplasma gondii* and *Neospora caninum* in *Apodemus sylvaticus* in an area relatively free of cats. Parasitology 138, 1117–1123. https://doi.org/10.1017/S0031182011000904

Turner, D.C., Bateson, P., 2013. The Domestic Cat: The Biology of its Behaviour. Cambridge University Press.

Turner, M., Lenhart, S., Rosenthal, B., Zhao, X., 2013. Modeling effective transmission pathways and control of the world's most successful parasite. Theor Popul Biol 86, 50–61. https://doi.org/10.1016/j.tpb.2013.04.001

Vyas, A., Kim, S.-K., Giacomini, N., Boothroyd, J.C., Sapolsky, R.M., 2007. Behavioral changes induced by *Toxoplasma* infection of rodents are highly specific to aversion of cat odors. Proc. Natl. Acad. Sci. U.S.A. 104, 6442–6447. https://doi.org/10.1073/pnas.0608310104

Webster, J.P., 1994. Prevalence and transmission of *Toxoplasma gondii* in wild brown rats, *Rattus norvegicus*. Parasitology 108 (Pt 4), 407–411.

Webster, J.P., Brunton, C.F., MacDonald, D.W., 1994. Effect of *Toxoplasma gondii* upon neophobic behaviour in wild brown rats, Rattus norvegicus. Parasitology 109 (Pt 1), 37–43.

Worth, A.R., Lymbery, A.J., Thompson, R.C.A., 2013. Adaptive host manipulation by *Toxoplasma gondii*: fact or fiction? Trends Parasitol. 29, 150–155. https://doi.org/10.1016/j.pt.2013.01.004

Zenner, L., Darcy, F., Cesbron-Delauw, M.F., Capron, A., 1993. Rat model of congenital toxoplasmosis: rate of transmission of three *Toxoplasma gondii* strains to fetuses and protective effect of a chronic infection. Infect. Immun. 61, 360–363.

Figure 1. Sampling sites of *Mus musculus domesticus* in Dakar and on Goree Island, Senegal.

sites	number of over-prevalence zones ($p < 0.05$)	p-value	SIR (95% CI)	number of mice within the over-prevalence zone	number of infected mice within the over-prevalence zone	total number of building units within the over-prevalence	number of building units with infected mice	radius of the over-prevalence zone (in meters)
Colobane	0							
Goree Island	1	0.007	2.71 (1.09 - 5.59)	21	7	9	6	97
Grand Dakar	1	0.027	5.77 (1.16 – 16.86)	3	3	3	3	40
Ngor	0							
Ouakam	1	0.003	10.71 (2.15 – 31.31)	3	3	3	3	18
Patte d'oie	0							
Plateau Faidherbe	0							
Plateau Rebeuss	1	0.014	2.89 (1.06 - 6.31)	17	6	10	5	55
Gueule Tapée	0							

Table 1. Over-prevalence zones of infected mice.

For each sampling site included in this study (n=9), purely spatial cluster analysis was conducted to identify significant over-prevalence zones using Kulldorf's statistic.

sites	number of admixed ; non-admixed genotypes	number of DAPC groups		
		total	within the over-prevalence	among infected mice within the over-prevalence
Colobane	17 ; 27	2		
Goree Island	5 ; 52	5	3	3
Grand Dakar	0 ; 29	3	2	2
Ngor		No clustering		
Ouakam	11 ; 32	2	2	2
Patte d'oie	10 ; 26	2		
Plateau Faidherbe		No clustering		
Plateau Rebeuss	5 ; 49	3	2	2
Gueule Tapée	10 ; 33	2		

Table 2. Genetic clustering of mice populations using discriminant analysis of principal components (DAPC).

DAPC was performed for each sampling site included in this study (n=9).

