Indication of spatially random infection of chlamydia-like organisms in *Bufo bufo* tadpoles from ponds located in the Geneva metropolitan area

Elia Vajana[1*], Ivo Widmer[1,2*], Estelle Rochat[1], Solange Duruz[1], Oliver Selmoni[1], Séverine Vuilleumier[3], Sébastien Aeby[4], Gilbert Greub[4†], Stéphane Joost[1,5,6#†]

[1] Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[2] Forum Biodiversity Switzerland (SCNAT), Berne, Switzerland

[3] La Source, School of Nursing, University of Applied Sciences and Arts Western Switzerland (HES-SO), Lausanne, Switzerland

[4] Institute of Microbiology, Lausanne University Hospital (CHUV) and University of Lausanne (UNIL), Lausanne, Switzerland

[5] Unit of Population Epidemiology, Division of Primary Care Medicine, Department of Community Medicine, Primary Care and Emergency Medicine, Geneva University Hospitals

[6] Institute of Social and Preventive Medicine (IUMSP), Division of chronic diseases (dMC), Lausanne University Hospital (CHUV), Lausanne, Switzerland

Running Head: Chlamydia-like organisms on *Bufo bufo*

Corresponding author: S. Joost, Ph.D, Laboratory of Geographic Information Systems (LASIG), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Email: stephane.joost@epfl.ch

*These authors contributed equally to this work and are considered co-first authors.

† Co-senior authors

Abstract

Occurrence of bacteria belonging to the order *Chlamydiales* was investigated for the first time in common toad (*Bufo bufo*) tadpole populations collected from 41 ponds in the Geneva metropolitan area, Switzerland. A *Chlamydiales*-specific Real-Time PCR was used to detect and amplify the *Chlamydiales* 16S rRNA-encoding gene from the tails of 375 tadpoles. We found the studied amphibian populations to be infected by "Chlamydia-like organisms" (CLOs) attributable to the genera *Similichlamydia*, *Neochlamydia*, *Protochlamydia* and *Parachlamydia* (belonging to the family *Parachlamydiaceae*), *Simkania* (family *Simkaniaceae*) and *Estrella* (family *Criblamydiaceae*); additionally, DNA from the genus *Thermoanaerobacter* (family *Thermoanaerobacteriaceae*) was detected. A global autocorrelation analysis did not reveal a spatial structure in the observed CLOs infection rates, and association tests involving land cover characteristics did not evidence any clear effect on CLOs infection rates in *B. bufo*. Despite preliminary, these results suggest a random and ubiquitous distribution of CLOs in the environment, which would support the biogeographical expectation "everything is everywhere" for the concerned microorganisms and their amoeba vectors.

Keywords

Bufo bufo, *Chlamydiales*, chlamydia-like organisms, chlamydia-related bacteria, emerging pathogens, intracellular bacteria, free-living amoebae, global spatial autocorrelation, hierarchical clustering, group comparison analysis, beta regression, public health, Geneva urban area.

Introduction

The order *Chlamydiales* consists of strict intracellular bacteria that replicate within eukaryotic cells of several animal hosts, among which humans [1]. Current molecular evidence suggests the

existence of two main lineages within *Chlamydiales*, which possibly diverged between 700 and 1400 million years ago from the last common ancestor [2]: the family *Chlamydiaceae*, and the "chlamydia-like organisms" (CLOs) belonging to the families *Piscichlamydiaceae*, *Clavichlamydiaceae*, *Simkaniaceae*, *Rhabdochlamydiaceae*, *Waddliaceae*, *Parachlamydiaceae*, *Criblamydiaceae*, and *Parilichlamydiaceae* [1,3,4].

Chlamydiaceae were first described in the sixties, and since then they were discovered to cause a wide variety of diseases affecting over 400 documented animal species [1,5,6]. Among the most eminent representatives of the family are (i) *Chlamydia trachomatis*, the etiological agent of the human visually-impairing trachoma [1], (ii) *C. psittaci*, responsible for pneumonia and hepatitis in birds [6], as well as zoonotic lung infections in humans [7] and equine infections [8], (iii) *C. abortus* causing abortion in sheep, goats, cattle and swine, and with proved transferability to humans [6,9], and (iv) *C. pneumoniae*, responsible for respiratory infections and atherosclerosis in humans, rhinitis in koalas and horses, and conjunctivitis in reptiles [1,3,6]. Notably, *C. pneumoniae* was also reported to infect amphibian populations of African clawed frogs (*Xenopus tropicalis*), great barred frogs (*Mixophyes iteratus*), blue mountains tree frogs (*Litoria citropa*), and common frogs (*Rana temporaria*), which were also positive for *C. abortus* and *C. suis* [10–12]; furthermore, the Candidatus *Amphibiichlamydia ranarum* was found at high prevalence in invasive bullfrog populations, and is considered an emerging pathogen possibly contributing to the current amphibian biodiversity crisis [13].

CLOs discovery is more recent, and dates back to the end of the eighties, when *Waddlia chondrophila* was first isolated from an aborted bovine foetus [1,3]. As exhaustively reviewed [3,6,14], most research has focused on identifying emerging pathogens among CLOs, with *W. chondrophila* being suspected to trigger human miscarriage [15] and ruminant abortion [16,17], *Parachlamydia acanthamoebae* human miscarriage [18], *Simkaniaceae*, *Parachlamydiaceae* and

69 *Rhabdochlamydiaceae* respiratory diseases in humans and cattle [19–22], ocular infections in cats [23]
70 and granulomatous inflammation in reptiles [24], *Parachlamydia* species to concur in the massive
71 mortality events affecting an highly endangered midwife toad (*Alytes obstetricans*) population [25],
72 and bacteria belonging to *Clavichlamydiaceae*, *Parachlamydiaceae*, *Parilichlamydiaceae*,
73 *Piscichlamydiaceae*, *Rhabdochlamydiaceae*, and *Simkaniaceae* having a recognized role in
74 epitheliocystis, a common gill disease in fish [3,26,27].

75 CLOs are also misnamed "environmental chlamydiae", as the majority of them were first
76 isolated from heterogeneous environmental sources spanning from water to soil [3,28,29]. Notably,
77 CLOs belonging to *Parachlamydiaceae*, *Simkaniaceae*, *Criblamydiaceae* and *Waddliaceae* have been
78 repeatedly observed as obligate endosymbionts of free-living amoebae species of the genera
79 *Acanthamoeba* and *Hartmanella* [5,10,29–31], which are therefore expected to play a central role in
80 guaranteeing CLOs survival and dispersal in the environment, as well as infections in new hosts [32].
81 Possibly reflecting the high ecological tolerance and dispersal capabilities of such vectors (which can
82 eventually rely on insects and wind to disperse over long distances), CLOs have been isolated from
83 several ecosystems, and are commonly considered ubiquitous in the environment [1,3,26,33,34], with a
84 highly diversified set of hosts including humans, marsupials and small mammals like the fruit bat,
85 reptiles like chelonians, lizards and snakes, fish species like the Leafy sea dragon, the Blue-striped
86 snapper, the Atlantic salmon, and the African catfish, as well as crustaceans like the Rough woodlouse
87 [1,3,9,24,26].

88 However, biogeographical end ecological studies have been conducted to elucidate possible
89 links between environmental conditions and composition of protozoa communities and distributions,
90 leading to reject the paradigm "everything is everywhere, but, the environment selects" associated with
91 free-living protozoa in some species (see the cases of *Nebela vas* and *Badhamia melanospora*) [35,36].

Particularly, local trends in precipitation [37] and soil characteristics related to moisture, temperature, pH, dissolved oxygen, and land cover (especially in terms of bryophyte species occurring in the crust) showed association with some testate [38–40] and protosteloid amoebae [37] occurrence. Such evidences would suggest – or at least do not exclude – a possible and still largely unexplored environmental influence on CLOs vectors and their endosymbionts at a local geographical scale.

In the present work, we investigated the occurrence of CLOs infection in the widespread common toad (*Bufo bufo*) for the first time, and tested the "everything is everywhere, but, the environment selects" principle with the observed infection patterns [41,42]. Particularly, we first tested *B. bufo* tadpole populations from the Geneva metropolitan area (Switzerland) for infection, and then the resulting infection rates for random distribution and association with land cover characteristics. In a public health perspective, we also derived human population density around sampling sites and studied a possible relationship with *B. bufo* infection, given the CLOs ability to infect humans from environment [3,21,43].

Materials and Methods

Sampling

In the context of the URBANGENE project, sampling locations were chosen in the state of Geneva on the basis of the MARVILLE (http://campus.hesge.ch/mareurbaine/) and of the *Centre de coordination pour la protection des amphibiens et des reptiles de Suisse* (http://www.karch.ch/) ponds databases, and also by means of a crowdsourcing campaign to include private ones (http://urbangene.heig-vd.ch). One hundred and fifty ponds were identified and then inspected. Tadpoles were finally sampled from April 9 to 22, 2015, in a subset of 41 ponds (Figure 1). Sampled ponds differed in size, species composition, and in the typology of the surrounding environments, some

of them being located in close proximity to the densely inhabited Geneva downtown (and placed in urban parks and private grounds), some others in the more rural Geneva suburbs, characterized by a higher degree of naturalness. Overall, 375 tadpoles were sampled, with an average of 9.2 samples per pond (sampling range: 4-15 tadpoles per pond). In order to characterize the whole tadpole population present in a pond, sampling privileged tadpoles coming from different frogspawns, whenever present; in such a case, tadpoles were collected shortly after they hatched from their frogspawn to reduce the chance of sampling siblings.

DNA extraction

Sample preparation

After sampling, tadpoles were put individually in a water dish with Tricaine methane sulphonate (MS-222), which caused tadpoles' rapid anaesthesia and decease. The apical part of the tail was carefully clipped in order to avoid contamination by bacteria from the intestinal tract. After the freeze-drying of the tails, DNA was extracted at the LGC laboratories in Berlin (Germany), using the sbeadex^{TM} tissue kit (LGC, Teddington, UK), and following the manufacturer's instructions.

Pan-Chlamydiales *real-time PCR assay*

A *Chlamydiales*-specific Real-Time PCR [44] was used to detect and amplify the DNA fragment from 207 to 215 bp belonging to the *Chlamydiales* 16S rRNA-encoding gene. Quantification was performed using a plasmidic 10-fold-diluted positive control tested in duplicate. Amplification reactions were performed in a final volume of 20 µl, containing: (i) iTaq Universal Probes Supermix with ROX (Bio-Rad, Reinach, Switzerland); (ii) 0.1 µM concentration of primers panCh16F2 (5′-CCGCCAACACTGGGACT-3′) (the underlined bases representing locked nucleic acids) and

panCh16R2 (5′-GGAGTTAGCCGGTGCTTCTTTACC-3′) (Eurogentec, Seraing, Belgium); (iii) 0.1 µM concentration of probe panCh16S (5′-FAM [6-carboxyfluorescein]-CTACGGGAGGCTGCAGTCGAGAATC-BHQ1 [black hole quencher 1]-3′) (Eurogentec); (iv) molecular-biology-grade water (Five Prime, Hilden, Germany); (v) 5 µl of sample DNA. Amplification started with an initial step of activation and denaturation at 95°C for 3 min, followed by 40 cycles at 95°C, 67°C and 72°C, each lasting 15 s, and was performed in a StepOne Plus real-time PCR system (Applied Biosystems, Zug, Switzerland). Samples with a threshold cycle value (C_T) <35 were finally sequenced, as this is the observed limit for amplicon sequencing (Aeby & Greub, unpublished).

DNA sequencing of the PCR-positive samples

According to the manufacturer's instructions, amplicons from positive samples were purified using the MSB Spin PCRapace (STRATEC Molecular, Berlin, Germany). The sequencing PCR assay was performed using a BigDye Terminator v1.1 cycle sequencing kit (Applied Biosystems, Zug, Switzerland), and with specific inner primers panFseq (5′-CCAACACTGGGACTGAGA-3′) and panRseq (5′-GCCGGTGCTTCTTTAC-3′). Amplification was performed after an initial denaturation step at 96°C for 1 min followed by 25 cycles at 96°C for 10 s and 60°C for 4 min. Purification of the sequencing PCR products was done using the SigmaSpin Sequencing Reaction Clean-up (Sigma-Aldrich, Buchs, Switzerland), and the sequencing was performed in a 3130xL genetic analyzer (Applied Biosystems). Sequences were analysed and blasted using the Geneious software [45,46].

Global spatial autocorrelation

Following taxonomic assignments, tadpoles' infection rates (IRs) were computed for each pond. A global spatial autocorrelation analysis was then conducted to investigate the presence of clusters or dissimilarities (i.e. the existence of spatial groups with similar IRs values, or, on the contrary, the

157 tendency of similar values to stay far away in space), under the null expectation of a random spatial
158 distribution of infection. Moran Scatter plots [47,48] were then constructed testing different weighting
159 criteria (in particular, using the mean IR from the first two, four, six, eight and ten nearest ponds,
160 respectively), and Moran's I [49] was estimated in each weighting scenario as the slope of the linear
161 regression between weighted and observed IRs. Both observed and weighted IRs were centred prior to
162 the analysis. Under the null hypothesis (h_0) of a Moran's I equal to zero and a significance threshold (α)
163 set to 0.05, statistical significance of observed Moran's I was derived by permuting IRs over the
164 landscape for 9999 times, re-estimating Moran's I in each permutation, deriving a Moran's I reference
165 distribution, and computing the pseudo *p*-value associated with the observed Moran's I in each
166 weighting scenario. Analysis was performed using a self-made script written in the R programming
167 language [50,51].

168 **Human population density around ponds**

169 The number of inhabitants residing within 1 km radius from each pond (surface: ~3.14 km^2)
170 was derived from the Federal Statistical Office database for the Republic and Canton of Geneva
171 (www.bfs.admin.ch; see Table 1), and for the year 2013. To test for the existence of a spatial
172 relationship between human and CLOs occurrence, the Pearson's product moment correlation
173 coefficient (r) was estimated between the number of inhabitants and the observed IRs, and a correlation
174 test (h_0: r=0; α=0.05) was performed through the function `cor.test` as implemented in the `stats` R
175 package [50]. Ponds 1, 32, 40 and 42 were discarded from analysis given partial information about the
176 number of inhabitants in the surrounding area.

177 **Group comparison analysis**

178 Thirty-two categories describing land cover were derived from the territorial information
179 system in Geneva (SITG) database (http://ge.ch/sitg/sitg_catalog/geodataid/1133) at a 10 m resolution.
180 Proportions of each land cover category were then computed around the sampling sites as a function of
181 a selected radius (i.e. buffer). In particular, buffers from 20 m up to 3 km (total: 299) were tested, and
182 the R function `extract` [52] was used to extrapolate the land cover categories from the buffer circles.

183 To group ponds with similar characteristics, hierarchical clustering was performed on the
184 resulting land cover proportions with the R function `hclust` [50], by relying on the "average",
185 "complete", "single", "Ward1" and "Ward2" clustering methods [53,54]. The Silhouette method [55]
186 was used to identify the optimal number of clusters, as well as the ponds' membership within the
187 groups.

188 When two groups of ponds were identified, independent t-tests or Wilcoxon rank sum tests
189 were run with the R functions `t.test` and `wilcox.test`, respectively; on the contrary, one-way
190 ANOVAs or Kruskal-Wallis rank sum tests were performed with `aov` or `kruskal.test` [50] in the
191 presence of more than two clusters. Test choice was driven by firstly checking IRs for normality and
192 homoscedasticity. *P*-values from the analyses performed with the same cluster method were corrected
193 for multiple testing with the "Benjamini-Hochberg" (BH) method through `p.adjust` [50]. Buffer
194 scenarios with at least one group composed by a single pond were discarded given impossibility of
195 assessing normality.

196 **Beta regression models**

197 Association between land cover and IRs was also investigated through univariate beta
198 regression models [56]. To ease interpretation of regression coefficients (β), previously obtained land

cover proportions were aggregated into five classes prior to analysis (Supplementary Table 1): "Managed vegetation", accounting for human-managed green areas; "Vegetation near water", referring to species communities well-adapted to live into/or in close proximity with water surfaces; "Forests", grouping forest species; "Urban environment", encompassing highly urbanized areas; and "Open fields", enclosing natural vegetation different from forests. Association tests (h_0: β=0; α=0.05) were then performed separately for each class, and for each buffer used. The function `betareg` [56] was used to perform the tests, and observed IRs were transformed prior to the analyses to account for the presence of extreme values (zeros and ones) [57]. As before, *p*-values from the tests involving the same aggregated land cover class were corrected for multiple testing with the BH method.