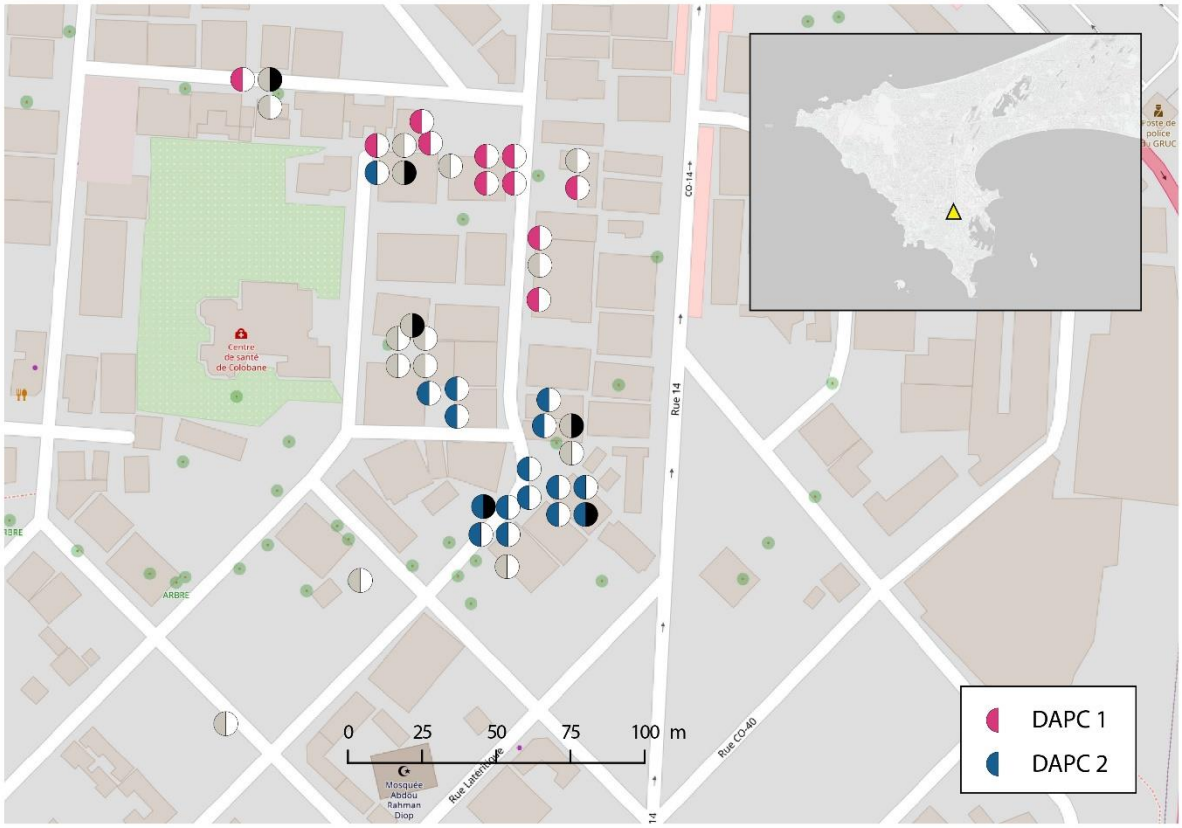
Centre de santé de Colobane
Rue 14
Rue CO-40
Mosquée Abdou Rahman Diop
0
25
50
75
100 m
DAPC 1
DAPC 2

Fig. S 1

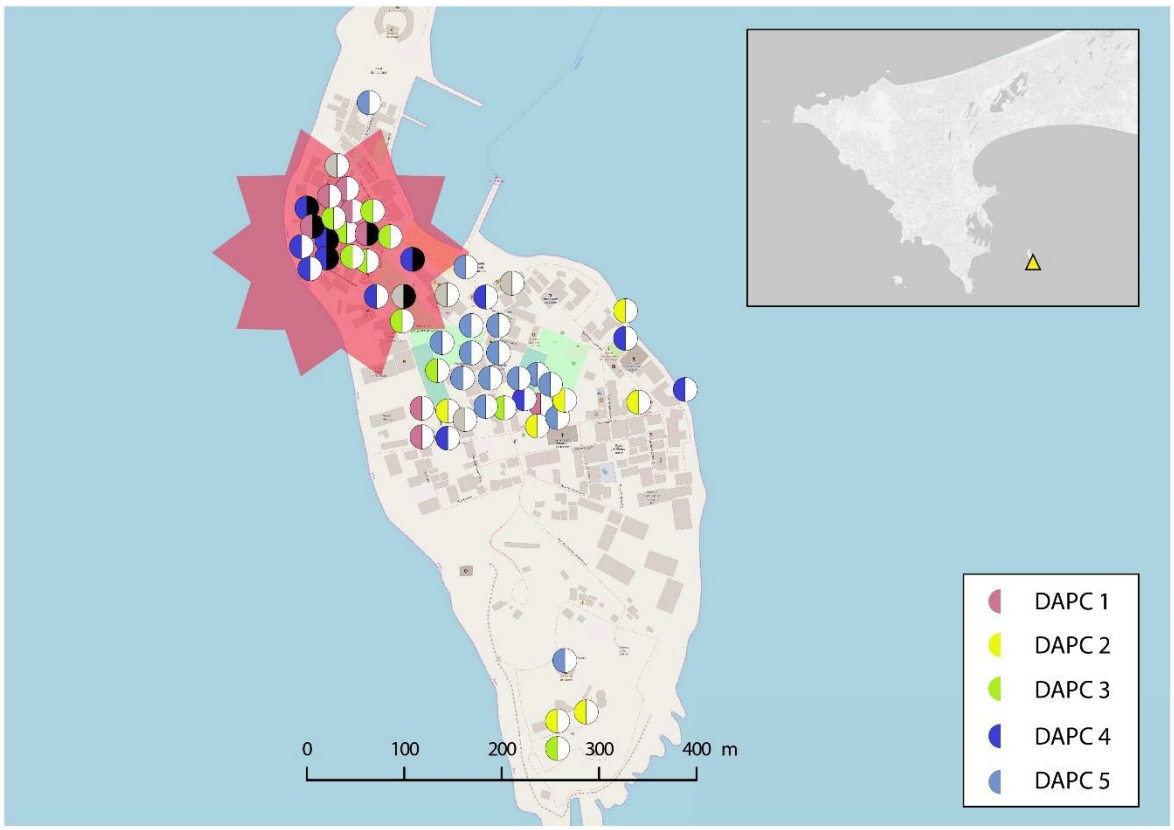

Fig. S 2

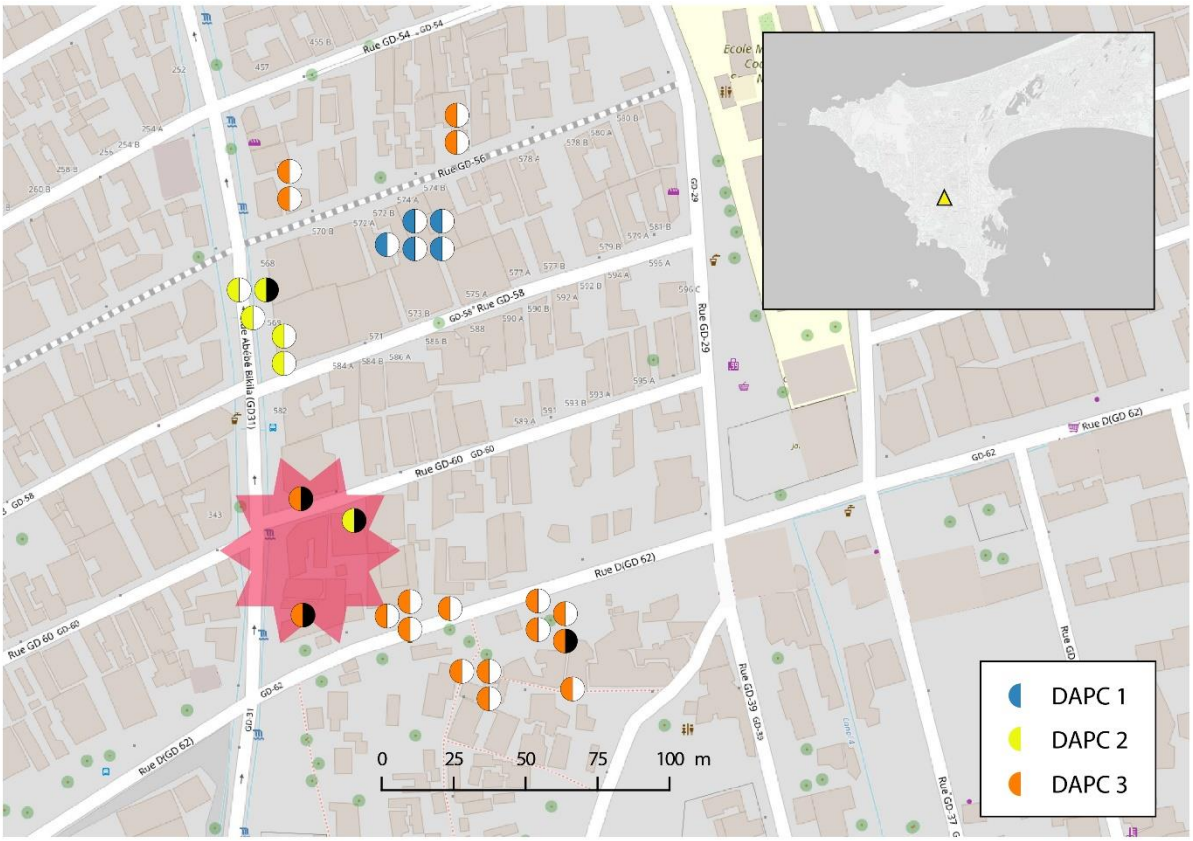
Rue GD-54
GD-54
Rue GD-56
GD-56
GD-58
Rue GD-58
Rue GD-60
GD-60
Rue D(GD 62)
GD-62
Rue GD-29
GD-29
Rue GD-39
GD-39
Rue D(GD 6z)
Rue GD 60
0
25
50
75
100 m
DAPC 1
DAPC 2
DAPC 3