Results

We found 145 tadpoles (i.e. 38.7% of the samples) positive for the presence of *Chlamydiales* infection. Positive samples occurred across 36 sampling sites (i.e. 87.8% of the ponds sampled), with 5 ponds (12.2%) displaying no evidence of *Chlamydiales* occurrence, and infection rates spanning from 0 to 100% (Table 1 and Figure 2). Moran's I analysis indicated the absence of significant spatial autocorrelation in the observed infection rates, regardless of the weighting scenarios used (Figure 3). In addition, no significant relationship was observed between human presence and observed IRs (r=−0.145; 95% confidence intervals: −0.448, 0.187; t=−0.869; df=35; p-value=0.391), with ponds displaying IR>0.5 being located in both scarcely and highly populated areas (see ponds 39 and 23 with ~50 and 900 inhabitants/3.14 km^2, respectively, and ponds 31 and 17 with ~6000 and 7100 inhabitants/3.14 km^2, respectively; Supplementary Figure 1).

Only Wilcoxon and Kruskal-Wallis rank sum tests were performed on the groups of ponds defined with clustering analysis due to non-normality and/or heteroscedasticity. After multiple testing

correction, none of the land cover-based group showed any significant difference in IRs (Figure 5). Likewise, none of the aggregated land cover variable showed a significant association with IRs after multiple testing correction in the beta regression analysis (Figure 6).

Out of the 145 positive samples, 16 presented a C_T value <35, and were subsequently sequenced at the 16S ribosomal RNA gene for taxonomic identification. Taxonomic attribution was possible at a family-level lineage for 13 samples. In particular, six were found to be positive for *Parachlamydiaceae*, three for *Simkaniaceae* and two for *Criblamydiaceae*. The remaining two samples were found positive for the family *Thermoanaerobacteriaceae*, genus *Thermoanaerobacter*, which does not belong to *Chlamydiales*. Among the six samples infected by *Parachlamydiaceae*, one was positive for the genus *Similichlamydia* (as retrieved from Pond 35), one for the genus *Neochlamydia* (from pond 3), one for the genus *Protochlamydia* (from Pond 17), and three for the genus *Parachlamydia* (as observed in pond 30, with two sequences being highly similar with less than 1% divergence in the 16S rRNA sequence). All the samples infected by *Simkaniaceae* were assigned to the genus *Simkania* (from ponds 7, 15 and 38, respectively). Finally, samples infected by *Criblamydiaceae* and positive for *Thermoanaerobacteriaceae* were assigned to the genera *Estrella* and *Thermoanaerobacter*, respectively, and retrieved from Ponds 23 and 25 (Table 1 and Figure 4). Due to low sequencing quality, no lineage could be identified for the samples coming from Ponds 1, 34 and 36.

Discussion

The order *Chlamydiales* comprises bacterial agents of important human and animal diseases, as well as emerging pathogens which affect a broad spectrum of hosts [1,3,26]. To our knowledge, the

242 present study reports the first observation of CLOs infection in tadpoles' populations of the common
243 toad species *B. bufo*.

244 Notably, among the operational taxonomic units found are CLOs assigned to the genus
245 *Parachlamydia*, which characterized the microbiome of a Pyrenean midwife toad population unable to
246 recover from an infection by the highly aggressive fungus *Batrachochytrium dendrobatidis* [25].
247 Considering co-occurrence of Chlamydiae and *B. dendrobatidis* was also observed in a *X. tropicalis*
248 population undergoing epizootic disease dynamics [12], an association was proposed between the skin
249 microbiome of amphibians and *B. dendrobatidis* infection outcome [25]. Given the emerging role of *B.*
250 *dendrobatidis* in the current global amphibian biodiversity crisis [58], *Parachlamydia* occurrence might
251 pinpoint a potential vulnerability for the *B. bufo* populations under study which should deserve
252 attention for conservation.

253 The genera *Simkania* and *Neochlamydia* encompasses recognized emerging pathogens for both
254 humans and animals. Particularly, *Simkania* species were associated with respiratory deficit in humans
255 and epitheliocystis in fish, and *Neochlamydia* species with ocular diseases in domestic cat and
256 epitheliocystis [3]. So far, there is weak evidence about *Estrella* involvement as a human pathogen,
257 even if this recently discovered bacterial genus is still understudied [59]. Such findings would suggest a
258 potential role for *B. bufo* as a host reservoir for CLOs, and an ad hoc monitoring program might be
259 beneficial for both biodiversity conservation and public health purposes. Nevertheless, no evident
260 relationship seems to exist between observed infection rates and population density in the study area,
261 even if further studies would be advisable relying on a bigger sample size to obtain more robust
262 evidences in this regard; furthermore, particular consideration should be accorded to *Le Marais* and
263 *Étang Hutins* (Pond 17 and Pond 31, respectively) given their combination of high IRs and number of
264 inhabitants (Supplementary Figure 1).

Several studies support the paradigm of microbial biogeography "everything is everywhere, but, the environment selects" for *Chlamydiales* [1,3,42]. Coherently with such literature statements, the apparent absence of a clear spatial pattern, together with the rather ubiquitous occurrence of infection in *B. bufo*, would suggest a cosmopolitan distribution for the vectors and their endosymbionts even at a local geographical scale (see Figures 2 and 3), thus comforting the expectation "everything is everywhere". Nevertheless, our attempt to investigate how "the environment selects" was more tricky, and no association was found with land cover typologies able to explain the observed CLOs distribution (see Figures 5 and 6). At this regard, we believe further studies should focus on the influence of local environmental conditions on *Chlamydiales* occurrence, especially tacking into consideration alternative variables with proved effects on amoebae distributions (e.g. precipitation, temperature, moisture, pH and dissolved oxygen) [36–40]. Ideally, the description of CLOs-specific niches would provide fine-scale predictive distribution maps of *Chlamydiales* and their vectors, with straightforward applications in both public health preventive strategies, and prioritization of susceptible animal host populations for conservation.

To conclude, the present work candidates the amphibian species *B. bufo* as a new host for CLOs, provides a first estimate of CLOs infection rate in *B. bufo* tadpole populations from a urban environment, agrees with literature findings concerned with *Chlamydiales* ubiquitous distribution, while apparently excluding the land cover as a selective variable for *Chlamydiales* occurrence in the studied area.

Acknowledgments

We warmly thank the residents of the city of Geneva who allowed us to access their private pond for sampling. The study was funded by the GELBERT Foundation in Geneva, Switzerland

(URBANGENE project no 088-2013), by the Grand Genève cross-border metropolitan area, and by the Direction générale Agriculture & Nature (DGAN) du Département de l'Environnement, des Transports et de l'Agriculture (DETA) of the State of Geneva, Switzerland.

References

[1] Horn M. Chlamydiae as Symbionts in Eukaryotes. Annu Rev Microbiol 2008;62:113–31. doi:10.1146/annurev.micro.62.081307.162818.

[2] Greub G, Raoult D. History of the ADP/ATP-Translocase-Encoding Gene, a Parasitism Gene Transferred from a Chlamydiales Ancestor to Plants 1 Billion Years Ago. Appl Environ Microbiol 2003;69:5530–5. doi:10.1128/AEM.69.9.5530-5535.2003.

[3] Taylor-Brown A, Vaughan L, Greub G, Timms P, Polkinghorne A. Twenty years of research into Chlamydia-like organisms: a revolution in our understanding of the biology and pathogenicity of members of the phylum *Chlamydiae*. Pathog Dis 2015;73:1–15. doi:10.1093/femspd/ftu009.

[4] Pillonel T, Bertelli C, Salamin N, Greub G. Taxogenomics of the order *Chlamydiales*. Int J Syst Evol Microbiol 2015;65:1381–93. doi:10.1099/ijs.0.000090.

[5] Lienard J, Croxatto A, Prod'hom G, Greub G. *Estrella lausannensis*, a new star in the *Chlamydiales* order. Microbes Infect 2011;13:1232–41. doi:10.1016/j.micinf.2011.07.003.

[6] Borel N, Polkinghorne A, Pospischil A. A Review on Chlamydial Diseases in Animals: Still a Challenge for Pathologists? Vet Pathol 2018;55:374–90. doi:10.1177/0300985817751218.

[7] Lamoth F, Greub G. Fastidious intracellular bacteria as causal agents of community-acquired pneumonia. Expert Rev Anti Infect Ther 2010;8:775–90.

[8] Polkinghorne A, Greub G. A new equine and zoonotic threat emerges from an old avian pathogen, *Chlamydia psittaci*. Clin Microbiol Infect 2017;23:693–4.

[9] Longbottom D, Coulter LJ. Animal Chlamydioses and Zoonotic Implications. J Comp Pathol 2003;128:217–44. doi:10.1053/jcpa.2002.0629.

[10] Horn M, Wagner M. Bacterial Endosymbionts of Free-living Amoebae. J Eukaryot Microbiol 2004;51:509–514.

[11] Blumer C, Zimmermann DR, Weilenmann R, Vaughan L, Pospischil A. Chlamydiae in Free-Ranging and Captive Frogs in Switzerland. Vet Pathol 2007;44:144–50. doi:10.1354/vp.44-2-144.

[12] Reed K. Chlamydia pneumoniae Infection in a Breeding Colony of African Clawed Frogs (*Xenopus tropicalis*). Emerg Infect Dis 2000;6:196–9. doi:10.3201/eid0602.000216.

[13] Martel A, Adriaensen C, Sharifian-Fard M, Vandewoestyne M, Deforce D, Favoreel H, et al. The novel Candidatus *Amphibiichlamydia ranarum*' is highly prevalent in invasive exotic bullfrogs (*Lithobates catesbeianus*): 'Candidatus Amphibiichlamydia ranarum' in bullfrogs. Environ Microbiol Rep 2013;5:105–8. doi:10.1111/j.1758-2229.2012.00359.x.

[14] Greub G. *Parachlamydia acanthamoebae*, an emerging agent of pneumonia. Clin Microbiol Infect 2009;15:18–28. doi:10.1111/j.1469-0691.2008.02633.x.

[15] Baud D, Goy G, Osterheld M-C, Croxatto A, Borel N, Vial Y, et al. Role of *Waddlia chondrophila* Placental Infection in Miscarriage. Emerg Infect Dis 2014;20:460–4. doi:10.3201/eid2003.131019.

[16] Blumer S, Greub G, Waldvogel A, Hässig M, Thoma R, Tschuor A, et al. *Waddlia*, *Parachlamydia* and *Chlamydiaceae* in bovine abortion. Vet Microbiol 2011;152:385–93. doi:10.1016/j.vetmic.2011.05.024.

[17] Henning K. *Neospora caninum* and *Waddlia chondrophila* strain 2032/99 in a septic stillborn calf. Vet Microbiol 2002;85:285–92. doi:10.1016/S0378-1135(01)00510-7.

[18] Ammerdorffer A, Stojanov M, Greub G, Baud D. Chlamydia trachomatis and chlamydia-like bacteria: new enemies of human pregnancies. Curr Opin Infect Dis 2017;30:289–96.

[19] Wheelhouse N, Longbottom D, Willoughby K. Chlamydia in cases of cattle pneumonia in Scotland. Vet Rec 2013;172:110–110.

[20] Friedman MG. Detection of *Simkania negevensis* by culture, PCR, and serology in respiratory tract infection in Cornwall, UK. J Clin Pathol 2006;59:331–3. doi:10.1136/jcp.2004.025601.

[21] Lamoth F, Jaton K, Vaudaux B, Greub G. *Parachlamydia* and *Rhabdochlamydia*: Emerging Agents of Community-Acquired Respiratory Infections in Children. Clin Infect Dis 2011;53:500–1. doi:10.1093/cid/cir420.

[22] Niemi S, Greub G, Puolakkainen M. Chlamydia-related bacteria in respiratory samples in Finland. Microbes Infect 2011;13:824–7. doi:10.1016/j.micinf.2011.04.012.

[23] von Bomhard W, Polkinghorne A, Lu Z, Vaughan L, Vögtlin A, Zimmermann D, et al. Detection of novel chlamydiae in cats with ocular disease. Am J Vet Res 2003.

[24] Soldati G, Lu ZH, Vaughan L, Polkinghorne A, Zimmermann DR, Huder JB, et al. Detection of Mycobacteria and Chlamydiae in Granulomatous Inflammation of Reptiles: A Retrospective Study. Vet Pathol 2004;41:388–97. doi:10.1354/vp.41-4-388.

[25] Bates KA, Clare FC, O'Hanlon S, Bosch J, Brookes L, Hopkins K, et al. Amphibian chytridiomycosis outbreak dynamics are linked with host skin bacterial community structure. Nat Commun 2018;9. doi:10.1038/s41467-018-02967-w.

[26] Borel N, Polkinghorne A, Pospischil A. A Review on Chlamydial Diseases in Animals: Still a Challenge for Pathologists? Vet Pathol 2018;55:374–90. doi:10.1177/0300985817751218.

[27] Taylor-Brown A, Pillonel T, Bridle A, Qi W, Bachmann NL, Miller TL, et al. Culture-independent genomics of a novel chlamydial pathogen of fish provides new insight into host-specific adaptations utilized by these intracellular bacteria: Novel chlamydial epitheliocystis agent in grouper. Environ Microbiol 2017;19:1899–913. doi:10.1111/1462-2920.13694.

[28] Corsaro D, Feroldi V, Saucedo G, Ribas F, Loret J-F, Greub G. Novel Chlamydiales strains isolated from a water treatment plant. Environ Microbiol 2009;11:188–200. doi:10.1111/j.1462-2920.2008.01752.x.

[29] Schmitz-Esser S, Toenshoff ER, Haider S, Heinz E, Hoenninger VM, Wagner M, et al. Diversity of Bacterial Endosymbionts of Environmental Acanthamoeba Isolates. Appl Environ Microbiol 2008;74:5822–31. doi:10.1128/AEM.01093-08.

[30] Michel R, Müller K-D, Hoffmann R. Enlarged Chlamydia-like organisms as spontaneous infection of *Acanthamoeba castellanii*. Parasitol Res 2001;87:248–51. doi:10.1007/s004360000318.

[31] Thomas V, Casson N, Greub G. *Criblamydia sequanensis*, a new intracellular *Chlamydiales* isolated from Seine river water using amoebal co-culture. Environ Microbiol 2006;8:2125–35. doi:10.1111/j.1462-2920.2006.01094.x.

[32] Barker J, Brown MRW. Trojan Horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. Microbiology 1994;140:1253–9. doi:10.1099/00221287-140-6-1253.

[33] Pilloux L, Aeby S, Gaümann R, Burri C, Beuret C, Greub G. The High Prevalence and Diversity of Chlamydiales DNA within Ixodes ricinus Ticks Suggest a Role for Ticks as Reservoirs and Vectors of Chlamydia-Related Bacteria. Appl Environ Microbiol 2015;81:8177–82. doi:10.1128/AEM.02183-15.

[34] Glime JM. Protozoa Ecology. Chapt. 2-6. In: Glime, J. M. Bryophyte Ecology. Bryological Interaction. vol. 2. n.d.

[35] Smith HG, Wilkinson DM. Not all free-living microorganisms have cosmopolitan distributions - the case of *Nebela* (*Apodera*) *vas* Certes (Protozoa: Amoebozoa: Arcellinida). J Biogeogr 2007;34:1822–31. doi:10.1111/j.1365-2699.2007.01733.x.

[36] Aguilar M, Fiore-Donno A-M, Lado C, Cavalier-Smith T. Using environmental niche models to test the 'everything is everywhere' hypothesis for *Badhamia*. ISME J 2014;8:737–45. doi:10.1038/ismej.2013.183.

[37] Aguilar M, Lado C. Ecological niche models reveal the importance of climate variability for the biogeography of protosteloid amoebae. ISME J 2012;6:1506–14. doi:10.1038/ismej.2012.12.

[38] Mieczan T, Adamczuk M. Ecology of testate amoebae (Protists) in mosses: distribution and relation of species assemblages with environmental parameters (King George Island, Antarctica). Polar Biol 2015;38:221–30. doi:10.1007/s00300-014-1580-0.

[39] Bobrov AA, Charman DJ, Warner BG. Ecology of Testate Amoebae (Protozoa: Rhizopoda) on Peatlands in Western Russia with Special Attention to Niche Separation in Closely Related Taxa. Protist 1999;150:125–36. doi:10.1016/S1434-4610(99)70016-7.