Fig. S 3

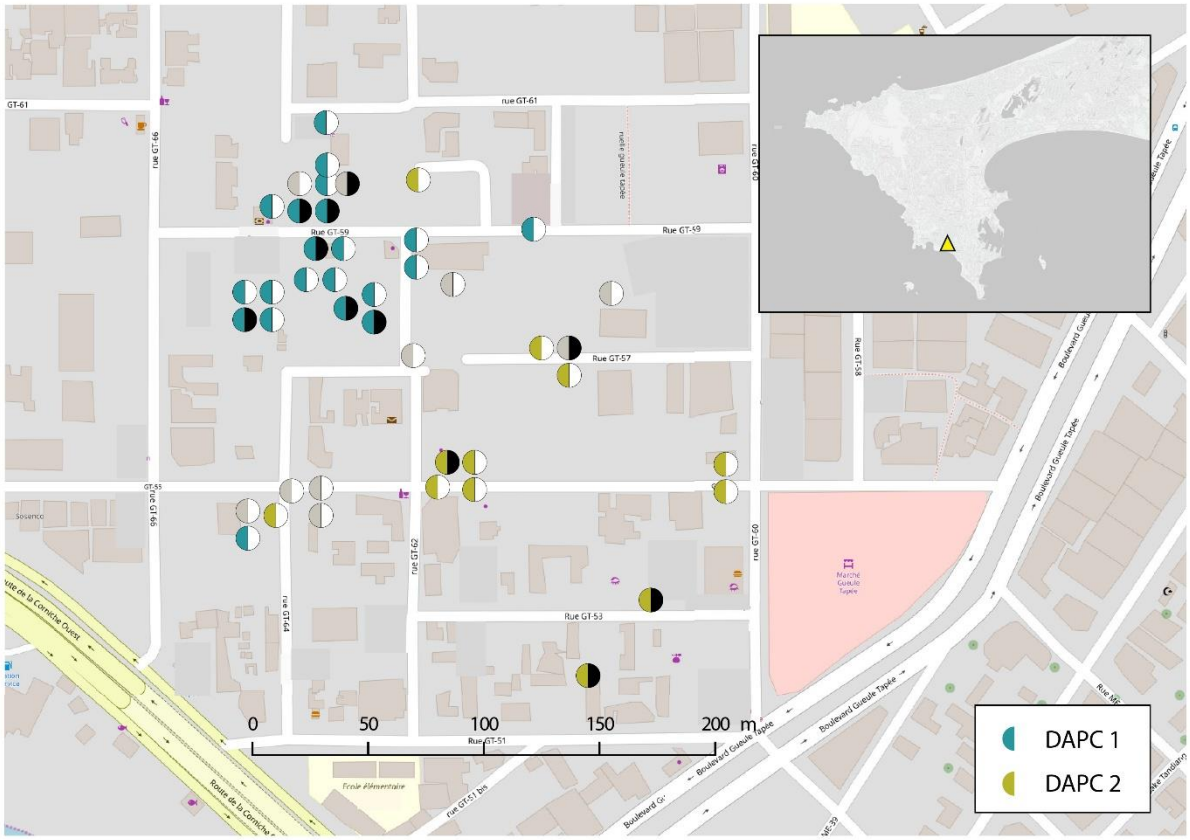
rue GT-61
Rue GT-59
Rue GT-49
Rue GT-57
Rue GT-53
Rue GT-51
Boulevard Gueule Tapée
0
50
100
150
200 m
DAPC 1
DAPC 2