[40] Barry G. Warner, Taro Asada, Noel P. Quinn. Seasonal Influences on the Ecology of Testate Amoebae (Protozoa) in a Small *Sphagnum* Peatland in Southern Ontario, Canada. Microb Ecol 2007;54:91–100.

[41] Garcia-Porta J, Litvinchuk SN, Crochet PA, Romano A, Geniez PH, Lo-Valvo M, et al. Molecular phylogenetics and historical biogeography of the west-palearctic common toads (*Bufo bufo* species complex). Mol Phylogenet Evol 2012;63:113–30. doi:10.1016/j.ympev.2011.12.019.

[42] de Wit R, Bouvier T. "Everything is everywhere, but, the environment selects"; what did Baas Becking and Beijerinck really say? Environ Microbiol 2006;8:755–8. doi:10.1111/j.1462-2920.2006.01017.x.

[43] Greub G, Boyadjiev I, Scola B, Raoult D, Martin C. Serological Hint Suggesting That *Parachlamydiaceae* Are Agents of Pneumonia in Polytraumatized Intensive Care Patients. Ann N Y Acad Sci 2003;990:311–9. doi:10.1111/j.1749-6632.2003.tb07381.x.

[44] Lienard J, Croxatto A, Aeby S, Jaton K, Posfay-Barbe K, Gervaix A, et al. Development of a New Chlamydiales-Specific Real-Time PCR and Its Application to Respiratory Clinical Samples. J Clin Microbiol 2011;49:2637–42. doi:10.1128/JCM.00114-11.

[45] Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 2012;28:1647–9. doi:10.1093/bioinformatics/bts199.

[46] Widmer I, Rochat E, Duruz S, Aeby S, Selmoni O, Vajana E, et al. Chlamydia-like organisms in *Bufo bufo* tadpoles from ponds located in the Geneva metropolitan area (10.5281/zenodo.1404188). 2018.

[47] Anselin L. The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association. In Spatial Analytical Perspectives on GIS in Environmental and Socio-Economic Sciences. London: Taylor; Francis: Manfred Fischer, Henk Scholten, and David Unwin; 1996.

[48] Cliff A, Ord JK. Spatial Autocorrelation. London: Pion; 1973.

[49] Moran P. The Interpretation of Statistical Maps. J R Stat Soc Ser B Methodol 1948;10.

[50] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/ 2015.

[51] Vajana E, Widmer I, Rochat E, Duruz S, Selmoni O, Vuilleumier S, et al. Chlamydia-like organisms in *Bufo bufo* tadpoles from ponds located in the Geneva metropolitan area: R scripts (10.5281/zenodo.1404197). 2018.

[52] Hijmans RJ. raster: Geographic Data Analysis and Modeling. 2016.

[53] Legendre P, Legendre L. Numerical Ecology. 3rd English Ed. Amsterdam: Elsevier Science BV; 2012.

[54] Murtagh F, Legendre P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? J Classif 2014;31:274–95. doi:10.1007/s00357-014-9161-z.

[55] Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20:53–65. doi:10.1016/0377-0427(87)90125-7.

[56] Cribari-Neto F, Zeileis A. Beta Regression in R. J Stat Softw 2010;34.

[57] Smithson M, Verkuilen J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. Psychol Methods 2006;11:54–71. doi:10.1037/1082-989X.11.1.54.

[58] Olson DH, Aanensen DM, Ronnenberg KL, Powell CI, Walker SF, Bielby J, et al. Mapping the Global Emergence of *Batrachochytrium dendrobatidis*, the Amphibian Chytrid Fungus. PLoS ONE 2013;8:e56802. doi:10.1371/journal.pone.0056802.

436 [59] de Barsy M, Bottinelli L, Greub G. Antibiotic susceptibility of *Estrella lausannensis*, a potential
437 emerging pathogen. Microbes Infect 2014;16:746–54. doi:10.1016/j.micinf.2014.08.003.

438

Tables

Table 1. Information is reported for each pond including: experimental and actual name (Pond ID and Pond name, respectively), corresponding municipality, geographical coordinates, number of inhabitants (Nr. inh.s) within 1 km radius, number of *B. bufo* tadpoles sampled (N), number of CLOs-positive samples (P), observed infection rates (IR), and *Chlamydiales* taxonomic assignments at the genus-level.

Pond ID	Pond name	Municipality	Lon.	Lat.	Nr. inh.s	N	P	IR	Genus
1	Étang de l'ancien château de St-Victor	Avully	5.98	46.17	394[*]	10	6 (1)[a]	0.600	?[b]
2	Étang de la pépinière Jacquet	Satigny	6.05	46.22	107	13	4	0.308	
3	Étang du signal de Bernex	Bernex	6.07	46.17	2038	12	3 (1)	0.250	*Neochlamydia*
4	Grand Etang des Mouilles	Bernex	6.08	46.19	149	10	4	0.400	
5	Étang des Evaux	Confignon	6.09	46.19	1441	10	9	0.900	
6	Étang de la ferme Lignon (PAF)	Vernier	6.09	46.21	2725	12	2	0.167	
7	Étang Autrichien	Onex	6.11	46.18	4750	10	2 (1)	0.200	*Simkania*
8	Bassin du Parc Louis Bertrand	Lancy	6.12	46.19	6340	4	0	0.000	
9	Étang Zimmermann	Lancy	6.12	46.19	5740	5	0	0.000	
10	Étang Perfetta	Lancy	6.12	46.19	5829	4	1	0.250	
11	Étang Lescaze	Genève	6.12	46.2	6393	10	3	0.300	
12	Bassin du cimetière de St Georges	Genève	6.12	46.2	5963	9	2	0.222	
13	Étang Spring	Lancy	6.12	46.19	5048	10	0	0.000	
14	Bassin du Parc Chuit	Lancy	6.12	46.19	4609	10	4	0.400	
15	Étang du Pré-d'œufeuf	Plan-les-Ouates	6.14	46.16	1574	10	6 (1)	0.600	*Simkania*
16	Étang Pinchat	Carouge	6.15	46.18	7000	15	2	0.133	
17	Le Marais	Grand-Saconnex	6.11	46.23	7116	15	10 (1)	0.667	*Protochlamydia*
18	Étang du chemin des Prjins	Grand-Saconnex	6.12	46.23	3635	7	4	0.571	

19	Bassin du château de Penthes	Pregny-Chambésy	6.14	46.23	467	5	0	0.000	
20	Étang Berthier	Pregny-Chambésy	6.15	46.23	380	7	0	0.000	
21	Étang du Jardin Botanique 2 (Serres)	Genève	6.15	46.23	2137	10	6	0.600	
22	Étang du jardin botanique	Genève	6.15	46.23	2336	9	2	0.222	
23	Étang Vieux-Clos	Chêne-Bougeries	6.18	46.18	911	12	10 (2)	0.833	*Estrella*; *Thermoanaerobacter*
24	Étang Flory	Chêne-Bougeries	6.18	46.19	946	7	1	0.143	
25	Bassin Waldvogel	Chêne-Bougeries	6.18	46.19	2560	6	5 (2)	0.833	*Estrella*; *Thermoanaerobacter*
26	Bassin de la Station de Zoologie	Chêne-Bougeries	6.18	46.19	2641	11	2	0.182	
27	Étang Richard	Chêne-Bougeries	6.18	46.19	1345	5	5	1.000	
28	Étang Paradis Haake	Chêne-Bougeries	6.18	46.2	4661	8	2	0.250	
29	Étang route de Chêne	Chêne-Bougeries	6.19	46.2	4342	8	1	0.125	
30	Étang Broud	Chêne-Bougeries	6.19	46.2	3161	10	7 (3)	0.700	*Parachlamydia*
31	Étang Hutins	Chêne-Bourg	6.2	46.2	6043	8	7	0.875	
32	Étang Loutan	Thônex	6.2	46.2	6349*	6	1	0.167	
33	Étang de la clinique Bel-Air	Thônex	6.21	46.21	1543	9	1	0.111	
34	Étang de Miolan	Choulex	6.21	46.23	318	8	2 (1)	0.250	?
35	Bassin du Centre horticole Lullier 2	Jussy	6.25	46.23	216	8	4 (1)	0.500	*Similichlamydia*
36	Étang des Dolliets	Jussy	6.28	46.23	23	12	6 (1)	0.500	?
37	Bois du Faisan amont	Versoix	6.15	46.28	436	10	1	0.100	
38	Étang Bon-séjour	Versoix	6.16	46.28	3396	10	5 (1)	0.500	*Simkania*
39	Étang des Douves	Versoix	6.14	46.29	51	12	7	0.583	
40	Étang Est Pré-Béroux	Versoix	6.14	46.3	3*	6	4	0.667	
41	Étang de Combes-Chapuis	Versoix	6.12	46.3	13*	12	4	0.333	

444 [a]In brackets are the number of samples (among the positive ones) for which a sequence attributable to *Chlamydiales* was obtained.
445 [b]Question marks highlight unsolved taxonomic assignations in the samples for which a sequence attributable to *Chlamydiales* was

obtained. *The reported number of inhabitants can be partial when the pond is located either close to the French or to the Canton of Vaud border (see Figure 1), for which demographic information was not considered.