Fig. S 4

Fig. S 5

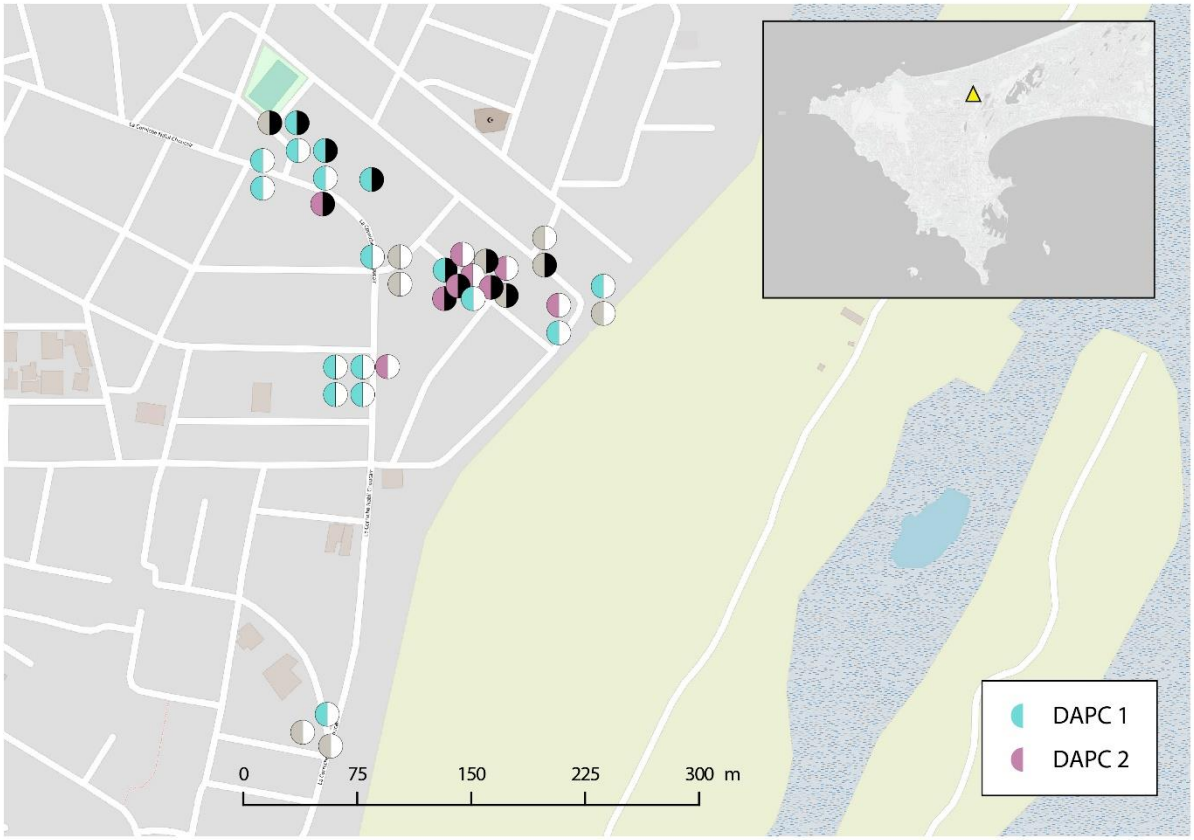

Fig. S 6

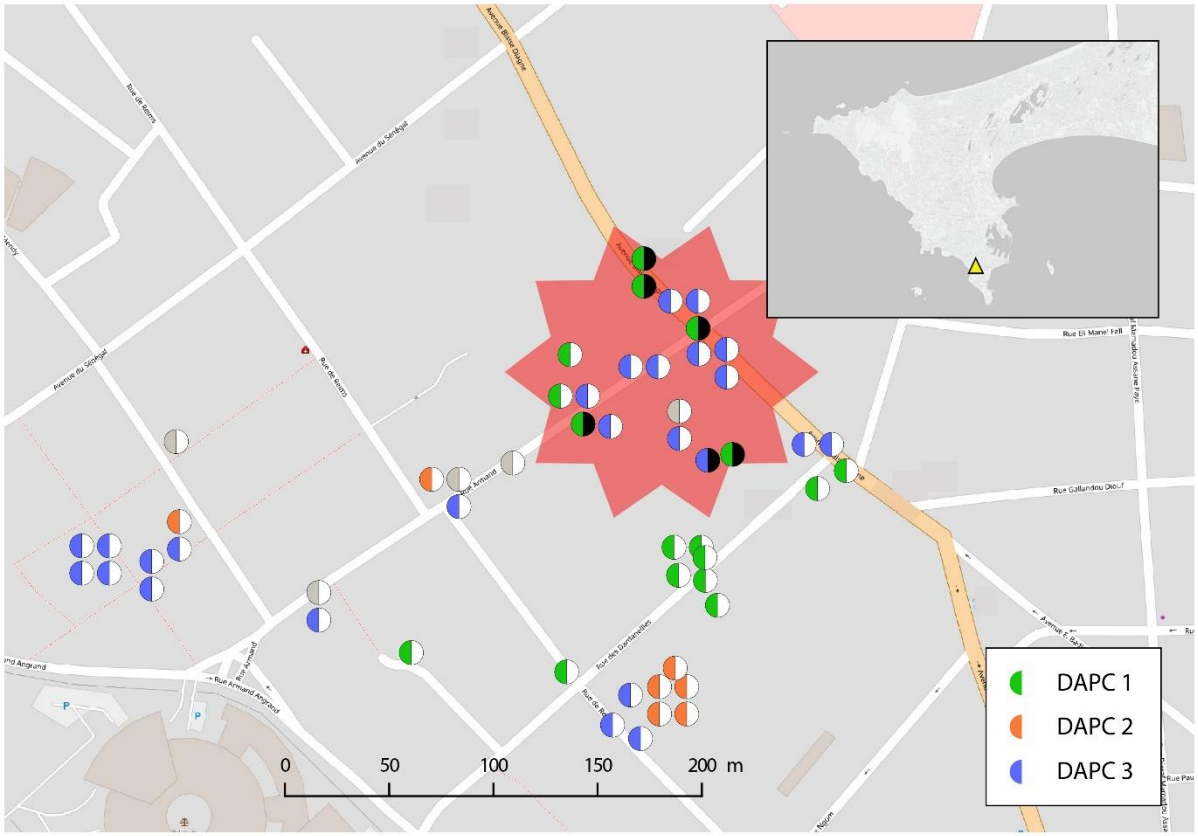

Fig. S 7

Fig. S1-S7. Mapping of mice from different sites of Dakar and on Goree Island according to DAPC cluster and *Toxoplasma gondii* infection.

Each circle represents an individual on its trapping localization. Different colours are attributed to distinct DAPC clusters in the left half of the circle. Mice with admixed genotypes are always represented in grey colour. The infected individuals and the non-infected individuals are represented in the right half of the circle in black and white, respectively. The over-prevalence zones are surrounded by a red circle with a dotted contour.

sites	sampled population	number of infected individuals	prevalence in % (95% CI)
Colobane	44	6	13.64 (3.50–23.78)
Goree Island	57	7	12.28 (3.76–20.80)
Grand Dakar	29	5	17.24 (3.49–30.99)
Ngor	29	4	13.79 (1.24–26.34)
Ouakam	43	3	6.98 (0.00–14.59)
Patte d'oie	36	12	33.33 (17.93–48.73)
Plateau Faidherbe	25	9	36.00 (17.18–54.82)
Plateau Rebeuss	54	6	11.11 (2.73–19.49)
Gueule Tapée	43	11	25.58 (12.54–38.62)

Table S1. Sampling size, number of infected individuals and prevalence data per site.