Figures

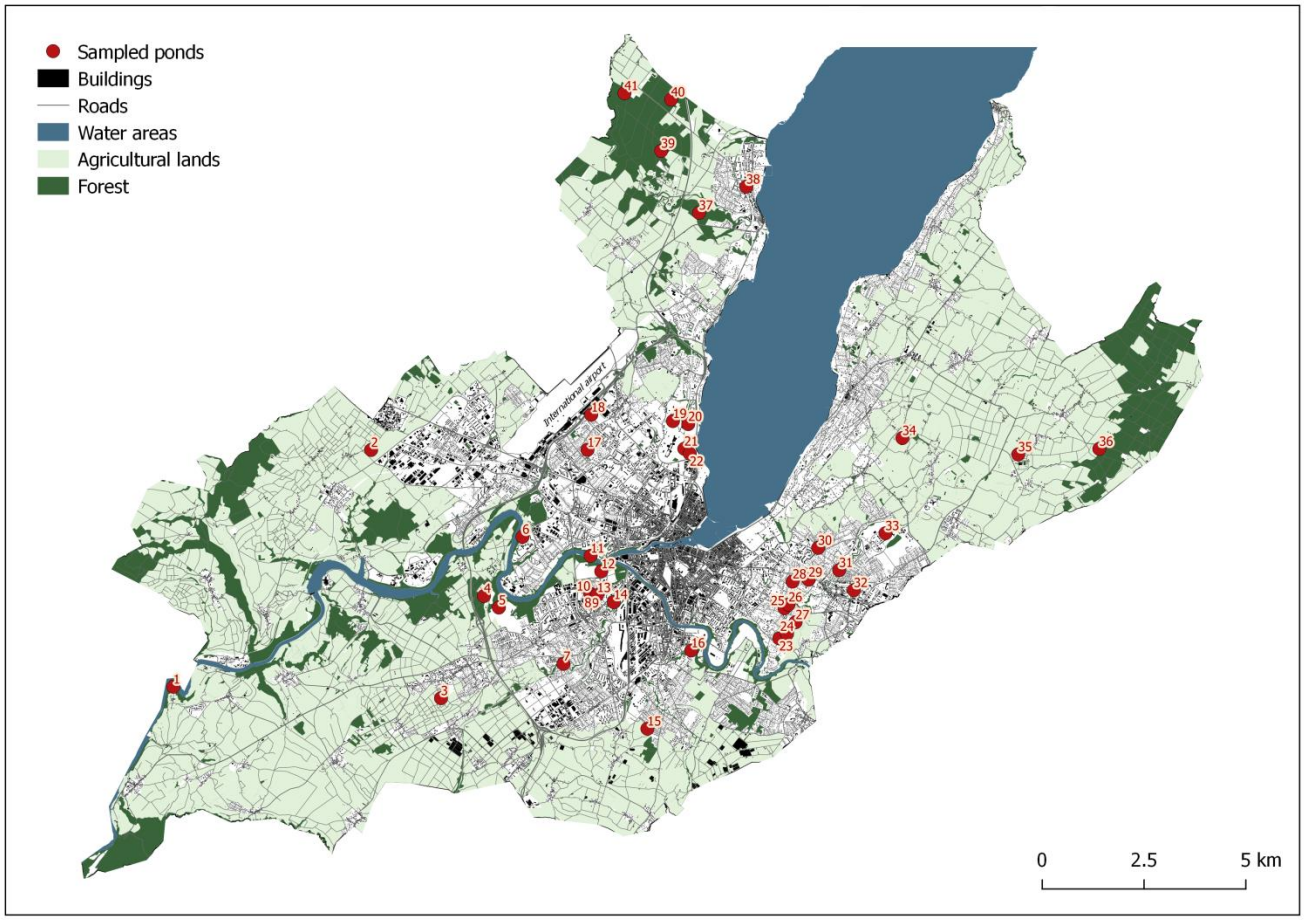


Figure 1. Spatial representation of the sampled ponds. Ponds are represented in red, with numbers corresponding to Pond IDs in Table 1. Background contextual information represents Geneva metropolitan area, and highlights forests, agricultural, water, as well as urbanized areas.

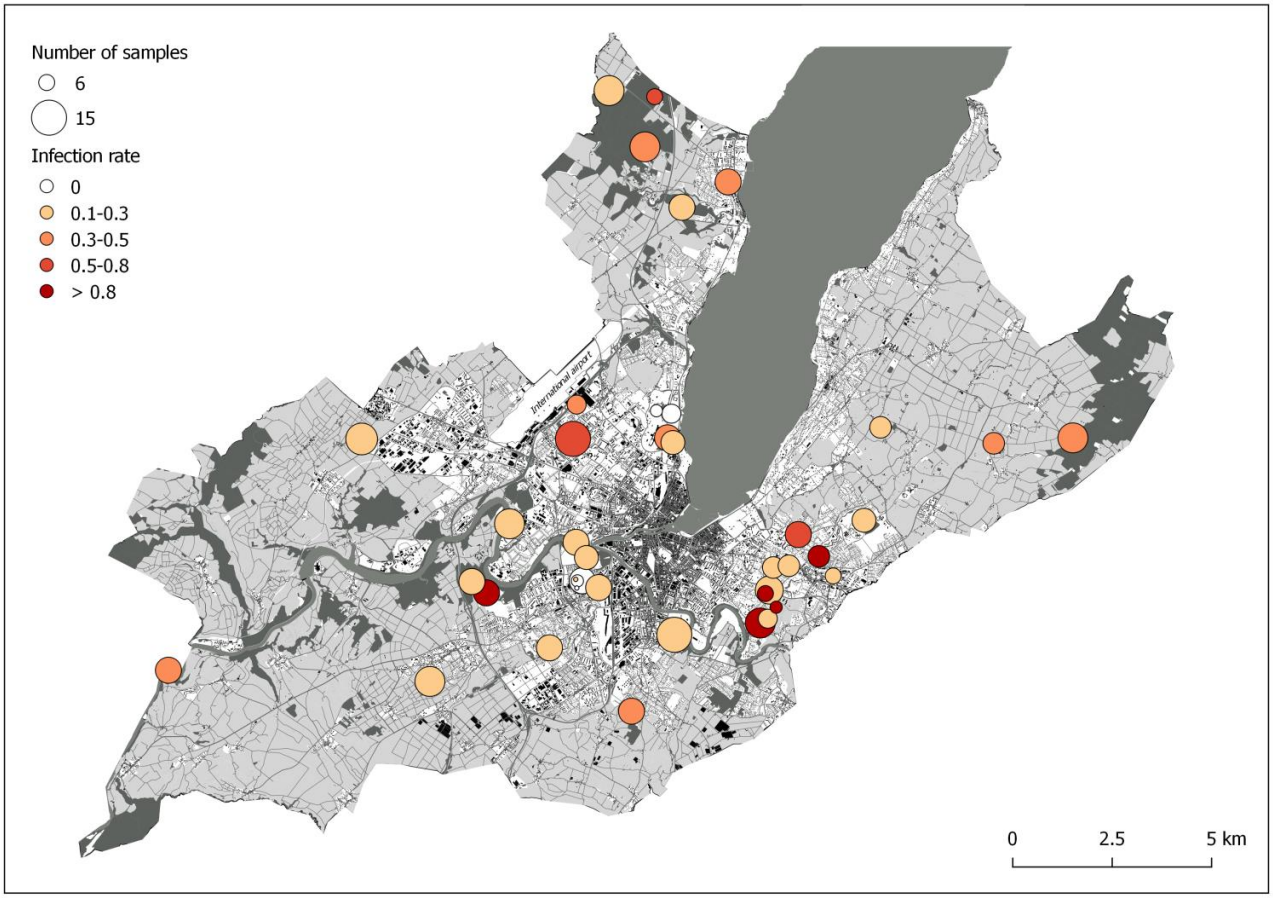
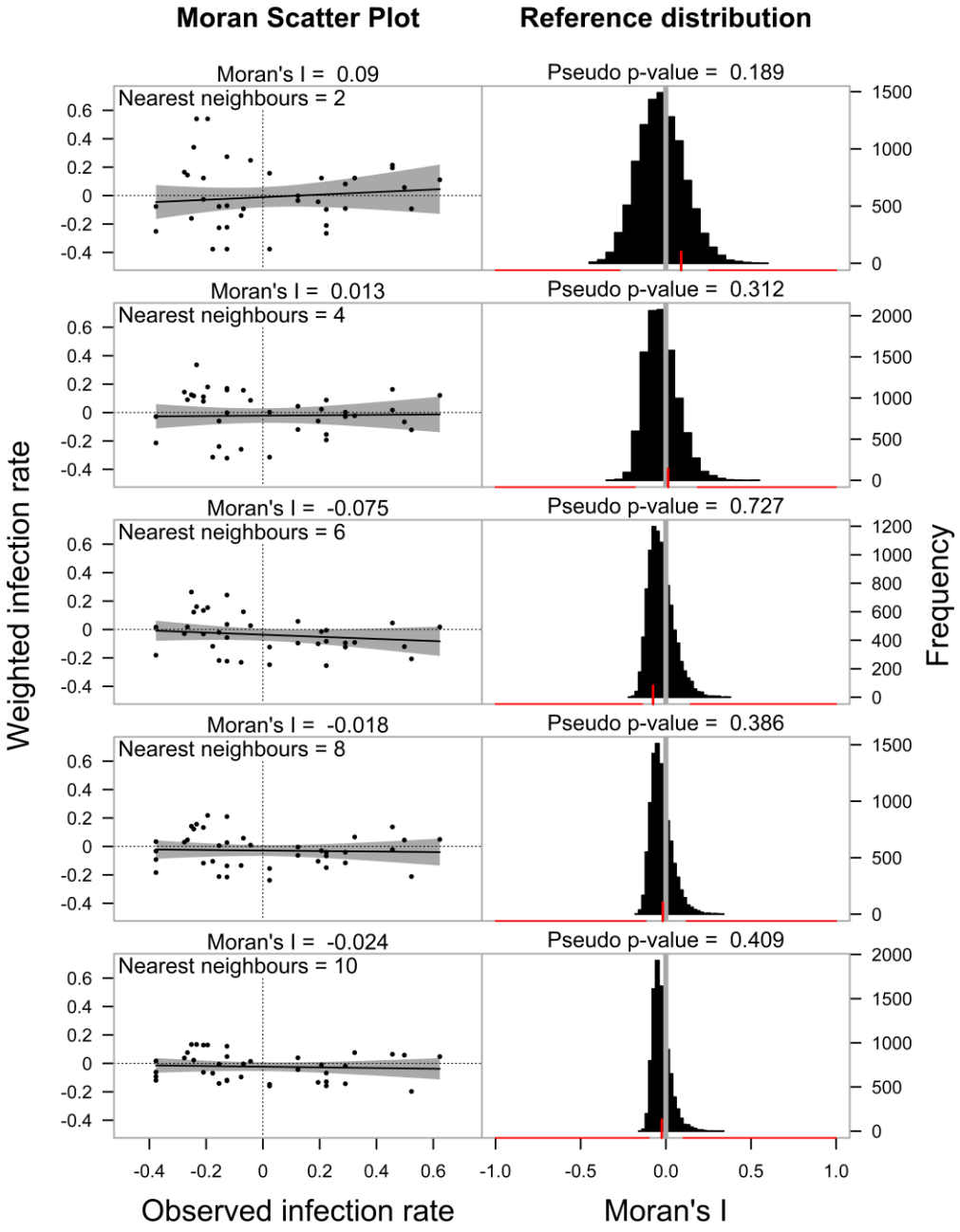
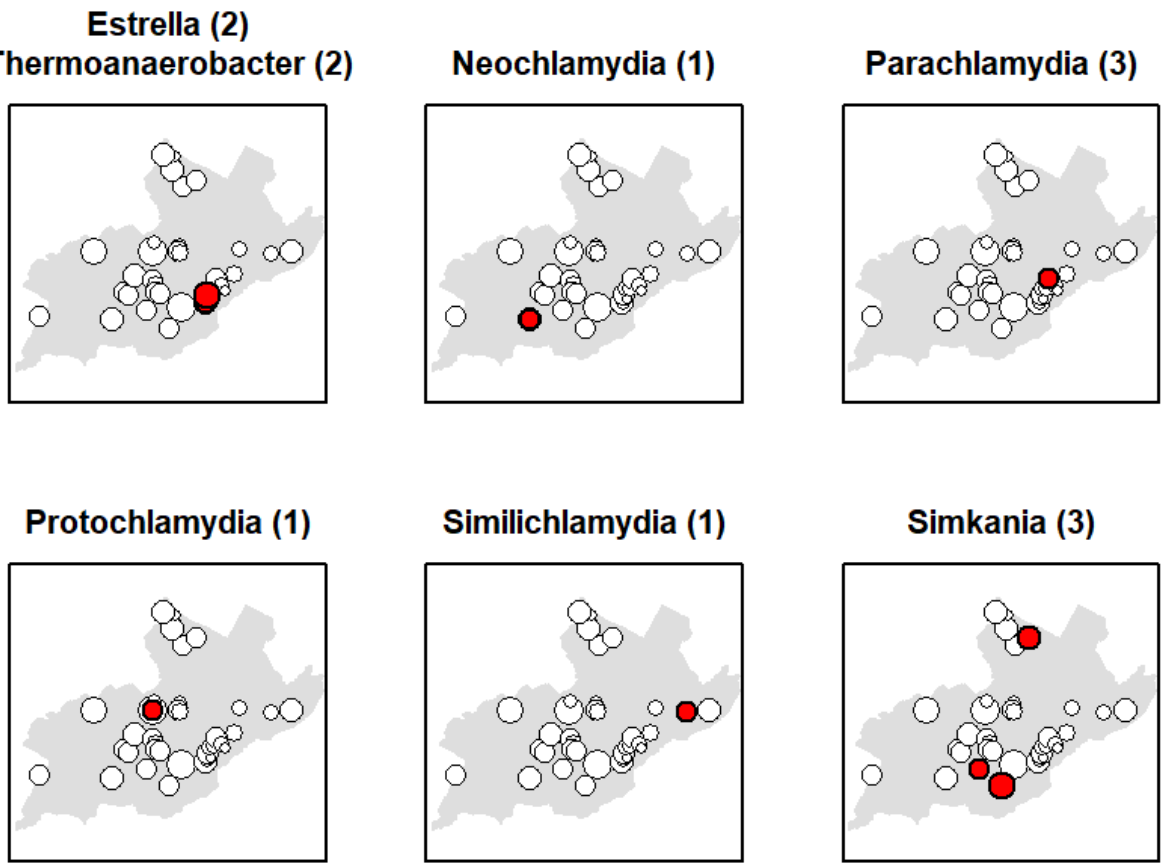


Figure 2. Observed infection rates over the Geneva metropolitan area. The circles represent sampling sites (see Figure 1), with size proportional to the number of tadpoles sampled, and hue intensity following the gradient in the observed infection rates.



457

458 **Figure 3.** Global autocorrelation analysis results. Left column reports the Moran Scatter Plots obtained
459 using the first two, four, six, eight and ten nearest neighbours (i.e. ponds), respectively. Right column
460 reports the Moran's I reference distributions as obtained for each weighting scenario by permutation
461 tests. The red vertical tick highlights the position of the observed Moran's I in the reference
462 distribution; a grey vertical line is drawn to show I=0 (i.e. the null hypothesis). Red horizontal lines
463 pinpoint percentiles 2.5 and 97.5 of the reference distributions, underlining the range of significant I
464 values.