		group 1 DAPC1/DAPC1	group 2 DAPC1/DAPC2	group 3 DAPC1/DAPC3	group 4 DAPC1/DAPC4	group 5 DAPC1/DAPC5	group 6 DAPC2/DAPC2	group 7 DAPC2/DAPC3	group 8 DAPC2/DAPC4	group 9 DAPC2/DAPC5	group 10 DAPC3/DAPC3	group 11 DAPC3/DAPC4	group 12 DAPC3/DAPC5	group 13 DAPC4/DAPC4	group 14 DAPC4/DAPC5	group 15 DAPC5/DAPC5
group 1	DAPC1/DAPC1	*	582.036	241.145	797.624	618.659	-626.155	917.707	562.188	644.321	-521.593	685.320	883.815	-131.469	666.313	-305.484
group 2	DAPC1/DAPC2	< 0.0001	*	-340.891	215.588	36.624	-1208.190	335.671	-19.848	62.286	-1103.629	103.285	301.780	-713.505	84.277	-887.519
group 3	DAPC1/DAPC3	0.031	0.000	*	556.479	377.515	-867.299	676.563	321.043	403.177	-762.738	444.176	642.671	-372.614	425.168	-546.628
group 4	DAPC1/DAPC4	< 0.0001	0.009	< 0.0001	*	-178.965	-1423.779	120.083	-235.436	-153.302	-1319.217	-112.303	86.192	-929.093	-131.311	-1103.108
group 5	DAPC1/DAPC5	< 0.0001	0.668	< 0.0001	0.010	*	-1244.814	299.048	-56.471	25.662	-1140.252	66.661	265.156	-750.128	47.653	-924.143
group 6	DAPC2/DAPC2	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	*	1543.862	1188.343	1270.476	104.562	1311.475	1509.970	494.686	1292.467	320.671
group 7	DAPC2/DAPC3	< 0.0001	0.000	< 0.0001	0.116	0.000	< 0.0001	*	-355.519	-273.386	-1439.300	-232.387	-33.892	-1049.176	-251.394	-1223.191
group 8	DAPC2/DAPC4	< 0.0001	0.813	< 0.0001	0.001	0.431	< 0.0001	< 0.0001	*	82.133	-1083.781	123.132	321.627	-693.657	104.125	-867.672
group 9	DAPC2/DAPC5	< 0.0001	0.477	< 0.0001	0.034	0.735	< 0.0001	0.001	0.270	*	-1165.914	40.999	239.494	-775.790	21.991	-949.805
group 10	DAPC3/DAPC3	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.436	< 0.0001	< 0.0001	< 0.0001	*	1206.913	1405.408	390.124	1187.906	216.109
group 11	DAPC3/DAPC4	< 0.0001	0.194	< 0.0001	0.071	0.316	< 0.0001	0.002	0.057	0.554	< 0.0001	*	198.495	-816.790	-19.008	-990.804
group 12	DAPC3/DAPC5	< 0.0001	0.000	< 0.0001	0.189	0.000	< 0.0001	0.657	< 0.0001	0.001	< 0.0001	0.001	*	-1015.285	-217.503	-1189.299
group 13	DAPC4/DAPC4	0.224	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	*	797.782	-174.015
group 14	DAPC4/DAPC5	< 0.0001	0.278	< 0.0001	0.028	0.458	< 0.0001	0.000	0.095	0.743	< 0.0001	0.735	0.000	< 0.0001	*	-971.796
group 15	DAPC5/DAPC5	0.008	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.012	< 0.0001	< 0.0001	< 0.0001	0.028	< 0.0001	< 0.0001	0.029	< 0.0001	*

Table S2. Pairwise comparisons of the distribution of r values between groups of mice pairs from Goree Island defined according to their DAPC clusters of belonging. Pairwise comparisons between the mean ranks of the different groups and their p-values (in italics) using the Dunn's procedure (Bonferroni corrected significance level: 0.0005).

		group 1 DAPC1/DAPC1	group 2 DAPC1/DAPC2	group 3 DAPC1/DAPC3	group 4 DAPC2/DAPC2	group 5 DAPC2/DAPC3	group 6 DAPC3/DAPC3
group 1	DAPC1/DAPC1	*	520.718	466.945	-846.943	350.128	99.000
group 2	DAPC1/DAPC2	< 0.0001	*	-53.774	-1367.661	-170.591	-421.719
group 3	DAPC1/DAPC3	< 0.0001	0.240	*	-1313.887	-116.817	-367.945
group 4	DAPC2/DAPC2	< 0.0001	< 0.0001	< 0.0001	*	1197.070	945.942
group 5	DAPC2/DAPC3	< 0.0001	0.001	0.003	< 0.0001	*	-251.128
group 6	DAPC3/DAPC3	0.030	< 0.0001	< 0.0001	< 0.0001	< 0.0001	*

Table S3. Pairwise comparisons of the distribution of r values between groups of mice pairs from *Plateau Reubeuss* defined according to their DAPC clusters of belonging.

Pairwise comparisons between the mean ranks of the different groups and their p-values (in italics) using the Dunn's procedure (Bonferroni corrected significance level: 0.003).