465

466 **Figure 4.** Spatial occurrence of the observed CLOs and *Thermoanaerobacteriaceae* genera is
467 highlighted by red circles. The number of infected samples (i.e. tadpoles) is reported for each
468 bacterial genera in brackets (see Table 1), and refers to the highlighted ponds (e.g. three tadpoles
469 are positive for the genus *Parachlamydia* from the same highlighted pond). The grey area in the
470 background represents the Geneva metropolitan area, and the size of the circles is proportional to
471 the number of tadpoles sampled in each sampling site (see Figure 2).

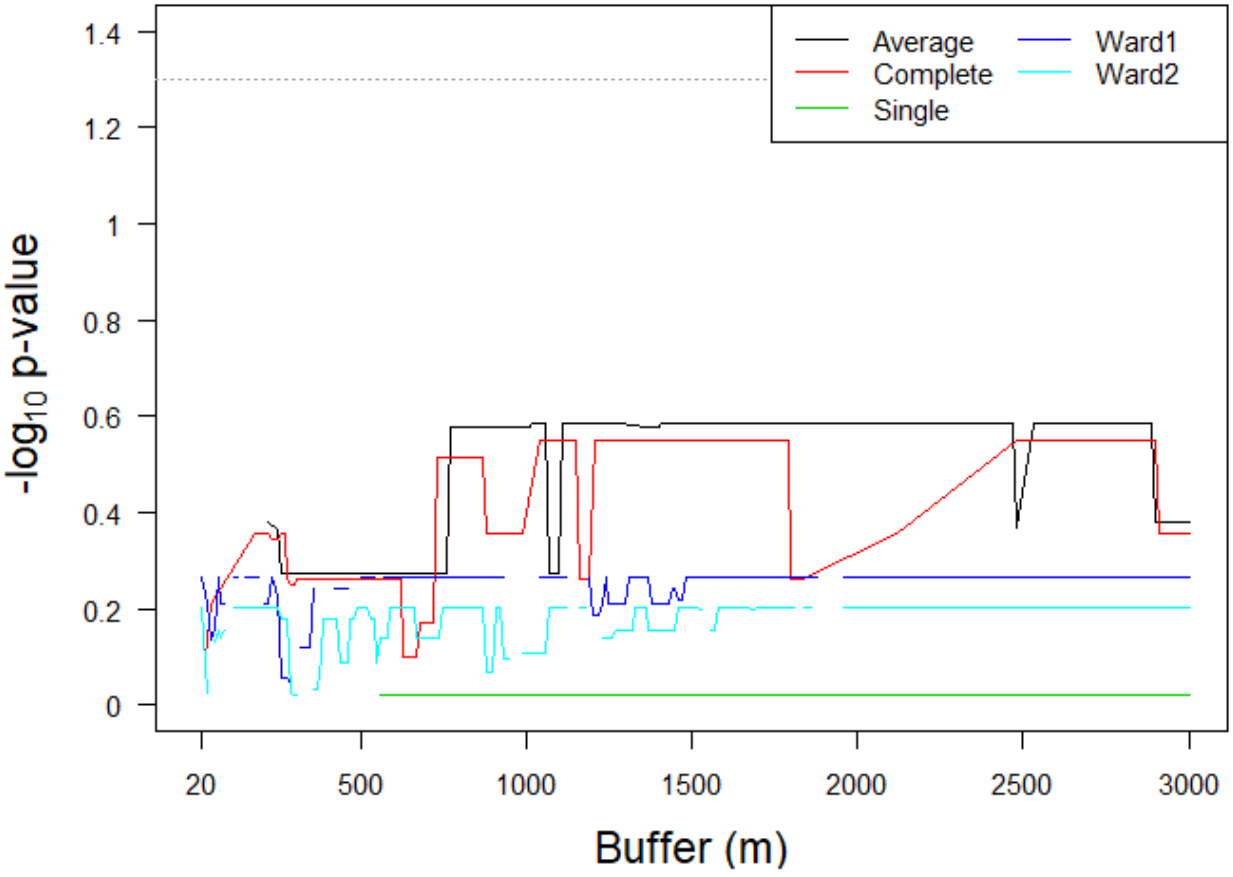
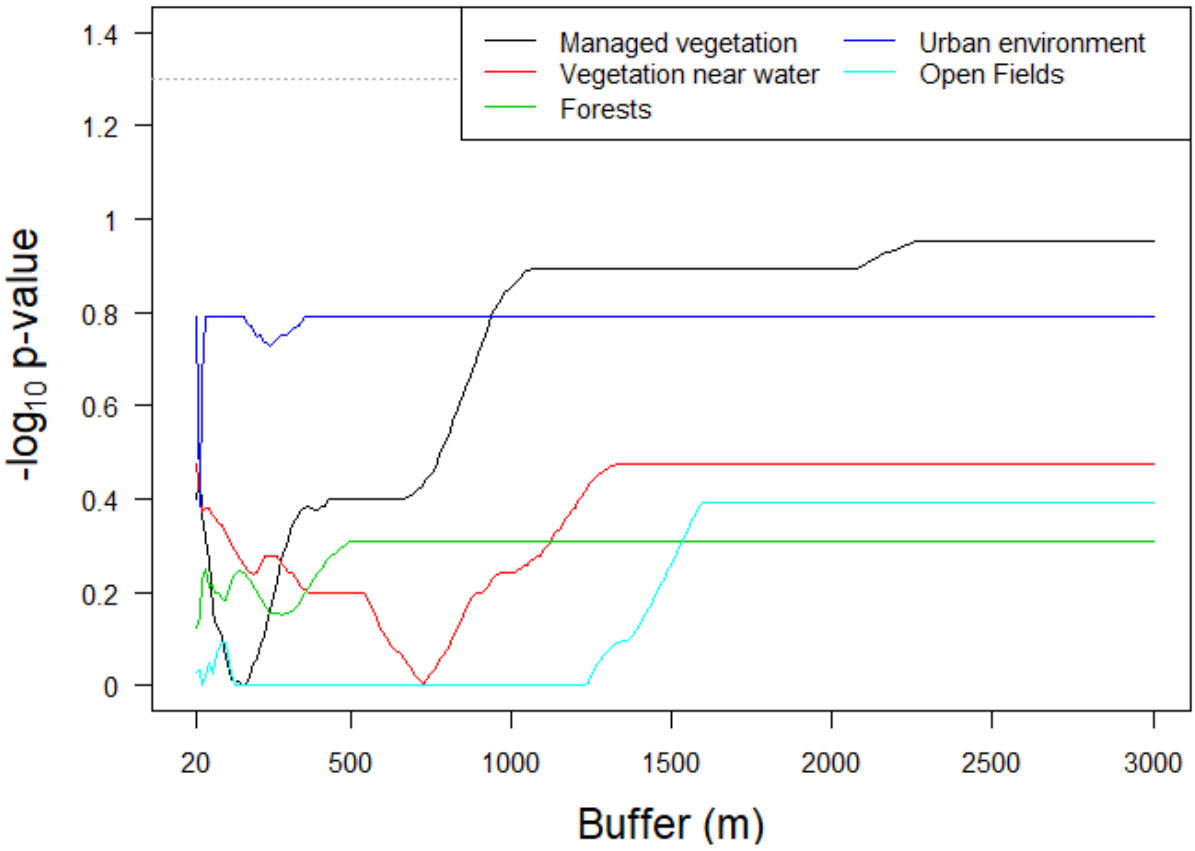


Figure 5. Results of the group comparison tests. *P*-values are reported on the logarithmic scale, after multiple testing correction, and as a function of both the buffer (i.e. radius) used for characterizing land cover around the sampling sites, and the clustering method used to classify ponds into environmental groups. In the uppermost part of the plot, the dotted line indicates the used significance threshold. Line discontinuities depict tests not run (i.e. where at least one group was constituted by a single pond; refer to main text for explanation).



480

481 **Figure 6**. Results of the beta regression analysis. *P*-values associated with the estimated regression
482 coefficients are reported on the logarithmic scale, after multiple testing correction, and as a function of
483 both the buffer (i.e. radius) used for characterizing land cover around the sampling sites, and the
484 aggregated land cover category. In the uppermost part of the plot, the dotted line indicates the used
485 significance threshold.